Emotion in Motion

Rebecca Anne Smith (BSc, MSc)

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

July 2022

School of Psychology

University of Glasgow

62 Hillhead Street

Glasgow

G12 8QB

TABLE OF CONTENTS

ABSTRACT

The central question of this thesis concerns how humans can communicate emotional expressivity through whole-body movement. Previous social communication research has focused mainly on the communicative potential of facial expressions, but until we understand the role of movement on a larger, whole-body scale, our appreciation of social signalling will remain incomplete.

To address this gap, in this thesis I present four experiments designed to examine different elements of the communicative potential of whole-body human movement, using dance as a model paradigm. These experiments explore various aspects of social decision-making when observing human movement through the use of several complementary methodologies including; motion capture, forced-choice emotion recognition tasks, slider scale aesthetic evaluation tasks, and semi-structured interviews.

Overall, *Emotion in Motion* argues for the integrated role of affective and aesthetic processing as a key aspect of evaluating the emotional content of whole-body human movement. In addition, it establishes a role of general and specific kinaesthetic empathy processes, and of previous physical and observational movement experience, in social decision-making for whole-body movement stimuli.

List of Figures

List of Tables

Table S.2	Summary of the methods and results of previous emotion recognition, and stimulus library validation experiments.
Table S.3	Summary of average recognition rates (%) for different types of stimuli reported in the previous emotion recognition tasks (from publications summarised in Supplementary Table S.1).
Table S.4	Information about each individual clip within the full, original McNorm Library.
Table S.5	Summary of the average durations and recognition rates for clips from each emotion category in the McNorm Library.

DEDICATION

I dedicate this thesis to 18-year-old me, who loved to dance and didn't know what to do with her life.

ACKNOWLEDGEMENTS

Compiling a list of people who have helped me get to where I am today would require a whole extra thesis chapter, but there are several people (and places) who have worked behind the scenes to make this thesis happen, and I would like to take the time to mention them here.

To my funders, the Economic and Social Research Council and the Scottish Graduate School for Social Sciences, thank you for providing me with the means to carry out the work contained in this thesis. I would also like to give special thanks to my host institute the University of Glasgow, and to the Department for Psychology and Neuroscience in particular. Glasgow is where my academic career started with a BSc in Psychology, and is the place I have lovingly called home for almost 10 years. A huge thank you to any and all staff members who have supported me throughout my journey, with a special acknowledgement of the kindness of Professor Niamh Stack, and the wisdom of my progression reviewers Professor Rachael Jack and Dr Christoph Scheepers.

I would also like to thank the various cafes and bars in the West End of Glasgow, for giving me a place to work, procrastinate, and celebrate the little victories. In particular I would like to thank the staff of Kelvingrove Café for bringing me cappuccinos before I even asked, for letting me steal their Wi-Fi and electricity supply during the last few months of my write-up, and for being there with a glass of prosecco when I finished this thesis.

I owe my gratitude to a number of academics who have shaped my journey. First, I give my warmest thanks to my supervisor Professor Emily Cross, for being a constant source of enthusiasm, support, and Paesano pizza. It is due to her boundless patience, and encouragement to let go of my perfectionism demons, that I have published anything at all, and her unwavering positivity in the face of unprecedented setbacks that I have been able to submit this thesis. Finally, I thank her for being an incredible role model. As an intelligent, and empathetic woman at the top of her field, she has shown me that anything is possible with enough hard work, passion, and dedication.

Second, I thank Professor Frank Pollick, who has been my supervisor since my undergraduate degree, for encouraging me to pursue my own research interests, and for always being there with words of wisdom, support, and kindness. Without him, I would have never realised that I could blend my passion for dance and psychology in my academic life, and his guidance set me on the path to where I am today.

Next, to Dr Julia Christensen, I fear I'll never be able to convey the extent of my gratitude. I thank you for encouraging me to always "try harder" in the face of adversity, for always doing research that inspires me, for being the kindest soul, and for being a most wonderful friend.

Finally, I also offer a massive thank you to the Sobots Team (in Glasgow, Sydney, and beyond) for their thought-provoking questions, you have all helped me to do better science. In particular, I would like to thank my office mates for letting me cover our (*shared*) walls with headache-inducing mind maps, and for being the most patient soundboards whenever I needed advice. And to Dr. Katie Riddoch and Henry Powell, my most heartfelt thanks for being my ear when I needed it, my brain when mine wasn't working, and my dear friends always.

There are many people, from life outside of academia, to whom I owe an endless debt of gratitude. First, to the guys and gals of Dancemania, thank you for giving me life-long friendships, and perpetuating my love for dance. Second, to members of The Damaged Goods, and The Silver Den; you are Heroes in every sense of the word. Thank you for enriching my life, keeping me sane, bringing joy and laughter to the everyday, and for being my found family.

To Max, my other half. Your proofreading has kept me humble, your cooking has kept me nourished, your patience has kept me sane, and your love has kept me going. I couldn't have asked for a better hand to hold through this process. I only hope that at some point, in the rest of our lives together, I can return the favour.

Finally, to the Smith family, I owe everything to you. To my dad John, thank you for keeping my brain switched on by doing the crossword and pub quiz questions with me every week, for reading endless drafts of my thesis chapters, and for forgiving me for snapping at you whenever I took your feedback too personally. Our phone calls have been an endless source of laughter and love. They are my favourite part of the day, and I couldn't have done this without you.

To my mum Anne Marie, thank you for the coffee/wine money, the hugs, and your unfaltering belief in me. You have shown me the true value of working hard, the payoff of perseverance, and what can be achieved by a strong, badass woman with a dream. You are the most empathetic person I know, and I thank you for being my most vocal cheerleader, and biggest idol.

To my little sister Jessica, you are the brightest ray of sunshine on a cloudy Glaswegian day. Thank you for keeping my use of semi-colons in check, for always letting me rant down the phone at unsociable hours, for making me laugh through tears, and for being so bloody clever that I'm always playing catch up to you. You are amazing.

Research Output

Published Output

Chapter 2 – Smith, R. A., & Cross, E. S. (2022). The McNorm library: creating and validating a new library of emotionally expressive whole body dance movements. *Psychological Research*, 1-25. https://doi.org/10.1007/s00426-022-01669-9

Contribution breakdown – RAS performed data collection and analysis, and wrote the manuscript. ESC and RAS edited the manuscript. ESC approved the final version for publication.

Forthcoming

Smith, R. A., & Pollick, F. E. (forthcoming). The role of dance experience, visual processing strategies, and quantitative movement features in recognition of emotion from whole-body movements. In C. Fernandes, V. Evola & C. Ribeiro (Eds.), Dance Data, Cognition, and Multimodal Communication (Chapter 16). Routledge.

Contribution breakdown – RAS conceptualized and wrote the first version of the manuscript. RAS and FEP reviewed and edited the manuscript drafts prior to submission.

Cross, E. S., & Smith, R. A. (forthcoming). The embodied neuroaesthetics of watching dance. In C. Fernandes, V. Evola & C. Ribeiro (Eds.), Dance Data, Cognition, and Multimodal Communication (Chapter 11). Routledge.

Contribution breakdown – ESC conceptualised the chapter. ESC and RAS wrote the first version of the manuscript, and edited drafts prior to submission. ESC submitted the final version.

Under Review

Chapter 3 – Smith, R. A., & Cross, E. S. (2022). Moving me, moving you: Emotional expressivity, empathy, and prior experience shape whole-body movement preferences. *Under review at Psychology of Aesthetics, Creativity, and the Arts.*

Contribution breakdown – RAS conceptualized and wrote the first version of the manuscript. RAS and ESC critically reviewed and edited the manuscript drafts prior to submission.

In Preparation

Chapter 4 – Smith, R. A., Fernandez, A., & Cross, E. S. (2022). Read Between the Lines: Behavioural and Computational Exploration of Factors Involved in the Emotional and Aesthetic Evaluation of Motion-Generated Continuous Line Drawings.

Contribution breakdown – RAS performed the statistical analysis and wrote the first version of the manuscript. AF performed the computational image analysis. AF and ESC critically reviewed and edited manuscript drafts.

Chapter 5 – Smith, R. A., & Cross, E. S. (2022). "A Little More Conversation": A Qualitative Assessment of Decision-Making by Dancers and Non-Dancers During a Whole-Body Movement Emotion Recognition Task.

Contribution breakdown – RAS wrote the first draft of the manuscript. ESC critically reviewed and edited manuscript drafts.

Contributors

In the following section, contribution summaries are listed for each of the thesis chapters.

Chapter 1 – General Introduction

RAS conceptualized and wrote the first version of this chapter. ESC critically reviewed drafts of the chapter.

Chapter 2

RAS and ESC conceptualized the experiment. EM created and performed movements captured within the McNorm Library. RAS performed data collection and analysis and wrote the first draft of the manuscript. ESC critically reviewed and edited multiple drafts of the manuscript. Two anonymous peer reviewers provided feedback on an earlier draft of the manuscript. ESC approved the final version for publication.

Chapter 3

RAS conceptualized the study idea, analysed the data, and wrote the manuscript. ESC critically reviewed and edited the manuscript.

Chapter 4

RAS conceptualized and wrote the first version of the manuscript. RAS and AF devised the analysis plan. AF performed the computational image analysis; RAS contributed to the computational analysis, conducted data collection and performed the statistical analysis. ESC and AF critically reviewed manuscript drafts.

Chapter 5

RAS conceptualized and wrote the first draft of the manuscript. ESC critically reviewed and edited manuscript drafts.

Chapter 6 – General Discussion

RAS conceptualized and wrote the chapter. ESC critically reviewed the chapter.

CHAPTER 1

GENERAL INTRODUCTION

GENERAL INTRODUCTION

1. Defining Emotion and Emotion-Related Concepts in this Thesis

Before I begin discussion of historical developments in the study of emotion, it may first be useful to establish the definition I will use for *emotion*, and the related concept of *affect*, throughout this thesis.

It should be noted here that controversy surrounding how to operationally define emotions is pervasive throughout much of the literature, and different fields (e.g., Philosophy, Physiology, Neuroscience, Psychology) and sources (e.g., academic papers, dictionary entries) vary widely in both the scope and technicality of definitions offered. I provide a more thorough discussion of the challenges associated with defining emotions in Chapter 5, but I will now briefly clarify the lens through which I will be discussing emotions. As this thesis sits within the discipline of Psychology, I will be using the American Psychological Association's (APA) definition which states that an emotion is a, "complex reaction pattern, involving experiential, behavioural, and physiological elements, by which an individual attempts to deal with a personally significant matter or event".

Throughout this thesis I will also make reference to the *affect*, therefore it is also important to clarify what I mean by this term. A number of researchers in this field use the terms emotion and affect interchangeably, but it is important to note that the predominant account of affect is as more of an umbrella term which can be used to describe a number of valenced phenomena, including emotion, but also encompassing feeling, mood and attitude, for example (Clore & Schnall, 2005). Affect tends to be used to describe phenomena that exist along two fundamental continuums – valence and arousal (including at a neutral point along either, or both of these dimensions), and this circumplex model is most commonly used in empirical work for testing emotion-related words, facial expressions or affective states as stimuli (Kuppens, Tuerlinckx & Russell, 2013; Russell, 1980; Barrett, 2006; Barrett & Bliss-Moreau, 2009; Posner, Russell & Peterson, 2008; Remington, Fabrigar & Visser, 2000). Therefore, this is the approach I will be taking when I discuss affect in this thesis. As I make use of specific emotion-labelling tasks (rather than collecting measures of perceived valence and arousal) throughout the experiments presented in this thesis, I will use the term emotion unless I am making specific reference to the dimensions of valence or arousal, or to the experience of valenced phenomena more broadly.

2. A Historical Perspective on Emotions

Successful interpretation of emotional states plays an essential role in interpersonal communication as it allows us to understand the goals and motivations of others. It enables us to predict the subsequent behaviour of our interaction partners, and prompts us to align our own behaviours to respond appropriately in different social situations (Blair, 2003). Discourse around the concept of emotion can be traced back to the great philosophers of Ancient Greece, and throughout history a number of competing accounts about the nature of the emotional experience have been proposed. Early philosophical accounts, such as those proposed by Plato and Aristotle, describe emotions in relation to aspects of the *soul*, and as a specific kind of subjective experience comprising of various constituent parts (like the pleasure-seeking *epithumētikon*, and the aggressive and self-affirming *thumoeides* in Plato's account, and the *physiological* and *cognitive* components suggested by Aristotle; Knuuttila, 2014; Leighton, 1982; Konstan, 2006). Later, in 1982, Darwin's seminal evolutionary theory of emotion provided some suggestion for the function of emotions (Darwin, 1872/1965). He proposed that emotions are adaptive, and that they serve to aid both survival (where fear, for example, facilitates the avoidance of danger), and the continuation of our species (where feelings of love and affection lead to us finding a mate and procreating). Perhaps as a result of this historical interest in the experience of emotions, a number of influential psychologists and physiologists have sought to develop and refine these theories to suggest mechanisms by which emotions may occur within the human body. Within this framework, there a several key theories to discuss (James-Lange, Cannon-Bard, Schachter and Singer), and all can be illustrated more clearly using the vignette of encountering a bear in the woods.

The James-Lange theory of emotion suggests that emotions occur as a result of physiological responses to a stimulus in our environment (James, 1884; Lange, 1885). According to this perspective, when we encounter a bear in the woods, our bodies produces physiological responses (e.g., sweating, trembling), which prompt us to experience an emotion; in this case, fear. Walter Cannon proposed several criticisms of this theory, as he argued that such physiological responses can occur as a by-product of other situations that do not induce an emotional experience (e.g., sweating or increased heart rate may occur after exercise without producing fear), and that emotional responses may occur too rapidly to result from physiological changes directly. These criticisms, along with contributions from physiologist James Bard, prompted the development of the Cannon-Bard theory of emotion (Cannon, 1927; Bard, 1928). This theory proposes that physiological responses to and cognitive processing (i.e., the creation and manipulation of mental representations of information; Krch, 2011) of a stimulus occur simultaneously, and have distinct contributions to the experience of an emotion. An update on this theory proposed by Schachter and Singer (1962) combines elements from the James-Lange and Cannon-Bard theories and emphasises the influence of both physiological *and* cognitive responses to a stimulus in a two-factor model of emotion. Through their theory, they propose that we experience a physiological response that must be situated within the context of the environment, and subjected to cognitive processing, before an emotion is experienced (Schachter & Singer, 1962). Returning to the bear vignette, according to the two-factor model, we would encounter a bear in the woods and experience physiological responses (e.g., sweating, racing heart), evaluate the situation (i.e., we are in the presence of a wild animal that may be aggressive), and experience an emotion which is appropriate to that situation (i.e., fear). However, if we experience those same physiological responses in a different situation, such as on a first date, we will interpret those signals in a different manner and experience a contextually-relevant emotion (e.g., excitement, sexual arousal).

While each of these theories diverge slightly with respect to the specific mechanisms of emotion production, they all share a number of common features. First, they all emphasise the role of some external stimulus, or trigger, to produce an emotional response. Secondly, they support the idea that emotions occur as a result of both cognitive and physiological processes. Finally, they all suggest that emotions serve some evolutionary purpose; as the vignette illustrates, the function of fear is safeguarding against dangers or threats (which also harks back to the adaptive function of emotions suggested by Darwin, 1872/1965). Therefore, through extensive development and refining of these theories, our understanding of how emotions are produced and experienced within the human body has advanced substantially. However, a substantial gap still exists in our understanding of human emotional experiences. These traditional theories of emotion have focused exclusively on the elicitation or production of emotional responses to a stimulus in our environment. This is problematic, because these theories do not account for the typically *social* nature of emotions and affective communication. This emphasis on the function and mechanism of emotions within the individual, in isolation, represents a substantial limitation of existing emotion frameworks. Until we have developed an equally sophisticated framework for establishing how these emotions can be successfully communicated and interpreted by others in social contexts, our understanding of emotions as an experience will remain incomplete.

3. *Emotion Communication Through Facial Expressions*

Some researchers have attempted to bridge this gap in our understanding by exploring the communicative potential of facial expressions. Faces have received substantial attention in social communication research, and as a result we have identified a number of social cues that individuals can interpret from them. For example, it has been found that we can rapidly attribute personality traits (e.g., dominance, trustworthiness) and judgements of attractiveness, physical health, and intelligence to individuals based on facial cues (Hung et al., 2016; Tan et al., 2018; Todorov & Duchaine, 2008; Mileva et al., 2019; Mignault & Chaudhuri, 2003; Sofer et al., 2015; Santos & Young, 2005). In addition to these factors, a wealth of evidence comes from studies exploring cues observers use to interpret emotional expressivity from small movements of the facial muscles (Ekman & Friesen, 1978; Ekman & Oster, 1979; Adolphs, 2002; Ekman, 1992; Matsumoto et al., 2008). This research has been conducted in participants ranging widely in age, and from a variety of cultural backgrounds, which prompted the development of the Universality Hypothesis (Ekman et al., 1969; Tomkins, 2008; Ekman et al., 1978). This theory proposes that a set of basic emotions (happiness, sadness, anger, fear, surprise, and disgust) exist that produce unique facial expressions, which are recognised consistently across all cultures.

A number of studies have yielded support for this theory, as it has been found that these six expressions are recognised at significantly greater than chance level (see review by Izard, 1994) by individuals around the globe. However, as a result of continued research attention, more recently this theory has been challenged, and a number of criticisms of the methodologies employed to obtain these results have been proposed. For example, several researchers have highlighted inconsistent recognition rates reported for facial expressions related to fear and disgust between Western and non-Western participant samples (Jack et al., 2009; Chan, 1985; Jack, 2013). More specifically, the typical facial movements (called action units) associated with disgust (i.e., wrinkled nose, raising of the upper lip, lowering of the eyebrows) and fear (i.e., raising the eyebrows, widening the eyes, opening of the nostrils) elicit significantly higher levels of recognition among participants from Western cultures, than among participants from non-Western cultures (who often misclassify these expressions as anger and fear, respectively; Jack et al., 2016). To address this, Jack and colleagues conducted a follow-up experiment to explore the recognition of over 60 emotions in participants from the UK (Western sample) and China (non-Western sample). After applying a multivariate reduction technique across the recognition data, these researchers found that only four facial expressions (with unique valence, arousal and dominance dimensions) could be reliably recognised by participants from both cultural backgrounds (Jack et al., 2016). These results were the first to suggest that the widely-accepted Universality Hypothesis model may require further refinement in order to more accurately reflect the idiosyncrasies of emotion communication through facial expressions. Considering this finding more broadly, this relatively recent challenge to the Universality Hypothesis suggests that our understanding of even well-established concepts in the field of social communication need further research attention, and that our comprehension of the intricacies of human social signalling is, currently, incomplete.

If this is the case for faces, then this issue possibly exists to a greater extent for social cues that have received less empirical attention. To date, the role of whole-body human movement has been somewhat neglected in social perception research; particularly in comparison to other kinds of social communication tools (e.g., facial expressions, voices, language), and to my knowledge, there are no dominant or well-established theories about the unique mechanisms that facilitate recognition or transmission of our internal states through movement of the human body. A number of papers, however, have explored disparate aspects of the communicative potential of movement, and the individual differences that can facilitate or impair the communication of emotional expressivity. While no complete framework for the role of body movement in social communication exists to date, findings from these studies (and related theories from other research domains), reviewed below, should be useful to inform the development of such a framework.

4. Previous Movement Research

Human movement provides a rich source of information which is of value to social observers, and, in fact, recent research has suggested that the body may actually provide a more salient source of social information than the face (Aviezer et al., 2012; Wang et al., 2017). It has been suggested that human bodies reflexively capture our attention, because they move in a biological fashion (Fox & McDaniel, 1982). Biological motion perception refers to the ability to recognise the characteristic movement patterns of organic entities, including humans (Voos et al., 2013). Research points to a wide and distributed brain network dedicated to this fundamental perceptual ability – including a series of cortical structures (posterior temporal sulcus, inferior frontal gyrus, extrastriate and fusiform body areas) and subcortical brain areas (including the ventral lateral nucleus, a portion of the thalamus that projects to motor and premotor cortices; Grosbras et al., 2012; Chang et al., 2018). Given the range of associated cognitive structures, it is not unreasonable to conclude that biological motion perception must confer some benefit beyond recognition of others. Indeed, research suggests that the salience of biological motion enables us to infer a variety of high-level social cues from other bodies in motion. This has been observed across a variety of kinds of movement (e.g., everyday actions like walking and knocking motions, dance sequences), and across various stimulus presentation methods (e.g., live, real-time movement, full light display videos). However, even without immediately apparent morphological information, research has shown that humans are able to identify a range of features from Point-Light Display (PLD) motion. For example, it has been found that people are highly accurate at identifying the gender (up to 71% accuracy; see review by Pollick et al., 2005) and personal identity (Mitchell & Curry, 2016) of PLD walkers. In addition, it has been found that individuals can reliably identify personality traits and the affective states through PLD gait representations (Gunns et al., 2002; Heberlein et al., 2004; Schneider et al., 2014). (Smith & Pollick, forthcoming; Cross & Smith, forthcoming).

While some evidence suggests these social decisions can be made rapidly, and often with high levels of reliability (e.g., Pollick et al., 2005; Gunns et al, 2002; Heberlein et al., 2004; Schneider et al., 2014), it remains unclear how exactly these judgements occur, but examining techniques employed in face research may provide some insights. For example, it has already been mentioned that, in theory, specific action units are involved in the production of distinct facial expressions (Ekman & Friesen, 1978; Ekman & Oster, 1979; Adolphs, 2002; Matsumoto et al., 2008). A number of researchers have applied a similar principle in the study on human movement, by examining whether there are unique profiles of movement kinematics (at the whole-body level, rather than for only specific muscles in the face) that are associated with particular emotional states. There are several ways in which we can extract this kinematic information from human motion, some of which employ computational methods (applied to 3-dimensional movement data), and some which are based on subjective perceptions of naïve observers, or specifically-trained movement experts (e.g., Laban Movement Analysts). So far, studies employing both of these methods have produced mixed results in terms of which characteristics of movement are associated with the expression of different emotional states.

From movement data, it is possible to compute features such as velocity and acceleration profiles, fluidity, contraction, and quantity of motion from coordinate data obtained in motion capture. A number of researchers with backgrounds in both computing science and psychology have employed such methods in the exploration of which characteristics of motion contribute to the recognition of emotional expressivity. For example, Roether and colleagues collected motion capture data from 25 individuals who performed non-expressive walking motions, and walking movements imbued with happiness, sadness, anger, and fear. They compared the average flexion angles, and spatio-temporal structures of the joint trajectories, and identified unique kinematic profiles which resulted from the communication of each of these emotion categories (Roether et al., 2009). They found that happy walking movements were associated with faster gait cycles, and a higher degree of movement activity. Similarly, angry walking movements were also faster and contained a higher degree of movement activity than sadness or fear. Angry movements also contained more angular arrangements of the body segments compared to happy movements. Sad walking movements, on the other hand, were relatively slower with lower levels of movement activity, and involved a downward angling of the head (i.e., tilted downward toward the chest). Fear, too, was associated with slower gaits and lower movement activity, but was also found to result in reduced motion of joints in the lower body (i.e., the knees and hips) (Roether et al., 2009).

Alternatively, some studies have explored movement dynamics from a perceptual standpoint. Laban Movement Analysis (LMA), for example, is a tool for describing, visualising (and ultimately quantifying) kinematic and expressive features of human movement. LMA is often employed in the context of movement training (e.g., dance, theatre, sports), and by health practitioners (e.g., physiotherapists, occupational therapists; to diagnose dysfunction or correct imbalances in body alignment), but it has also been used in human movement research. There are four dimensions along which movement can be classified, these are *body* (physical characteristics of the dynamic human body), *effort* (dynamics of human motion), *shape* (a deeper examination of how the body changes during movement), and *space* (relationship between movement and the environmental space which it occupies). The effort dimension has probably received the most attention in this area of research, and is itself comprised of the following subcategories; *space* (a measure of movement direction, ranging from direct to indirect), *weight* (a measure of the heaviness of movement, ranging from strong to light), *time* (a measure of the speed of movement, ranging from sudden to sustained), and *flow* (a measure of movement fluidity, ranging from bound to free). It is likely that effort is the most often-used LMA category because these features (i.e., space, weight, time, flow) can be mapped more clearly onto computational measures of movement kinematics.

Experiments using both subjective annotation (e.g., LMA) and computational analysis to identify kinematic characteristics of emotionally expressive movement have reported sometimes contradictory results. For example, sadness has been associated with both direct (Sawada et al., 2003) and indirect (Wallbott, 1998) movement trajectories, happiness has been associated with both relaxed (Gross et al., 2010) and high energy (Brownlow et al., 1997) motion, and these opposing profiles are even more pronounced when examining more complex emotional states (e.g., hot anger, grief, anxiety; Wallbott, 1998; Camurri et al., 2002; Gross et al., 2010). A more detailed overview of the previous results in this area can be found in Supplementary Table S.1.

The only parameter with an apparently consistent influence on emotion perceptions is speed of motion (Smith & Pollick, in preparation). Faster movements tend to produce higher intensity emotions (such as anger, and joy), while slower movements appear to influence the perception of lower intensity emotions (e.g., sadness). However, if speed is the only consistent determining factor in the emotions we assign to human movement, then these results suggest that observers would not be able to distinguish between fury and elated joy, nor between sadness and anxiety. We know that this is not how emotion communication occurs in real-world social situations. As such, it would appear that other factors must also be involved in our ability to communicate non-verbal social signals to others in our social world. If kinematics alone do not provide a comprehensive explanation for the reasons we assign emotion judgments to human motion, it may be useful to also examine the factors that can either impede or facilitate our ability to assign appropriate emotional states to the movement of others.

5. *Individual Differences Impeding Emotion Recognition: Mood Disorders and Autism Spectrum Disorder*

Research suggests that neurodevelopmental and mood disorders can influence our ability to successfully recognise and respond to the emotions being expressed by others more generally, and specifically through movement of their body. For example, in a meta-analysis conducted by Demenescu and colleagues (2010), they explored results from 10 studies exploring emotion recognition capabilities of individuals with major depression, and anxiety disorders. They observed that adults with anxiety disorders provided less accurate judgements in emotion recognition tasks, and this effect was even more pronounced among adults with major depression. However, children with anxiety disorders did not experience such severe deficits in emotion recognition capabilities. The authors concluded that the specific deficit in emotion recognition observed in adults with depression may reflect a “lack of attentiveness to others” (Demenescu et al., 2010). Considering this conclusion alongside evidence which suggests a negative relationship between depression and perspective-taking (Wegemer, 2020), these results suggest that perhaps the ability to attune to the internal states of others (via perspective taking or empathy, for example) plays a key role in the success of emotion recognition. Therefore, it may be useful for any framework describing the communication of emotional expressivity to account for individual differences empathy.

In addition, it has been found that individuals with ASD experience difficulties in the perception of biological motion (Mandy et al., 2012). In a meta-analysis conducted by Todorova and colleagues, results from 52 papers revealed that individuals with ASD showed a decreased performance in tasks involving the perception and interpretation of biological motion, and that this difference between individuals with and without ASD was more pronounced in younger children (Todorova et al., 2019). Further research suggests that individuals with ASD recognise emotions with significantly lower accuracy than typically developing (TD) individuals, and this finding has been observed across a range of sensory modalities. For example, Yeung (2021) recently conducted a systematic review and meta-analysis of emotion recognition from facial expressions in individuals with ASD, in comparison to individuals without ASD, and those with other clinical conditions. The results of 148 studies revealed that individuals with ASD showed significant impairments in recognising all basic facial emotions, and that these impairments were more severe in individuals with ASD than in individuals in other clinical populations (including schizophrenia, ADHD, dyslexia, and anxiety). Results from these studies also suggest that impairments are more pronounced when emotion recognition tasks involve more complex emotions, and holistic processing of facial stimuli (Yeung, 2021).

In the case of emotion recognition from body movements, similar patterns of impairment have been observed for recognition of target emotion categories. In a study conducted by Metcalfe and colleagues, children with and without ASD observed avatars depicting eight emotional states (anger, boredom, disgust, fear, happiness, sadness, surprise, and worry) through movement of the body, which sometimes included emotionally-salient gestures. They found that the children with an ASD provided significantly lower levels of emotion recognition than the TD children (Metcalfe et al., 2019). Similar results have been obtained from a number of studies exploring emotion recognition capabilities of adults with ASD (Philip et al., 2010; Kothari et al., 2013; Alaerts et al., 2014). Results from these experiments suggest that ASD produces specific deficits in the recognition and interpretation of emotional expressivity (across various modalities), and results of the meta-analysis conducted by Yeung (2021) suggest that, at least for faces, ASD produces more pronounced impairment in tasks using more complex emotions, and tasks which involve the holistic processing of emotionally-salient stimuli.

In review, if there are general impairments in emotion recognition in individuals with an ASD, and more specific impairments with holistic processing of more complex emotions, this may suggest that these facets of human motion are particularly salient in social communication. Therefore, it may be useful to focus on emotion recognition from whole body movement sequences (rather than isolated movements of the arms – e.g., Sawada et al., 2003; or hands – e.g., Tripathi et al., 2018), and to explore differences in the perception of more complex emotions (rather than simply positive versus negative valence affective states – e.g., DeMeijer, 1989) when developing a framework for the role of human movement in communicating emotional expressivity.

6. *Individual Differences Facilitating Emotion Recognition: Empathy and Previous Experience*

In addition to the individual differences found to impede emotion recognition capabilities, researchers have obtained evidence to suggest that possessing certain traits can facilitate these skills. Research has shown that individuals with higher levels of trait empathy (as measured by, for example, the Toronto Empathy Scale – Spreng et al., 2009; the Emotion Specific Empathy Questionnaire – Olderbak et al., 2014; the EmQue – Rieffe et al., 2010) provide more accurate judgements of emotion than individuals reporting lower levels of empathy. This result has been observed across a variety of social cues, including facial expressions (Besel & Yuille, 2010; Konrath et al., 2014; Balconi & Bortolotti, 2012; Holland et al., 2021; Jospe et al., 2018) and to a lesser extent in voices (Tsiourti et al., 2019; Golan et al., 2007; Kleinman et al., 2001; Rutherford et al., 2002). Empathy has also been suggested to play a role in responding to movement of the human body.

The human brain has evolved specialised networks for processing human action and motion. One of these networks has been termed the human mirror neuron circuit (Rizzolatti & Craighero, 2004). "Mirror neurons" is the name given to cells in the primate brain that are activated when performing an action, as well as when observing another individual perform that same action. These individual cells are part of a larger network that encompasses a wide range of cortical structures. These cells were first identified by a team of researchers from the University of Parma in 1992, led by Italian neurophysiologist Giacomo Rizzolatti. These researchers conducted studies to explore activation of ventral premotor cortex (vPMc) neurons when Macaque monkeys perform hand- and mouth-related actions (e.g., grasping, placing food in the mouth). They observed activity in this area when the monkeys executed the grasping actions, but also, curiously, found that the same neurons fired when observing the experimenter's demonstration of the same grasping actions (Di Pellegrino et al., 1992; Rizzolatti et al., 1996). They concluded that these cells in the vPMc compose a network dedicated to both action-execution and action-observation for familiar movements. This discovery formed the basis for future studies that have identified a system with similar functionality within the human brain (Keysers & Gazzola, 2006; Caspers et al., 2010; Grezes & Decety, 2001; Mazurek et al., 2018; Press et al., 2012). This notion was later extended by Rizzolatti and colleagues, who suggest that "mirror neurons" may present a neurological approach for the study of empathy. In review of a number of works (e.g., Phillips et al., 1997; Adolphs, 2003; Damasio, 2003; Calder et al., 2000), Rizzolatti and Craighero suggest that direct mapping mechanisms are involved in our understanding of emotions being expressed by others; more specifically, that we perceive and interpret emotions successfully from others because those same emotions are activated within ourselves (Rizzolatti & Craighero, 2005). Therefore, in this regard, it seems plausible that empathy might play a key role in emotion recognition capabilities, particularly when these emotions are being communicated through the body in motion, because we simulate the expressivity we see in others within our own bodies to understand their internal states.

While the question of whether an equivalent "mirror neuron" circuit as that identified through neurophysiology work in the non-human primate brain also exists in the human brain, and whether their functions are absolutely comparable, may be a contentious topic within social cognition research (Corradini & Antonietti, 2013; Jacob, 2008; Tramacere et al., 2017; Spaulding, 2013; Kosonogov, 2012; Sinigaglia, 2008; Holmes, 2013; Corradini & Antonietti, 2013) the theory originally proposed by Rizzolatti and colleagues served as a basis for the development of more recent theories (e.g., simulation theory), and a growing body of evidence supports the role of the Action Observation Network (AON) in the social perception of human movement instead. Researchers who favour the latter terminology note that the AON refers to a broader circuit of structures involved in the processing of observable actions more generally, including but not exclusive to those that are believed to exist within the human "mirror neuron" system (i.e., parietal and premotor areas responsible for the observation and execution of movement, more specifically) (Cross et al., 2009). Despite the often contradictory discourse and terminology used to describe the neural correlates of observing and responding to human movement, these various perspectives suggest that empathic processes may play a specific role, at least from a neurocognitive perspective, in the processing of human motion.

Finally, this link between empathy and emotion recognition capabilities is clear when returning to the discussion of ASD. Individuals with an ASD tend to provide lower self-reported levels of empathy, and typically show poorer performance on empathy-related tasks compared to typically developed individuals (Bishop & Seltzer, 2012; Demurie et al., 2011; Kok et al., 2016; Pepper et al., 2019; Trimmer et al., 2017). Furthermore, a common characteristic of ASD is difficulty in interpreting social cues, including the expression of emotions. It has already been suggested that the specific impairments experienced by individuals with ASD in emotion recognition tasks may provide some insight into the processes which facilitate the social perception of human movement (Yeung, 2021), but this evidence suggests that trait empathy may be of particular relevance to this question.

Considering the findings from research explicitly examining the relationship between empathy and affective processing, and indirect evidence obtained from the discussion of "mirror neurons", and the differences in emotion recognition capabilities among individuals with and without an ASD, it seems likely that individual differences in empathy play a central role in the successful interpretation of social cues. Therefore, in generating a comprehensive framework for social signalling, it may be crucial to include some evaluation of the role played by empathy.

Another important factor to discuss in relation to emotion recognition from human movement is the role of previous experience. The majority of studies exploring the role of prior experience in emotion recognition capabilities has made use of expert dancer participants, and a wealth of evidence suggests an intimate relationship between motor expertise and the successful interpretation of bodily cues of affect. In one study, Christensen and colleagues explored how substantial dance experience influences sensitivity to recognising the affective expression present in the movements of other dancers (Christensen et al., 2016). In this study, 24 non-dancers (novices) and 19 professional ballet dancers were asked to rate a series of short (5–6 sec) movement sequences in terms of their affective valence on a scale from 0 ("very sad") to 100 ("very happy"). Each of these clips had previously been assigned a valence category (based on results from an earlier stimulus validation study), and an equal number of "sad" and "happy" clips were presented to participants. They found that the professional dancers were significantly more accurate in their emotion judgements for happy movements than the novices. Similar results were obtained in a later study, wherein participants from a variety of dance backgrounds were asked to observe a series of movement sequences and identify whether these movements were expressive or non-expressive (Christensen et al., 2019). In this experiment, years of dance experience was found to correlate with accuracy in affective recognition, providing further support for the idea that dance experience facilitates the ability to recognise expression from emotion whole-body movement displays. Given the wealth of evidence suggesting that previous motor experience can prompt improved performance in emotion recognition tasks, it would be useful for any framework of whole-body social signalling to account for the motor expertise in mediating successful social communication.

7. The Potential Role of Aesthetics?

Considering what we know about how emotions are produced (through a combination of physiological responses and cognitive processing) from traditional emotion theories, and the role of certain individual differences that can mediate the success of communicating and recognising affective states (e.g., higher trait empathy, previous experience), it is worth noting that another kind of social judgement bears striking similarities in both of these regards. Aesthetic judgements emerge through similar combinations of physiological and cognitive processes, and are also mediated by individual differences in empathy and previous experience.

Before beginning a more thorough discussion of the similarities between affective and aesthetic decision-making, it would be useful to establish what I mean by *aesthetics* within the context of this thesis. It should be noted that many definitions for aesthetics are problematic in that the terminology used often isolates one particular facet of aesthetic evaluation (usually beauty or pleasantness) and/or one particular sensory experience (most often sight), rather than providing a more general description of the term. For example, the first entry in the Merriam Webster dictionary defines aesthetics as a, "branch of philosophy dealing with the nature of beauty, art, and taste and with the creation and appreciation of beauty" (Merriam-Webster(a), n.d.). After evaluating these issues within many definitions put forth by a number of different sources, I have decided to generate my own definition for what I mean when I discuss aesthetics within this thesis. Therefore, I ultimately define aesthetics as, "a set of principles concerned with the nature and appreciation of artistic taste". Now that this term has been defined, I will return to discussing the similarities between aesthetic and affective processing.

The similarities between these two kinds of social processing can perhaps be understood more clearly through two thought experiments. Returning to the typical vignette used to describe the experience of emotions, regardless of whether you ascribe to the James-Lange, Cannon-Bard, or Schachter and Singer theory of emotion, in this scenario you encounter a bear in the woods (which serves as an input stimulus) which you process cognitively (growling, for example, conjures an association with animal aggression which you understand represents a threat) and react to physiologically (you begin sweating and trembling), which results in you experiencing an emotional state (i.e., fear). When considering aesthetic responses, I would like to propose the following vignette: you are visiting an art gallery and see a painting (in this case, the input stimulus) to which you react physiologically (you may experience "aesthetic chills", and your pupils may dilate), and process cognitively (it may have symmetry, for example, which you likely consider to be aesthetically pleasing), and this produces a positive aesthetic experience (e.g., "I like this"). A comparison of the processes involved in the elicitation of emotional and aesthetic experiences can be observed in Figure 1.1.

Indeed, there is a significant volume of evidence to support the role of cognitive processing and physiological responsiveness in the appreciation of aesthetically-relevant stimuli. For example, it has been found that individuals cognitively process the aesthetic value of visual artworks based on factors such as symmetry (Weichselbaum et al., 2018; Bertamini & Makin, 2014), subject matter (Kettlewell et al., 1990; Graham et al., 2010; Leder et al., 2013), art style (Mastandrea et al., 2011; Furnham & Walker, 2001; Van Paasschen et al., 2015), and visibility of brushstrokes (Leder et al., 2012; Ticini et al., 2014). For music, listeners cognitively process various auditory factors such as tone, harmony, timbre and timing (Hevner, 1935; Weld, 1912; Vernon, 1930; Müller et al., 2010; Valentine, 1962; Marin et al., 2016) of an excerpt before providing an aesthetic value judgement for a composition, and in dance factors such as symmetry (Orgs et al., 2013; Calvo-Merino et al., 2008), familiarity (Cross et al., 2011; Torrents et al., 2013; Orgs et al., 2016; Kirsch et al., 2013; Ticini et al., 2015), and synchrony (Orgs et al., 2016; Vicary et al., 2017) can influence how beautiful we find dance to be, and how much we like certain choreographic sequences. In terms of physiological responses, a number of studies have found during the process of evaluating stimuli in an aesthetic context, differences can occur in heart rate (Yalta et al., 2016; Tschacter et al., 2012; Neale et al., 2021), galvanic skin responses (Tschacter et al., 2012; Neale et al., 2021; Colver & El-Alayli, 2016; Farley & Dowling, 1972), visual fixation patterns (Khachatryan et al., 2020; Xu et al., 2020; Locher, 2006), and activity of facial and body musculature (Fiori et al., 2020; Jola et al., 2012; Jola & Grosbras, 2013; Michaelis et al., 2014). Therefore, there is a wealth of evidence to support the idea that the processes involved in the elicitation of an emotion response may also be involved in the production of an aesthetic experience.

Given these clear similarities in the elicitation and perceptual processes, in addition to the evidence suggesting their potential co-activation, this may merit the inclusion of aesthetic factors in any theoretical framework of affective social signalling. A further evaluation of these similarities, and evidence for their potential overlap in social processing, is provided in the following section.

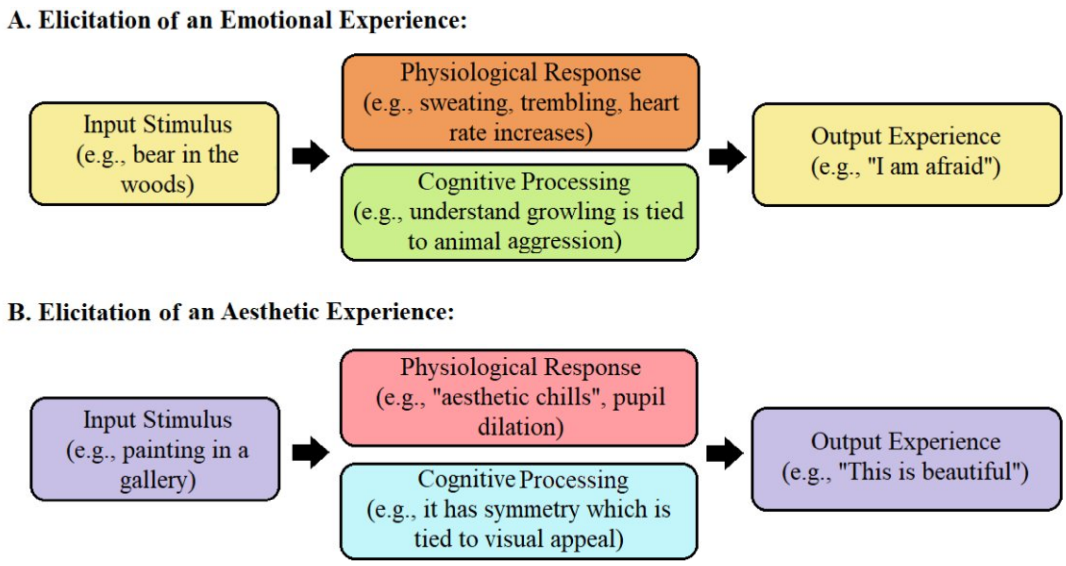


Figure 1.1: Flowchart A shows a visualisation of the processes involved in the elicitation of an emotional experience (based on general emotion theories; James-Lange, Cannon-Bard, Schachter & Singer), illustrated by the vignette of encountering a bear in the woods. Flowchart B shows a visualisation of the processes proposed to be involved in the elicitation of an aesthetic experience (based on empirical aesthetics and neuroaesthetics research), which is illustrated by the example of observing a painting in an art gallery.

8. The Relationship Between Affect and Aesthetics

Historical perspectives on emotion have already been discussed, but interest in the concept of aesthetics, and debate about its very definition (as previously noted), can also be traced back to the early Greek philosophers (e.g., Plato, Aristotle), and throughout history a number of competing accounts about the nature of art, beauty, and the aesthetic experience have been proposed. A long-held view in philosophy is that art appreciation is rooted in hedonism (i.e., the pursuit of pleasurable experiences and emotions). Some theorists support the notion that two distinct kinds of aesthetic hedonism exist; pure and mixed pleasure (De Clercq, 2019; Massin, 2018). Pure pleasure refers to the kind of pleasure which results from purely positive experiences, whereas mixed pleasure refers to the kind of pleasure that can be obtained through the experience of negative feelings (i.e., our love of the great tragedies, appreciation of sad music). I argue that this distinction between pure and mixed aesthetic hedonism, which necessitates affect, suggests an intimate link between affect and aesthetics.

Beyond philosophical accounts, a considerable amount of behavioural and neuroimaging evidence suggests that affective and aesthetic processes are intrinsically linked. A number of studies have been conducted across a variety of artistic mediums, including visual artworks (e.g., paintings, photographs), music and the performing arts that provide support for the relationship between affective and aesthetic perception. In one such experiment, participants provided responses on a measure of 'emotional contagion', which captures an individual's ability to "feel into" the emotions of others, before evaluating a series of representational and abstract artworks while facial electromyography and skin conductance data were collected (Gernot et al., 2018). Participants who reported experiencing greater emotional contagion demonstrated greater physiological reactions, and provided more positive aesthetic appraisals of the artworks (i.e., "being moved", valence, and interest). These results provide evidence for the notion that physiological responses are involved in the aesthetic experience as they are in emotional experiences, and adds some empirical weight to my idea that these processes may involve similar elicitation processes.

Additional evidence for the close relationship between affective and aesthetic processing can be found when examining similarities in the individual differences which can mediate these judgements. It has already been noted that individuals with high levels of empathy perform significantly better in emotion recognition tasks than those with lower levels of trait empathy (including individuals with ASD; Young, 2021; Metcalfe et al., 2019; Philip et al., 2010). However, individual differences in empathy have also been found the influence the aesthetic value judgements people assign to artworks, across a variety of artistic media. For example, in response to the observation of paintings, Wilkinson and colleagues (2021) observed that individuals recording higher empathy scores were significantly faster to respond to the artworks than those with lower scores. Additionally, these authors found that individuals scoring higher on empathy measures showed higher aesthetic interest and emotional engagement in response to the more emotionally evocative artworks (Wilkinson et al., 2021). Similar results have reported from neuroscientific experiments of aesthetic processing. In one study, participants observed a series of images which had been classified as either beautiful and sad, beautiful and happy, neutral, or ugly (Ishizu & Zeki, 2017). Neuroimaging data was collected across a series of brain regions, but with a particular focus on activity of the medial orbito-frontal cortex (mOFC), a brain region believed to play a key role in the experience of beauty. Greater activity of the mOFC was observed in response to both types of beautiful stimuli (sad and happy), than in response to the neutral or ugly stimuli. Through functional connectivity analyses, the authors further found that this activity was mediated by activation of the supplementary motor area and middle cingulate cortex; areas which are known to be involved in experiencing empathy when others are experiencing sadness (Ishizu & Zeki, 2017). Results from such experiments suggest that individuals with higher trait empathy may process visual artworks more efficiently and provide more positive appraisals of them, but they also suggest that emotional processing may be a key component of this process.

While the majority of empirical aesthetics and neuroaesthetics research in this area has focused on responses to visual artworks, some evidence also exists that suggests this effect is not limited to the processing of paintings or static visual artworks. A number of researchers have explored the relationship between empathy and aesthetic processing in response to musical compositions. For example, Vuoskoski and colleagues observed that participants with higher self-reported empathy scores provided higher liking scores and reported experiencing more intense emotional responses to sad music than individuals with lower trait empathy scores (Vuoskoski et al., 2011). A growing body of evidence suggests that general and domain-specific empathetic skills are involved in the processing of human movement stimuli. Calvo-Merino and colleagues (2008), for example, were the first to obtain neuroscientific evidence for the relationship between kinaesthetic empathy and aesthetic processing. In this experiment, participants observed a series of whole-body dance movements while fMRI data were collected, and in a follow-up session, the same participants were asked to rate each of the dance clips in terms of interest and liking. They found that the clips which (as a consensus) received more positive liking scores produced greater activation of the premotor cortex than clips with lower liking scores. The authors suggest these results indicate an inherent link between kinaesthetic empathy (as evident from activation of the motor areas) and aesthetic evaluation of the human body in motion (Calvo-Merino et al., 2008).

The work carried out by Calvo-Merino and colleagues (2008) was conducted with dance-naïve participants, but concepts related to kinaesthetic empathy (i.e., movement simulation, mirroring, embodiment) are often explored within participants who have received some form of dance training. This, therefore, points to a relationship between previous experience and aesthetic evaluation of human movement stimuli. Indeed, there is a wealth of evidence to support the notion that movement experience produces greater aesthetic appreciation of whole-body movements (Cross et al., 2011; Kirsch et al., 2013; Ticini et al., 2015; Vicary et al., 2017). It has been reported that even short-term dance training can sculpt the aesthetic perception of dance sequences. In one such study (Kirsch et al., 2015), participants were invited to take part in a 4-day dance training session. Each day, the participants physically rehearsed one set of movement sequences, passively observed another set of movements, and listened to the music accompanying a third set of movements; a fourth set of movements remained untrained to allow for comparison. fMRI data were collected prior to and immediately following the training sessions, along with measures of movement capability and aesthetic value judgements. The participants liked the movements they had both physical and observational experience with significantly more than the music-only sequences and the untrained sequences. In addition, neurocognitive changes were observed following this dance training. Prior to the sessions, brain activity associated with aesthetic responses was centred around the dopaminergic reward system structures, such as the nucleus accumbens. However, after the training sessions, greater activity was observed in areas of the action observation network (AON), in particular, areas believed to be involved in multisensory integration and the processing of biological motion (pSTS and STG; Kirsch et al., 2015). These results suggest that even short-term dance training can result in changes to subjective aesthetic value judgements, and can lead to reorganisation of the corresponding functional brain responses associated with aesthetic evaluation of dance movements.

Considering the role of dance experience in aesthetic appraisal of human movement, and the aforementioned role this previous experience plays in mediating emotion recognition capabilities, this provides further support for the idea that affective and aesthetic processes are intrinsically linked. However, to my knowledge, no model of social communication has accounted for the role of both emotions and aesthetics in the perception of socially-salient stimuli, and certainly not within the context of human movement.

9. Summary of Related Work

In review of what has been presented above, a number of general emotion theories have been proposed (e.g., James-Lange, Cannon-Bard, Schachter and Singer), but these theories account for only the production of emotions within an individual, and do not provide any explanation for the social nature of emotional communication. Only in facial expression research has there been any attempt to consolidate our understanding of how emotions are communicated to, and interpreted by, others in our social world. However, even a relatively well-established theory in face research (Universality Hypothesis) has been the subject of recent valid criticism, thus highlighting the need for continued research in the field of social communication. As the role of the dynamic human body in social signalling has been somewhat neglected in this area of study, so far, no theoretical framework has been proposed or subjected to rigorous empirical evaluation. Until such a framework is developed, we cannot hope to gain a comprehensive understanding of the intricacies of human social communication.

A number of research works have explored disparate elements of the role human motion plays in social communication, and the individual differences which can sculpt the perception of movement. Empathy is one such trait that can influence the success with which we can interpret affective cues in others, and for recognition of these social cues from movement more specifically, the role of previous motor experience has been emphasised in much of the existing literature. A fairly well-established relationship between affective and aesthetic processing of the dynamic human body also exists. However, to my knowledge, no attempts have been made to examine these factors simultaneously, nor to suggest how these factors interact in a unified model of the social communicative potential of movement.

10. Using Dance to Explore Social Perception

I believe that dance poses a uniquely useful lens through which to explore these social judgements, for a number of reasons. Primarily, I believe that if affect and aesthetics are both at play in the social evaluation of human movement, as previous research would suggest, then movement stimuli used to jointly evaluate these factors should contain both affective and aesthetic components. Dance is, at its heart, the purposeful performance of expressive whole-body movements designed to communicate a narrative or meaning to observers, and while I am not the first to suggest the value of dance in social communication research (Orgs et al., 2018; Van Dyck et al., 2012; Aristidou et al., 2017; Van Dyck et al., 2017; Christensen et al., 2016; Christensen et al., 2019), I believe dance represents an ideal tool to explore how the dynamic body can transmit both affective and aesthetic information to individuals in our social world.

11. General Aim of This Thesis

Through this thesis, my aim is to address a major gap in the literature by exploring the combined influence of emotional and aesthetic cues on judgements of motion through the use of dance movements (which are both emotionally and aesthetically-salient). Taken together, the work contained in this thesis should provide an overview of how emotion, aesthetics, empathy, and previous motor experience interact to sculpt social perceptions of the human body in motion. I also hope that the contents of this thesis may contribute to the development of a framework for the role of human motion in whole-body social signalling. I will elaborate on these thoughts in the general discussion.

12. Synopsis of This Thesis

This thesis comprises four main chapters that are provided in the form of experimental manuscripts, with each exploring different elements of emotional and aesthetic evaluation of human motion. In the first chapter, we describe the creation and validation of the McNorm Library, a new set of emotionally-expressive whole-body dance movement sequences depicted in the form of point-light displays (Smith & Cross, 2022). In a series of experiments, participants completed two short questionnaires to provide information about their previous dance experience (the Goldsmith's Dance Sophistication Index, Gold-DSI; Rose et al., 2020) and a measure of their trait empathy (the Toronto Empathy Scale, TES; Spreng et al., 2009), before taking part in an emotion recognition task; providing forced-choice emotion labels to the movement clips from the McNorm Library. The main aims of this experiment were to validate whether the emotion communicated by the movement performer could be identified successfully by observers, and to explore whether individual differences in empathy and previous dance experience would facilitate successful recognition of the target emotion.

In the third thesis chapter, we explore aesthetic perceptions for the same McNorm Library whole-body movement videos (Smith & Cross, under review). In this experiment, participants also completed the Gold-DSI and TES, before being asked to provide judgements of beauty and liking (on 100-point slider scales) for the emotionally-expressive movement sequences. The main aims of this experiment were to determine whether the communication of specific, basic emotions would influence aesthetic evaluation of human motion, and to explore whether empathy and/or previous dance experience would mediate the value we assign to whole-body movements.

In the fourth thesis chapter, we explore the utility of a novel method for visualising human movement in emotion recognition and aesthetic evaluation tasks. More specifically, we examined whether relatively successful emotion recognition rates could be obtained from observers in response to continuous line drawings depicting the trajectories of the wrists and ankles of a dancer during the performance of emotionally expressive whole-body dance movements. We also examined whether these images would function as visual artworks in their own right, separate from their movement origins. The aim of this experiment was to evaluate a new potential method for visualising body movement; to potentially serve as an alternative to PLDs for researchers who wish to explore social perception from stripped back representations of the human body.

In the fifth chapter, we explore emotion judgements through a qualitative lens. In this experiment, both non-dancers and highly experienced dancers assigned emotion judgements to the McNorm Library movement clips and, through a series of interview questions, were asked to provide verbal justification for their decisions. Following the emotion recognition task, all participants were asked a number of additional questions around the topics of kinaesthetic empathy, emotion regulation, and visual attention, in addition to more general questions about their cognitive processes. The aim of this experiment was to delineate between what is important for emotion recognition from an empirical standpoint, and what is perceptually important to observers. An additional aim of this experiment was to examine whether the decision-making process of dancers and non-dancers are different, and if so, in what ways do they diverge. One of my main aims with this final experiment was to isolate a number of novel questions to explore in future social perception research.

In the final chapter, I will discuss the results obtained from the main empirical chapters, how they relate to the previous literature, and the implications of these findings in a wider context. I will also propose the initiation of a new framework for how emotional expressivity is communicated through movement of the dynamic human body; one which describes how aesthetic evaluation, movement kinematics, and individual differences in trait empathy and previous dance training contribute to the perception of emotion in our social world. Finally, in this chapter, I propose four new experiments that I believe would be worthwhile supplements for the contents of this thesis, and would contribute substantially to the framework I propose.

CHAPTER 2

THE MCNORM LIBRARY: CREATING AND VALIDATING A NEW LIBRARY OF EMOTIONALLY EXPRESSIVE WHOLE-BODY DANCE MOVEMENTS

The McNorm Library: Creating and Validating a New Library of Emotionally Expressive Whole-Body Dance Movements

Rebecca A. Smith & Emily S. Cross

ABSTRACT

The ability to exchange affective cues with others plays a key role in our ability to create and maintain meaningful social relationships. We express our emotions through a variety of socially salient cues, including facial expressions, the voice, and body movement. While significant advances have been made in our understanding of verbal and facial communication, to date, understanding of the role played by human body movement in our social interactions remains incomplete. To this end, here we describe the creation and validation of a new set of emotionally expressive whole-body dance movement stimuli, named the Motion Capture Norming (McNorm) Library, which was designed to reconcile a number of limitations associated with previous movement stimuli. This library comprises a series of point-light representations of a dancer's movements, which were performed to communicate neutrality, happiness, sadness, anger, and fear to observers. Based on results from two validation experiments, participants could reliably discriminate the intended emotion to clips in this stimulus set, with accuracy rates up to 60% (chance = 20%). We further explored the impact of dance experience and trait empathy on emotion recognition and found that neither significantly impacted emotion discrimination. As all materials for presenting and analysing this movement library are openly available, we hope this resource will aid other researchers in further exploration of affective communication expressed by human bodily movement.

INTRODUCTION

The ability to communicate with others plays a vital role in our ability to successfully navigate our social world. Successful interactions with others have a substantial impact on how we develop and maintain meaningful relationships, and also have consequences for our sense of belonging to a wider social community, as well as our general mental well-being (Caplan, 2003; Shankar et al., 2015). These interactions are multi-faceted, and researchers have made substantial advances in our understanding of how facial expressions (Jack et al., 2012; Jack et al., 2014; Du et al., 2014) and verbal communication (McAleer et al., 2014; Whiting et al., 2020) mediate the success of our interactions with others. However, key elements of non-verbal communication have been somewhat neglected in this area of research. Human body movement, for example, provides a rich source of information, whose social value to observers is only beginning to be explored (Williams & Cross, 2018; Williams et al., 2019). In fact, emerging research suggests that the human body may provide even more salient cues to emotion than the face (Aviezer et al., 2012a, 2012b; Wang et al., 2018). At present, however, no framework exists that attempts to explain the role played by human body motion in our social interactions. Until this field receives more dedicated research attention, our appreciation of human social signalling, in all of its complexities, will remain incomplete.

A number of movement libraries have been created to explore what observers can identify from simple, everyday human motions (Ma et al., 2006; Vanrie & Verfaillie, 2004; Dekeyser et al., 2002). These libraries vary greatly in the type and complexity of movements they capture, with some focusing on walking motions and others exploring the movement of isolated body areas (e.g., simple arm movements like pointing, waving, and grasping). They also vary in the way visual information about the body is presented, with some depicting movement in the form of full video recordings and others depicting the human form with a reduction of surface-level visual information about a person's form; for example, by rendering human movement as Point-Light Displays (PLDs). PLDs were originally developed by Johansson when he observed that the kinematics of biological motion could be depicted by attaching small light sources to the major joints of a model's body (Johansson, 1975; Krüger et al., 2017). The output of these displays are dot configurations that, when animated, leave the viewer with the impression that they are watching a person (or other animate being) in motion. This introduced the idea that the visual system can interpret animate motion from abstract representations of human figures which are devoid of form cues and other superficial visual information (Chang et al., 2018). Even without morphological cues, research has shown that humans are able to identify a range of features from PLD motion. For example, it has been found that people are highly accurate at identifying the gender (with up to 71% accuracy; see review by Pollick et al., 2005) and identity of PLD walkers (Mitchell & Curry, 2016), and can attribute higher-order social constructs like personality traits and sexuality to these abstract figures and dot configurations (Johnson et al., 2007; Heberlein et al., 2004). In addition, it has been found that individuals can rapidly and reliably attribute affective states to PLD representations of gait cycles and everyday actions (Gunns et al., 2002; Heberlein et al., 2004; Schneider et al., 2014; Atkinson et al., 2004).

PLDs are extremely useful in human movement and affective science research. It is standard practice in emotion research that aims to evaluate the contribution of one particular expressive cue to remove extraneous contextual information from stimuli. In movement libraries, auditory cues (e.g., music, and exertion sounds like breathing and gasping) are typically removed (Jola et al., 2014, are a notable exception), and the performer's face is blurred to remove the competing influence of facial expressions on social judgements (Christensen et al., 2014; Melzer et al., 2019). PLDs represent an extension of this rationale, and as such are a uniquely useful tool for exploring pure motion, in isolation from all other emotionally salient communicative cues (e.g., non-movement related visual cues like facial expressions, and the appearance of the dancer). a number of PLD movement libraries have been created to further the study of human emotion recognition from body movement.

While these stimuli libraries are inarguably useful for exploring how everyday human movements communicate emotional expression to observers, they also have several limitations worth considering. The first is that many of these stimuli depict emotion through pantomimed actions (e.g., shaking fist in anger). While humans do use these kinds of cues to extract high-level social information from others in the real world, we are also able to infer this information from far more nuanced demonstrations of authentic expression. Including these iconic gestural cues in movement sequences likely obscures the influence of subtler components of human movement that provide expressive information. The inclusion of more literal or iconic gestures in libraries may also raise questions about the content validity of studies using them to explore the relationship between motion and emotion recognition, as it could be argued that participants in these studies are providing measures of their ability to successfully recognise social gestures, rather than their ability to infer expression from the performative elements of movement.

Secondly, it has already been noted that these libraries vary greatly in their content and the way they represent the human form. The issue with the scope and variability of these libraries is that this presents significant difficulties for comparing the results across studies and for extracting the most salient results. In the case of exploring emotion recognition from human movement specifically, this has given rise to large inconsistencies in the reported recognition rates for different emotion categories. Often the range of emotions explored in these libraries is very narrow, with authors choosing to focus on recognition of only one or two basic emotions or only on the distinction between positive and negative emotional valence (Castellano et al., 2007; Michalak et al., 2009; Huis In 't Veld et al., 2014a; 2014b). This ignores the diverse range of emotions each of these categories encompass: in essence, equating more complex emotions like anger or fear with sadness. Alternatively, the scope of emotions used in other works may be too broad; with some studies including portrayals of more than 10 specific emotions (Walbott, 1998;

Paterson et al., 2001), and exploring more abstract concepts which are sensitive to cultural or interpretational variance (e.g., pride, shame, strength). However, even studies which account for the middle ground between these approaches (those which focus on several basic emotions - e.g., happiness, sadness, anger, and fear) report inconsistent recognition rates. Successful recognition of happiness, sadness, anger, and fear have been found to vary from 23–92% depending on the specific methodologies of each study, and recognition of neutral emotional displays (i.e., the absence of any clear emotional expression) also vary widely (Dael et al., 2012; Roether et al., 2009; Crane & Gross, 2013; Gross et al., 2010, Atkinson et al., 2004). A brief overview of average recognition rates obtained in response to various types of movement libraries reported in previous works can be found in Supplementary Table S.1 (Atkinson et al., 2004; Crane & Gross, 2013; Gross et al., 2012; Montepare et al., 1999; Pasch & Poppe, 2007; Bernhardt & Robinson, 2007; Dael et al., 2012; Roether et al., 2009; Gross et al., 2010; Camurri et al., 2003; Melzer et al., 2019; Alaerts et al., 2011; Atkinson et al., 2007; Bachmann et al., 2020; Christensen et al., 2016; Christensen et al., 2021; Christensen et al., 2019; Dittrich et al., 1996; Grezes et al., 2007; Dahl & Friberg, 2007).

To begin to address these problems, greater cohesion (or at least clearer correspondence) between different researchers' methodological approaches would help tremendously for building a more reliable and generalisable evidence base on the relationship between body movement and emotional expression. One of the most prominent issues to address in future movement libraries is the inclusion of socially relevant gestures in motion sequences. This is important because gestural communication is highly sensitive to cultural nuance. For example, in many Western countries (e.g., UK, USA) a 'thumbs up' is a positive symbol (indicating something is good) or is used to indicate that you are looking to share a car ride, but in other countries (e.g. Iran) this gesture is an insult, meaning something akin to "up yours" (Archer, 1997). It is likely that being presented with a positive symbol, or an obscenity (depending on cultural background) will influence a participant's emotional state. Therefore, movement libraries which include instances of gestural communication may be unsuitable for research to explore more universal features of emotion recognition, particularly when using culturally diverse samples. In order to overcome such issues, it would be useful to ensure movement sequences feature movements that are not culturally specific, nor inherently tied to any particular emotional state. This should allow researchers to explore the impact of more nuanced movement features, and the expressive quality of motion on perceptions of emotion in observers. One way to reduce the impact of contextual cues on recognition is to use more abstract movements, rather than everyday motions, when creating stimuli libraries. Recently, it has been noted that dance can be of great value to social communication research in general (Orgs et al., 2018; Van Dyck, 2012), and to emotion research in particular (Van Dyke et al., 2017; Aristidou et al., 2017). Dance is, at its heart, the purposeful performance of expressive whole-body movements designed to communicate a narrative or meaning to observers. The inherently communicative nature of dance and the flexibility of its design (at both the choreographic and performative level) make dance ideally suited for the exploration of emotion expression by the human body in motion.

Christensen and colleagues have created two such dance libraries that have been used in emotion research. For their first library, they took 203 movement taken from recordings of full ballet performances, and blurred the dancers' faces in order to remove the impact of facial cues (Christensen et al., 2014). These stimuli were evaluated for 25 movement characteristics, and were also rated for affective valence, arousal, and aesthetic appeal. The choreography was also annotated to provide detail about the specific movements contained in each sequence. Libraries like this (full light displays depicting the dancer in full costume, with set backdrops performing on stage) are particularly useful for aesthetic researchers, as these stimuli are more representative of authentic performative dance than others which contain reduced visual detail (e.g., PLDs). However, for use in emotion research, several critical limitations are associated with this type of stimuli. Using recordings of live dance performances, which contain a variety of visual confounds, may influence emotion judgements, particularly in a dance-experienced sample. It is extremely likely that experienced dancers will recognise elements of choreography from these stimuli and will have a number of these movements in their own repertoire. Familiarity with and intimate knowledge of such movement sequences may influence emotional responses and interfere with recognition data. In addition, costumes or stage furniture may provide cues about the narrative the dance is conveying, which even inexperienced observers can use to infer the emotional content, rather than simply relying on the movements themselves.

A more general issue in this area stems from many of these dance libraries assigning movements into emotion categories based solely on subjective perceptions made by observers, with no consideration given to the intention of the mover while the movement was being performed. By excluding intentionality from evaluating the components of expressive movement, this undermines the dyadic nature of authentic human interaction (Orgs et al., 2016). This issue was addressed in a more recent movement library created by Christensen and colleagues (2019). The Warburg Dance Movement (WADAMO) Library contains a series of movement sequences which were performed three times, each time with a different expressive intention (non-expressive, expressive-positive, or expressive-negative). These stimuli were then shown to dancers and dance-naïve observers, and the intended expression was compared to the expression perceived by observers (Christensen et al., 2019). The WADAMO library represents a substantial improvement on previous work by accounting for both the intention of the performer and the perception of the observer. Together, this makes for more robust classification of stimuli into different expressive categories. However, this library only accounts for the relationship between human motion and expressive valence, rather than providing information about the relationship between movement and the perception of specific emotions. Without creating libraries of this type, ones which address both sides of the interaction dyad and explore the expression of specific emotional states without the inclusion of overt gestures, our understanding of how body movement contributes to human expressivity will remain extremely limited. The Motion Capture Norming (McNorm) library developed in the present study aims to address these important gaps in the research.

The McNorm Library

Description

The Motion Capture Norming (McNorm) library contains a series of whole-body ballet and contemporary dance movement sequences depicted in the form of PLDs. These sequences were new, original pieces of choreography that were devised and performed by a professional dancer. The full library (at the time of recording) comprised **17** different dance sequences. Each sequence was performed 5 times, each time with the aim to communicate a different emotion to observers (neutral, happy, sad, angry, and fearful) while maintaining the same choreography across each performance of the same sequence. Therefore, at the time of recording, the McNorm library comprised 85 recordings (17 of each emotion category). Technical issues, and missing data from the recording phase meant two of these sequences (2 sequences x 5 emotion portrayals = 10 recordings), and 2 individual recordings (1 angry, and 1 fearful) could not be included in the final library validated in this experiment. Therefore, this validation experiment was conducted on 73 emotionally expressive dance movement sequences (15 neutral, 15 happy, 15 sad, 14 angry, and 14 fearful).

Rectifying Limitations of Previous Movement Libraries for use in Emotion Recognition Studies

It has already been noted that many pre-existing movement libraries rely on the emotion judgements of observers to assign expressive movement into different emotion categories. However, the subjective nature of this task creates issues for ensuring the movements assigned to different emotion categories are truly representative of that category. To rectify this issue, the McNorm library was created with expressive intention of the performer in mind. In the validation of this library, the intended emotion of the performer is compared with the subjective emotion categorisation decisions of observers. If intention and subjective judgements align in this validation study, then the McNorm library has generated a more representative sample of expressive motions; one that more effectively captures the dyadic nature of human body movement in communication (between the movement performer and movement observer).

Further, the McNorm library depicts the dynamic human body in the form of PLDs. While limitations are associated with reducing the human form to a configuration of moving dots, this serves to limit the influence of superficial visual cues on emotion recognition. This approach ensures that researchers can use the McNorm library to study motion in isolation from other visual and social cues that are salient to all observers (e.g., perceptions of facial expressions, attractiveness and race of the movement performer) and cues which are of particular relevance to dance-experienced observers (e.g., costumes, stage furniture, etc).

Aims and Predictions of the Validation Study

Our central aim with this validation study is to determine whether the performed emotion (i.e., the emotion intended on the dancer's part) was reliably perceived by observers. As such, we hypothesised that the intended emotion would be recognised at greater than chance level by the participants. Based on the previous literature (see Supplementary Tables S.2 and S.3 for an overview of previous work in this field), we also predicted differences in recognition rates across the different emotion categories. Fear is not frequently explored in relation to this research question, but recognition rates which have been reported tend to be relatively low. Therefore, it is likely that fear will be recognised with lowest accuracy, in comparison with the other emotions explored in this study (Pasch & Poppe, 2007; Camurri et al., 2003; Atkinson et al., 2007; Dittrich et al., 1996; Dahl & Friberg, 2007).

Another aim of this study was to explore the impact of dance experience and trait empathy on emotion recognition capabilities. Dance experience has been found to impact a number of neuropsychological and behavioural outcomes (including emotion recognition capabilities; for a review, see Bläsing et al., 2012). For example, it has been shown that dance training results in significant changes to activity and organisation of sensorimotor structures that compose the action observation network (AON) in the human brain. Several studies have shown that professional dancers show greater engagement within the AON when watching dance compared to non-dancers, and this activity is amplified when dancers observe movement styles they have extensive physical experience performing (Calvo-Merino et al., 2005; Cross et al., 2006). In addition, even purely visual experience with dance has been shown to shape AON responses (Cross et al., 2009; Kirsch & Cross, 2015). There is also evidence of synaptic pruning in subcortical structures associated with the AON and more symmetrical activation of relevant occipitotemporal regions, as a direct result of dance training, which have been argued to reflect more efficient communication between areas involved in this complex neural circuit (Hanggi et al., 2010; Orlandi & Proverbio, 2019).

Beyond neuroimaging studies, it has been shown that dance experience can impact how individuals perceive and respond to the human body in motion. Results from Stevens and colleagues (2010) suggest that dancers and choreographers may develop specific visual-search patterns which influence visual attention when observing complex dance movements, and that these are driven by their movement expertise and learned experiences (Stevens et al., 2010). Studies have also noted that dancers appear to recognise affective expression from human movement with greater accuracy than individuals with no prior dance experience. In one such study, 24 non-dancers and 19 expert ballet dancers observed a series of 5-6s ballet sequences and provided valence ratings on a slider scale from *very sad* (0) to *very happy* (100). They found that dancers were significantly more accurate in their emotion classifications for recognition of positive affect than non-dancers (Christensen et al., 2016). Similar results were observed in a later study, wherein participants from a variety of dance backgrounds were asked to observe a series of movement sequences and identify whether these movements were expressive or non-expressive (Christensen et al., 2019). The authors report that years of dance experience correlated with accuracy in recognition of the intended expression, providing further support for the idea that dance experience facilitates the ability to recognise expression from emotion whole-body movement displays. Therefore, in addition to the emotion recognition task, participants will be asked to complete a shortened version of the Gold-DSI (Rose et al., 2020) to provide detail about their physical and observational engagement with dance. These factors will be explored in relation to emotion recognition accuracy to determine whether different levels of dance experience mediate the ability to interpret expressive cues from the movements of others.

Individual differences in empathy were also identified as another relevant factor to explore. Many previous studies have observed a link between empathy and the ability to successfully interpret emotional cues presented by others in our social environment. This has been observed across emotion recognition tasks using a variety of social cues; including faces, voices, body postures, and body movements (Besel & Yuille, 2010; Jospe et al., 2018; Israelashvili et al., 2020; Soto & Levenson, 2009; Balconi & Bortolotti, 2012; Rizzolatti & Craighero, 2005; Holland et al., 2021; Neumann et al., 2014). This link is even clearer when considering Autism Spectrum Disorders (ASD). Individuals with ASD tend to provide lower self-reported levels of empathy, and typically show lower performance on empathy-related tasks compared to typically developed individuals (Kok et al., 2016; Pepper et al., 2019; Bishop & Seltzer, 2012; Trimmer et al., 2017; Demurie et al., 2011). Furthermore, a common characteristic of ASD is difficulty in interpreting social cues, including the expression of emotions. Considering these factors together, it is plausible to expect that individual differences in empathy may mediate the ability to successfully interpret emotion expression from social behaviours.

Furthermore, research suggests there may also be a link between movement expertise and empathy, and these factors may interact to produce changes in emotion recognition capabilities. For example, it has been found that dancers have higher interoceptive accuracy than individuals without such movement training (Christensen et al., 2018), and more skilled dancers provide higher self-reported levels of emotional intelligence (Petrides et al., 2006). While these studies do not provide a direct link between movement expertise and increased empathy, the increasing implementation of dance movement therapy in the management of ASD suggests that dance may facilitate the development of empathic behaviours (Behrends et al., 2016; McGarry & Russo, 2011; Federman, 2011; Mastrominico et al., 2018; Koch et al., 2015). To this end, participants in the McNorm experiments will also complete the Toronto Empathy Scale (Spreng et al., 2009), to generate individual trait empathy scores for consideration in the analysis.

METHODS

MCNORM LIBRARY CREATION

2.1 PARTICIPANTS

One female professional dancer, previously a principal dancer with the Scottish Ballet, participated in the stimuli creation procedure. The dancer was chosen due to her extensive training, totalling more than 9 years with professional companies and 4 years in a freelance capacity, as well as her additional choreographic experience. The dancer contributed over a 6-day period (2 days to generate choreography and a 4-day recording period) and was provided with an honorarium for her time.

2.2 STIMULI CREATION

Movement sequences were recorded in the University of Glasgow's motion capture lab, using 12 Vicon MXF40 cameras, recording at a rate of 120 frames per second (120fps), to produce point-light displays for a new library of expressive dance movements.

2.2.1 CREATION AND RECORDING OF MOVEMENT SEQUENCES

In advance of the filming dates, the dancer was asked to create a series of classical and contemporary ballet sequences. The dancer was informed that these movement sequences should contain only neutral movements (i.e., movements not inherently linked to a particular narrative, or emotional expression), in order to limit interference from previously acquired knowledge when performing the choreography. The dancer was also instructed that these movement sequences should have a minimum duration of 6 seconds. No maximum duration was explicitly defined, but the dancer was informed that longer sequences may be prone to more technical errors during recording and that she should not create sequences that, when performed repeatedly, would cause unnecessary fatigue.

The dancer was informed, prior to recording days, that the objective would be to perform each movement sequence five times, each time maintaining the same choreography but portraying a different emotional state; neutral (non-expressive), happy, sad, angry, and fearful. The experimenter did not show any movements, nor did she give examples or directions regarding expression; therefore, these expressive portrayals reflect the dancer's personal interpretation of the different emotional states. For the neutral performances, the dancer was instructed to perform the movements with the same technical accuracy (with the same level of mechanical precision) but without expressivity; as if she was "going through the motions" or "working through an exercise". To further limit interference from the experimenter, all communication about expression and technicalities of the sequences were discussed verbally, rather than through potentially leading gestures or body movement. Sequences were re-recorded at the dancer's request when she had made a misstep or felt that she had not portrayed the emotion in a satisfactory way, and only at the experimenter's request when there was a technical issue.

Finally, while each sequence had a different musical accompaniment (to aid with timing, recall of the choreography and the feeling of performance of these movements), the music remained constant across all emotion portrayals, and all audio was removed as part of the final stimuli set.

The first day of the four-day filming period served as a trial run. The dancer spent 30 minutes at the start of each day warming up to ensure she would not injure herself, before practicing the choreography. The dancer then performed each movement sequence, always starting with the neutral portrayal before moving on to the other expressive categories in an order selected by the dancer (to ensure the dancer was in the correct mindset and the resultant expression felt authentic).

During the filming stage, 39 retroreflective markers were placed on the dancer's body in anatomical regions defined by the Plug-in Gait Model, which is widely accepted for use in biological motion research (Kainz et al., 2017; Piwek et al., 2016). This placement was checked before and after each recording to ensure no markers had come loose or had fallen off and corrected where necessary.

In review, 17 sequences were performed, with each sequence performed five times (neutral, happy, sad, angry, fearful), creating a total of 85 recordings for the first iteration of this stimulus library. These recordings ranged in duration from 6.6 – 42.8 seconds, with an average duration of 22.1 seconds. Sad clips, on average were the longest in duration (average = 24.3 seconds), while angry clips were the shortest in duration (average = 20.6 seconds). The average duration of neutral, happy, and fearful clips were 21.1, 21.9, and 22.5 seconds respectively. Further detail about each of the stimuli recordings can be found in Supplementary Tables S.4 and S.5.

2.2.2 POINT-LIGHT DISPLAY CREATION

In the Nexus software, each marker was labelled on a frame-by-frame basis. A skeleton template was then applied over each recording to generate 15 new body markers from the original 39 placed on the dancer's body. These 15 new markers depict a simplified figure with less visual clutter, while maintaining the overall body form (see Figure 2.1 for more detail about marker transformation).

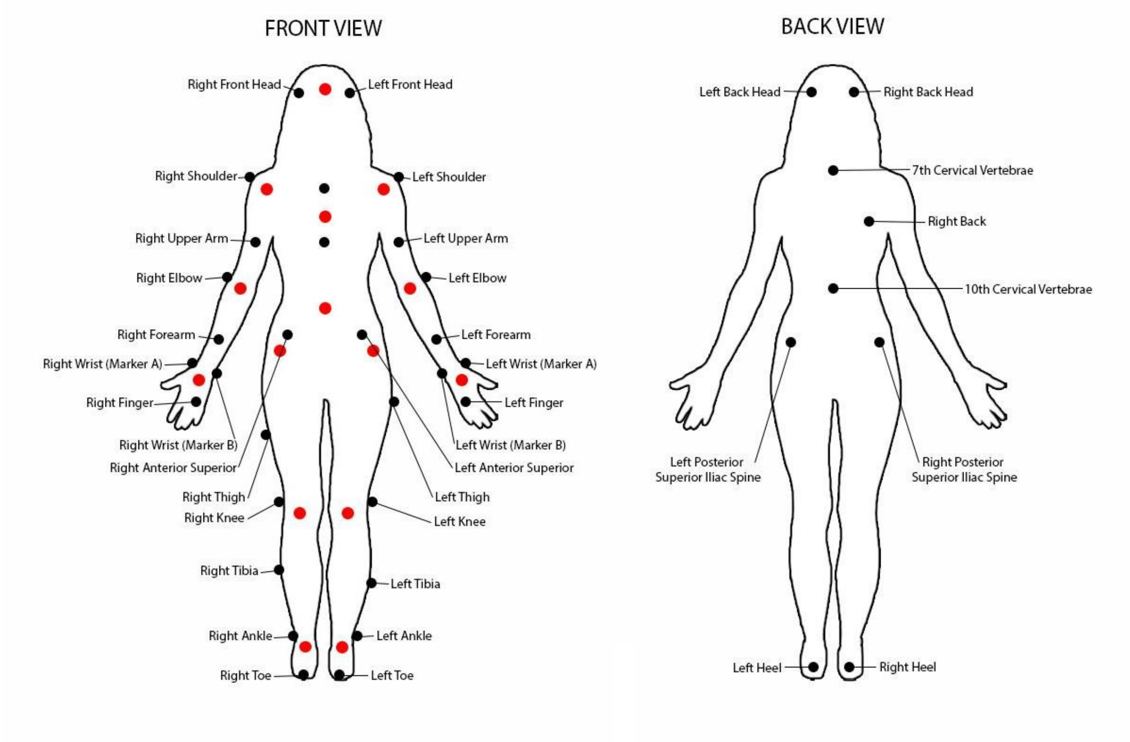


Figure 2.1: Placement of retroreflective markers on the body during the recording phase are denoted in black. The Plug-In Gait template in Vicon Nexus was used to convert the original 39 markers to point-light display figures (the final markers are denoted in red).

For each of the movement recordings, 3D coordinate information was extracted from the Vicon Nexus data files. Then for each time frame of each clip, the 2D coordinates depicting the upright, frontal perspective were extracted and plotted on a scatterplot (as white points on a black backdrop) to create a static point-light display image. In Python, these images were then displayed one after the other (in a manner akin to the creation of stop-motion animation) and presentation timestamps were adjusted by a factor of *0.23 (to visually match the speed of the original recording) to generate a dynamic point-light display for each sequence.

As a result of the highly dynamic nature of some of the movements and sequences, some movement information was lost during the recording process, resulting in significant gaps in the coordinate data. In order to overcome these issues, many '*gap-filling*' techniques exist, each varying in their accuracy depending on the type of data they applied to. For simpler movement sequences (e.g., knocking, walking) a linear extrapolation (from point x, where the gap begins, to point y, where the information returns) may be sufficient but for highly dynamic, complex dance movements linear extrapolation is too simplistic, and can result in movement distortion. For this data, it was determined that a combination of gap-filling techniques should be implemented on a case-by-case basis. Two visual examples

(Supplementary Figures S.1 & S.2) and more detail about this gap-filling procedure are available in Section A of the Supplementary Materials.

As mentioned in the description of this library in the previous section, for some sequences, the gaps were too large or numerous to create a complete, non-distorted point-light display clip. Therefore, 2 full movement sequences (2 sequences x 5 emotion portrayals = 10 recordings), and 2 individual recordings (1 angry, 1 fearful) were discarded at this stage, leaving 73 point-light display recordings to form the final stimuli set (15 neutral, 15 happy, 15 sad, 14 angry, 14 fearful).

MCNORM 1: FIRST VALIDATION EXPERIMENT

METHODS

OPEN SCIENCE STATEMENT

Consistent with open science practices widely adopted within psychological research (Open Science Collaboration, 2012), we report all manipulations and all measures in the study. In addition, following open science initiatives (Munafò, 2016), the data, stimuli, and analysis code associated with this study are freely available on the Open Science Framework. By making the data available, we enable others to pursue tests of alternative hypotheses, as well as more exploratory analyses.

PARTICIPANTS

Previous attempts to generate and validate movement libraries have participant samples sizes that have ranged between 12 and 80 participants (see Supplementary Table S.2 for an overview). While the average number of participants tested in this prior literature is around 40 participants, we aimed to collect data from at least that many participants in our two validation experiments. 1179 participants started the validation experiment, but due to the volume of incomplete responses, only those who completed the task in full were included in the analysis. Fifty participants completed the entire task, and thus made up the final sample for the first validation experiment. Participants ranged in the amount of previous dance experience they had, but they were assigned to one of five different dance-level groups (non-dancer, beginner, intermediate, advanced, or professional) by indicating which label they thought was most applicable to their previous experiences. A more detailed breakdown of sample demographics is presented in Table 2.1. The experiment was created in formr (Arslan et al., 2020) and was advertised through the University of Glasgow subject pool, and on a variety of social media channels (including Twitter, Facebook, and Instagram).

Table 2.1: Demographics information for participants McNorm Experiment 1 (N = 50).

Age	Mean (SD)	*29.06 (11.8)*
	Range	*18 – 63*
Gender	Female	*32*
	Male	*15*
	Transgender Female	*1*
	Transgender Male	*2*
	Gender Variant/Non-Conforming	*NA*
Dance Level	Non-Dancer	*13*
(Self-Reported)	Beginner	*16*
	Intermediate	*15*

Advanced	*4*
Professional	*2*

STIMULI

The 73 complete movement sequences outlined in the previous section were uploaded to Vimeo and were displayed through formr using embed codes. In Vimeo, the embed codes were manipulated to remove the standard online video handles (video title, uploader information, suggested videos, and Vimeo branding). The video controllers were also removed, and the video was set to loop automatically until participants had submitted their responses for the item. The stimuli were all presented on screen at a consistent size of 496 x 370 pixels.

PROCEDURE

Participants followed the formr link to begin the experiment. After reading the information and consenting to participate (in accordance with BPS guidelines), participants completed two questionnaires: a series of questions extracted from an early, unpublished, version of the Goldsmiths Dance Sophistication Index (Gold-DSI; Rose et al., 2020), and the Toronto Empathy Scale (TES; Spreng et al., 2009).

Selected questions from the early Gold-DSI included demographics measures which allowed participants to provide detail about their previous dance experience, covering both formal and informal experience, and both visual and physical experience. Participants were also asked to select which level of dance experience was most applicable to them from one of five options (non-dancer, beginner, intermediate, advanced, or professional. The version of the early Gold-DSI which was used in this experiment can be found on the OSF (https://osf.io/458sq/). The TES was used to generate a trait empathy score for each participant (ranging from 0 to a maximum of 40) and includes a variety of questions (both forward and reverse scored) covering facets of both emotional and cognitive empathy.

Testing began with a short practice block, containing 5 different point-light display movements, to ensure participants understood the instructions and that there were no issues with video playback. Following this, participants began the validation experiment wherein they watched each of the 73 movement sequences (presented in random order) and after observing each sequence were asked *"What emotion do you think the dancer is trying to convey?"* in a forced-choice paradigm (neutral, happy, sad, angry or fearful). For each clip, they also responded on a sliding scale to the questions *"How intensely is the emotion being expressed?"* from not intense (0) to very intense (100), and *"How sure are you of your decision?"* from very uncertain (0) to very certain (100). The whole experiment took between 45 and 60 minutes for the volunteers to complete and first-year psychology students at the University of Glasgow were given 3-4 participation credits for taking part.

As this experiment was conducted using an online sample, three catch trials were included to test whether participants were paying sufficient attention to the stimuli and task. For this purpose, a separate point-light display sequence was manually edited twice to depict two new scrambled dot motion videos. At random points throughout the duration of the experiment, these two scrambled motion videos and the original unedited version appeared and on the next screen participants were asked *"Did you perceive human motion in the previous clip?"*. The correct answer was only *"Yes"* for one of these three trials. However, in hindsight, it was decided that this question may have been too open to interpretation, as the scrambled motion videos were created from an actual motion clip and subjects may have observed, for example, an arm-like dot configuration that they perceived to be humanlike in motion. Moreover, when examining the impact of correct responses to the catch trials on task performance, no significant difference in accuracy was observed between those who passed (M= 26.66%) and failed (M= 27.13%) the attention checks: $t(47.63)= 0.38$, $p= 0.71$. Therefore, it was decided that failing to correctly respond to the attention checks would not warrant immediate exclusion of that participant from the analysis.

RESULTS

Perceived Emotion Results

The average percentage recognition of the intended emotion, across all participants and all emotions was 26.9% (SD = 4.49). With a baseline of 20% established as chance level (from a 1 in 5 chance of selecting the intended emotion), after normality testing ($p > 0.05$), a one-sample t-test revealed that overall participants identified the intended emotions at greater than chance level: $t(49) = 10.86$, $p < 0.001$, $d = 1.54$.

Examining each emotion individually, we found that all emotions with the exception of fear were recognised at significantly greater than chance level. See Table 2.2 for a summary of recognition rates. Further, the mean recognition rates for each expression category, and the distribution of responses can be observed in Figure 2.2.

Table 2.2: A summary of recognition rates for each emotion category (means and standard deviations), and the results of inferential tests performed on the data to determine whether they were recognised at greater than chance level.

Emotion	Normally Distributed?	Average Recognition Rate (%)	Greater than Chance?	Test	Results
Neutral	Yes	28.13% *(SD= 13.21)*	**Yes**	One-Sample t-test	t(49)= 4.35, p <0.001, d = 0.62.
Happy	No	29.73% *(SD= 10.1)*	**Yes**	One-Sample Wilcoxon Signed Rank Test	Z= 713, p <0.001, d = 0.96.
Sad	Yes	32.67% *(SD= 10.28)*	**Yes**	One-Sample t-test	t(49)= 8.71, p <0.001, d = 1.23.
Angry	No	26% *(SD= 10.78)*	**Yes**	One-Sample Wilcoxon Signed Rank Test	Z= 1023, p <0.001, d = 0.56.
Fearful	No	17.29% *(SD= 11.18)*	**No**	One-Sample Wilcoxon Signed Rank Test	Z= 465, p= 0.095, d = 0.24.

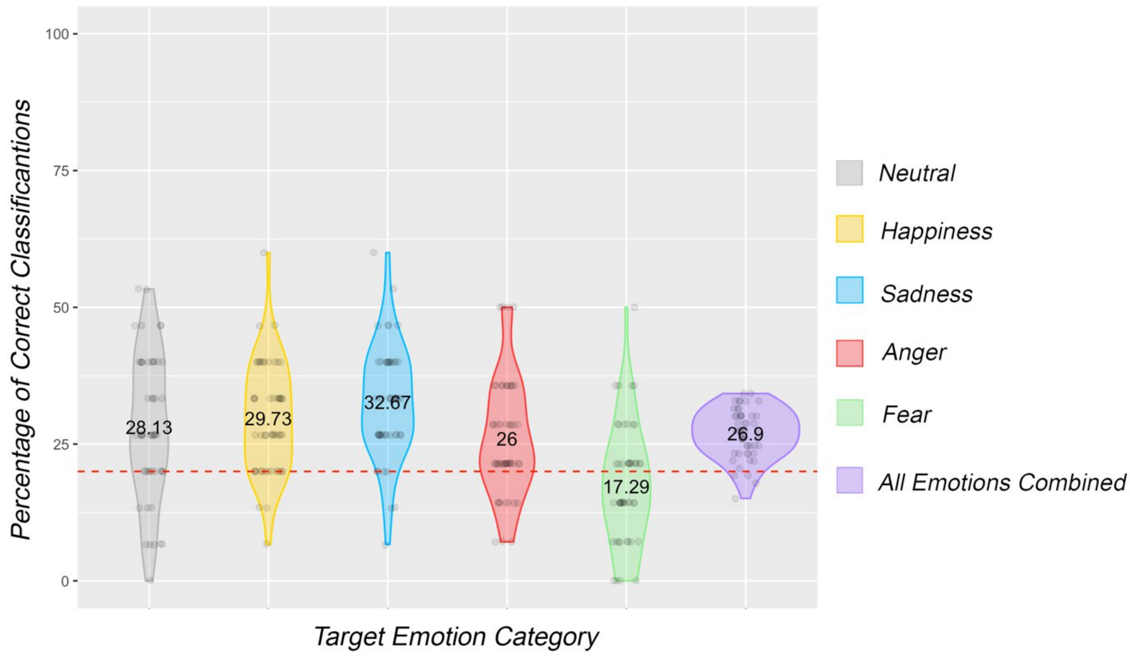


Figure 2.2: Violin plots depicting the distribution of recognition rates for each emotion category. Values presented in the centre of each of the violins represent the mean recognition rate, and the red dotted line indicates chance level of recognition. All emotions, with the exception of fear, were recognised at greater than chance level.

Although all emotions but fear were recognised at greater than chance level, a number of frequent misclassifications still occurred. Movements intended to communicate neutral expression were often perceived to communicate happiness (26.34%). Movements intended to express anger were also mistaken for happiness (27.63%), in fact even more often that they were labelled correctly (26.47%). Fear was confused with neutrality (23.21%), happiness (21.49%), and sadness (23.21%), more often than it was correctly labelled (17.34%). A more detailed breakdown of the correct classification and misclassification rates can be found in Table 2.3.

Table 2.3: Correct Classification and Misclassification Rates for Each Emotion Category in Experiment 1.

		Perceived Emotion				
		Neutral	Happy	Sad	Angry	Fearful
Intended Emotion	Neutral	**28.36%**	26.34%	19.49%	11.56%	14.25%
	Happy	21.10%	**29.97%**	21.1%	17.47%	10.35%
	Sad	19.41%	20.49%	**33.02%**	14.56%	12.53%
	Angry	23.45%	27.63%	10.79%	**26.47%**	11.65%
	Fearful	23.21%	21.49%	23.21%	14.76%	**17.34%**

These results provide support for the primary hypothesis, which was that participants should recognise the intended emotion at a level greater than chance. The data show that overall recognition accuracy across all participants was statistically greater than chance. It was also predicted that there would be variation in recognition rates for each specific emotion category, with movements expressing fear being recognised the least well. The data supports the secondary hypothesis as fear was the only emotion category not to be recognised at greater than chance level, and sadness was recognised at a higher rate than any other emotion.

The relatively high standard deviations for each emotion category suggested variation in recognition rates for each individual clip in the stimuli set. Exploratory analysis confirmed this idea, as recognition rates were found to vary widely for each individual clip in the movement library. Recognition of neutral expression varied from 4-62% across the 15 clips in the stimuli set. Similarly, recognition of happiness varied from 4-60% and recognition of anger varied from 2-66% from clip to clip. For both the sad and fearful clips, some stimuli were never assigned the intended emotion category (with recognition of sad ranging from 0-66%, and recognition of the fearful expressions ranging from 0-44%). See Supplementary Table S.4 for a detailed breakdown of recognition rates for each individual clip in the McNorm library.

<u>Exploratory Analysis</u>

Impact of Dance Experience on Recognition

Participants completed a number of measures from an early unpublished version of the Gold-DSI to provide detail about their prior dance experience. One of these measures asked participants to classify their level of dance experience according to the following options: non-dancer, beginner, intermediate, advanced, or professional. A Kruskal-Wallis Test was conducted to determine whether a participant's self-reported level of dance experience had an impact on their ability to correctly identify the target emotion from clips in the McNorm library. However, no significant differences were observed between these groups (Chi-Square = 2.597, $p = 0.627$, $df = 4$). It should be noted here that participants were not asked to provide numerical answers about their years of experience, the frequency of their training, or the age at which they first engaged with dance classes (in any form), so examination of these more specific aspects of physical dance experience was not possible with these data.

Participants completed a number of additional measures, three of which were related to physical dance experience ("How often do you currently go dancing for fun/social reasons?", "I have taken regular dance classes at least once a week for…", and "I have had formal training in any dance style for…") and two measures were related to visual experience ("How often do you watch dance performances/shows/videos on TV or Internet?", and "How often do you attend live dance performances?"). Therefore, these item responses were grouped together to create a more general score for each participant's physical dance experience and observational dance experience, respectively. An additional item from the questionnaire was not related to the frequency of experience, instead asking participants to indicate the number of dance styles they had experience with (using the options; 'none', 'one dance style', 'two dance styles', 'three dance styles', or 'more than three dance styles') and was therefore considered separately in the analysis. As the number of possible responses to these items in this questionnaire varied from 5-7, prior to further analysis the responses were transformed using min-max normalization to account for the differences in scales.

A three-stage hierarchical multiple regression was conducted with average recognition rate as the dependent variable. The physical dance experience factor was entered at the first stage of the regression, the observational dance experience factor was added in the second stage, and number of styles a participant was familiar with was added at the final stage to create the maximal model. As this analysis is exploratory in nature, the factors were added in this order to account for the frequency with which these facets of dance experience have been explored in previous research in this field. The results can be found in Table 2.4 below.

Table 2.4: Summary of Hierarchical Multiple Regression Analysis for Factors Predicting Recognition Accuracy in McNorm Experiment 1:

Model	Summary	Predictor	B	T	Std Err
Model 1	$F(1,48)= 0.367^{n.s.}$, $R^2_{Adj}= -0.013$, RSE = 4.524	Intercept	27.366		
		Physical Experience	$-1.162^{n.s.}$	-0.606	1.916
Model 2	$F(2,47)= 0.580^{n.s.}$, $R^2_{Adj} = -0.017$, RSE = 4.534	Intercept	26.605		
		Physical Experience	$-2.357^{n.s.}$	-1.006	2.343
		Observational Experience	$2.385^{n.s.}$	0.891	2.678
Model 3	$F(3,46)= 0.387^{n.s.}$, $R^2_{Adj} = -0.039$, RSE = 4.581	Intercept	26.642		
		Physical Experience	$-1.965^{n.s.}$	-0.577	3.406
		Observational Experience	$2.416^{n.s.}$	0.891	2.712
		Number of Styles	$-0.438^{n.s.}$	-0.160	2.734

Note: N = 50; *n.s.*p > .05, *p < .05, **p < .01, ***p< .001

Model 1: Percentage Recognition ~ Physical Dance Experience Factor

Model 2: Percentage Recognition ~ Physical Dance Experience Factor + Observational Dance Experience Factor

Model 3: Percentage Recognition ~ Physical Dance Experience Factor + Observational Dance Experience Factor + Number of Dance Styles a Participant Has Experience With

The hierarchical multiple regression revealed that at Stage one, the regression model was not a significantly better fit for the data than the null model $F(1,48)= 0.367$, $p= 0.547$, and that the physical dance experience factor only explained 1.3% of the variation in recognition accuracy. Introducing the observational dance experience factor at Stage two did not significantly improve the model when compared with the null model $F(2,47)= 0.580$, $p= 0.563$, and the observational dance factor only explained an additional 0.4% of the variation in recognition rates. Finally, the maximal model added the number of dance styles a participant had experience with to the model, but the addition of this predictor did not significantly improve the fit of the model to the data $F(3,46)= 0.387$, $p= 0.763$ and only explained an additional 2.2% of the variance in recognition rates. Therefore, previous dance experience (as measured by items from an early version of the Gold-DSI) did not have a significant impact on recognition of the target emotion in this experiment.

Relationship Between Empathy and Recognition of the Intended Emotion

Participants completed the Toronto Empathy Scale to provide a measure of trait empathy. Possible scores for this measure range between 0 – 64. A Pearson's correlation was conducted to explore the relationship between trait empathy and recognition of the intended emotion and no significant relationship was observed: cor= 0.0048, $p= 0.973$.

Intensity Ratings

In addition to assigning an emotion category to each movement sequence, participants also provided a measure of how intensely they felt the movements portrayed their chosen emotion on a slider scale from 0 (not intense) to 100 (very intense). The average overall intensity score provided by participants, across all emotions was 56.69 (± 10.5). The average intensity scores provided for clips in each emotion category can be found in Table 2.5.

Table 2.5: Average intensity ratings for clips in each emotion category in Experiment 1.

	Neutral	Happy	Sad	Angry	Fearful
Mean Intensity Score	54.36 *(± 12.32)*	57.44 *(± 11.72)*	57.54 *(± 11.1)*	59.02 *(± 11.33)*	55.18 *(± 10.6)*

From the mean intensity scores, it appears that portrayals of neutral expression were assigned the lowest intensity scores and portrayals of anger were assigned the highest intensity scores. A one-way ANOVA was conducted to explore the significance of these differences and no significant differences were observed in intensity scores across the different emotion categories: $F(5,249)= 1.132$, $p= 0.34$, $d= 0.14$.

Certainty Ratings

Participants also provided a measure of how certain they were of their emotion judgements on a slider scale from 0 (very uncertain) to 100 (very certain). The average certainty score for all participants across the entirety of the task was 54.12 (± 13.8). The mean certainty scores for each emotion category can be found in Table 2.6.

Table 2.6: Average certainty ratings for clips in each emotion category in Experiment 1.

	Neutral	Happy	Sad	Angry	Fearful
Mean Certainty Score	52.97 *(± 14.31)*	55.9 *(± 15.79)*	53.09 *(± 13.85)*	55.89 *(± 14.94)*	52.76 *(± 13.74)*

It appears that participants were least certain about their categorisation of neutral and fearful expressions and most certain about their perceptions of happiness and anger. However, a one-way ANOVA revealed that these differences were not significant: $F(5,294)= 0.51$, $p= 0.77$, $d= 0.104$.

Subset of the McNorm Library: Clips with Highest Agreement in Perception of the Intended Emotion

Due to the large variation in recognition rates across individual movement clips in the stimuli set, a subset of these clips was examined further. The four clips with the highest rate of agreement in emotion perceptions (i.e., the highest recognition rates for the intended emotion) across the different emotion categories were isolated for further analysis. The selected clips are highlighted in Supplementary Figure S.3, which also provides a more detailed breakdown of average recognition rates for these clips (and all other clips in the McNorm Library).

Recognition Results

For the new subset of the McNorm stimuli, the average overall percentage recognition of the intended emotion was 52.2% (± 12.42). See Figure 2.3 for a summary of recognition rates across each individual emotion category and the distribution of participants' responses. With a baseline of 20% established as chance level, after normality testing ($p > 0.05$), a one-sample t-test revealed that overall participants identified the intended emotions at greater than chance level: $t(49) = 18.33$, $p < 0.001$, $d= 2.59$.

Examining each emotion individually for the new subset of clips, all emotions were recognised at significantly greater than chance level. See Table 2.7 below for summary of recognition rates.

Table 2.7: A summary of descriptive and inferential statistics for the subset of 20 clips from the full McNorm library identified for future study. The table shows the mean and standard deviation of recognition rates for each emotion category, and the results of inferential tests performed on the data to determine whether they were recognised at greater than chance level.

Emotion	Normally Distributed?	Average Recognition Rate (%)	Greater than Chance?	Test	Results
Neutral	No	57.5% *(SD= 30.83)*	**Yes**	One-Sample Wilcoxon Signed Rank Test	Z= 1227, p <0.001, d= 1.22.
Happy	No	53.5% *(SD= 22.02)*	**Yes**	One-Sample Wilcoxon Signed Rank Test	Z= 1275, p <0.001, d= 1.52.
Sad	No	60% *(SD= 25.75)*	**Yes**	One-Sample Wilcoxon Signed Rank Test	Z= 1256, p <0.001, d= 1.55.
Angry	No	52% *(SD= 23.6)*	**Yes**	One-Sample Wilcoxon Signed Rank Test	Z= 1248, p <0.001, d= 1.36.
Fearful	No	38 % *(SD= 29.55)*	**Yes**	One-Sample Wilcoxon Signed Rank Test	Z= 1054, p <0.001, d= 0.61.

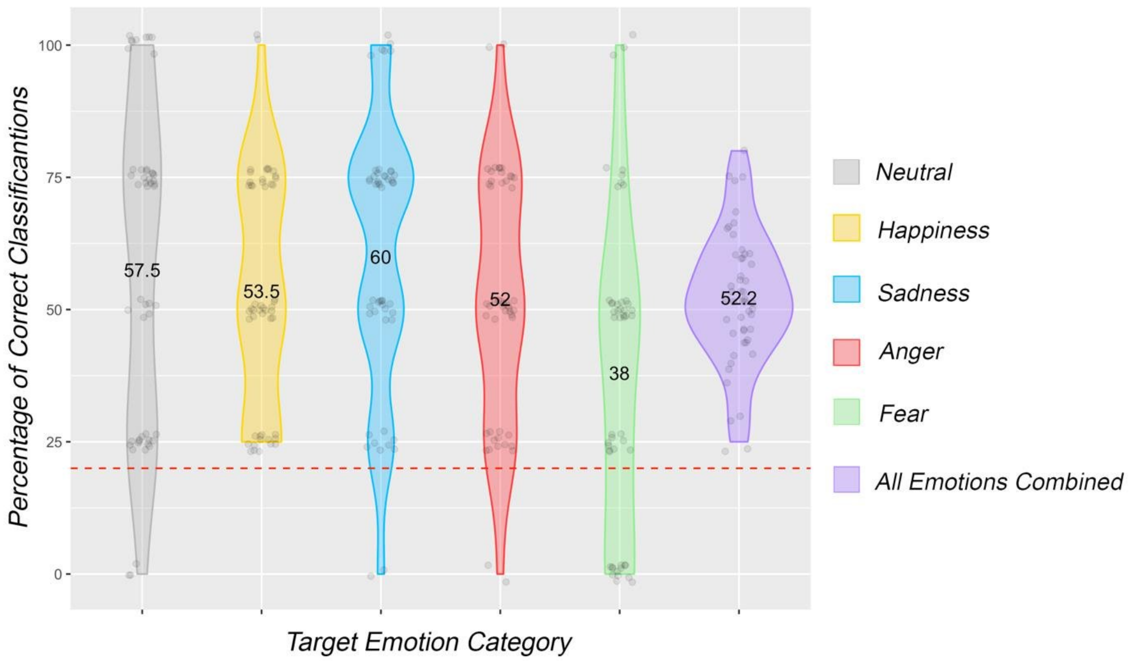


Figure 2.3: Violin plots depicting the distribution of recognition rates for each emotion category for ONLY the 20 clips from the full McNorm library identified for future study. Values presented in the centre of each of the violins represent the mean recognition rate, and the red dotted line indicates chance level of recognition. All emotion categories were recognised at greater than chance level for this subset of clips.

While this subset of clips from the McNorm library was recognised with greater accuracy, were a number of common misclassifications persisted. Expressions of happiness were often confused for expressions of anger (27.27%) and conversely, anger was often mistaken for happiness (31.66%). Fearful expressions were often considered to depict sadness (23%) and sadness was sometimes confused with neutral expression (19.29%). Neutrality was also often perceived to be happy (16.08%). Table 2.8 provides a more detailed breakdown of correct classifications and misclassifications attributed to this subset of the McNorm library.

Table 2.8: Correct classification and misclassification rates for each emotion category for the subset of 20 clips from the McNorm library identified for further study.

		Perceived Emotion				
		Neutral	Happy	Sad	Angry	Fearful
Intended Emotion	Neutral	**57.79%**	16.08%	11.56%	4.02%	10.55%
	Happy	7.07%	**54.04%**	8.59%	27.27%	3.03%
	Sad	19.29%	5.08%	**60.91%**	3.55%	11.17%
	Angry	5.53%	31.66%	4.02%	**52.26%**	6.53%
	Fearful	6.5%	14%	23%	18.5%	**38%**

Intensity Ratings

In addition to assigning an emotion category to each movement sequence, participants also provided a measure of intensity of expression on a slider scale from 0 (not intense) to 100 (very intense). The average overall intensity score provided by participants, across all emotions was 60.43 (± 11.3). The average intensity scores provided for clips in each emotion category can be found in Table 2.9.

Table 2.9: Average intensity ratings for each emotion category for the subset of 20 clips from the full McNorm movement library.

	Neutral	Happy	Sad	Angry	Fearful
Mean Intensity Score	46.94 *(± 18.02)*	66.08 *(± 15.61)*	58.81 *(± 14.59)*	72 *(± 12.78)*	58.01 *(± 14.29)*

From the mean scores outlined above, it appears that participants perceived certain emotions to be portrayed with greater intensity than others in this subset of clips from the full McNorm library. A one-way ANOVA confirmed that several of these differences were significant: F(5,249)= 16.74, p <0.001, d= 0.56. After correcting for multiple comparisons (with a Tukey HSD test), expressions of anger were perceived to be significantly more intense than expressions of fear (p <0.001), neutrality (p <0.001), and sadness (p <0.001), but not more intense than happy expression (p= 0.33). Neutral expressions were assigned significantly lower intensity ratings than sad (p <0.001), happy (p <0.001) and fearful expressions (p <0.005). There were no significant differences in intensity ratings for sequences intending to communicate happiness and fear (p= 0.07), or happiness and sadness (p= 0.13), or fear and sadness (p >0.9). No other significant differences were observed (all p values >0.05).

Certainty Ratings

Participants' provided certainty on their judgements based on a sliding scale from 0 (very uncertain) to 100 (very certain) throughout the experiment. The average certainty score for all participants across the entirety of the task was 56.64 (± 14.5). The mean certainty scores for each emotion category can be found in Table 2.10.

Table 2.10: Average certainty ratings for each emotion category for the subset of 20 clips from the full McNorm movement library.

	Neutral	Happy	Sad	Angry	Fearful
Mean Certainty Score	54.19 *(± 18.68)*	60.96 *(± 16.85)*	54.22 *(± 16.66)*	60.44 *(± 17.78)*	53 *(± 17.67)*

It appears that participants were most certain about movements in the happy and angry expression categories, and least certain about the fearful, sad and neutral expressions. However, the results of a one-way ANOVA found that these differences were not statistically significant: F(5,294)= 1.996, p= 0.08, d= 0.20.

Summary

The reason for the large variation in recognition rates across all clips in the full McNorm library is unclear. A lack of attention to the task may be one potential explanation. It should be noted here that this experiment took between 45 and 60 minutes to complete in full, and it is well established that engaging in long duration experiments which involve the presentation of repetitive stimuli can lead to a decrease in task attention (see Langner & Eikhoff, 2013, for a comprehensive meta-analysis of these findings). Catch trials were included in an attempt to exclude participants who did not fully attend to the stimuli, however in this experiment these trials were deemed to be too subjective to use as a benchmark for participant exclusion. This may too have been exacerbated by remote data collection. In non-laboratory settings, the possible influence of environmental distractors cannot be ruled out. Therefore, as it was not possible to control for distractions, and especially given nature of this task, it is possible that low attention levels (or a high number of distractions) contributed to the relatively low recognition rates seen in Experiment 1. Alternatively, given the fairly high recognition rates for some of the movement stimuli, it is possible that the subset of clips isolated for further analysis are simply more representative in their portrayals of the intended emotions. To explore this possibility, a second validation experiment was devised and conducted using only this new subset of clips.

MCNORM 2: SECOND VALIDATION EXPERIMENT

METHODS

PARTICIPANTS

A total of 722 participants started this experiment. After excluding participants who did not complete the task in full, the final sample size was 77. A more detailed breakdown of sample demographics can be found in Table 2.11. The experiment was also created in formr (Arslan et al., 2020) and was advertised through the University of Glasgow subject pool, and on social media channels. Participants who participated in Experiment 1 were informed that they were not eligible to take part in Experiment 2.

Table 2.11: Demographics information for participants who took part in the second McNorm Experiment (N = 77).

Age	Mean (SD)	*31.96 (13.09)*
	Range	*20 – 66*
Gender	Female	*66*
	Male	*10*
	Transgender Female	*NA*
	Transgender Male	*NA*
	Gender Variant/Non-Conforming	*NA*
	Missing	*1*
Dance Level	Non-Dancer	*24*
(Self-Reported)	Beginner	*29*
	Intermediate	*18*
	Advanced	*5*
	Professional	*NA*

STIMULI

The four clips with the highest recognition rates, for each emotion category (as identified in the first experiment) were used (i.e., 4 x each emotion category (5; neutral, happy, sad, angry, fearful) = 20 clips). As before, these clips were shown in formr using Vimeo embed links (with all video handles removed).

PROCEDURE

As in the first experiment, after reading the experiment information and providing informed consent, participants provided responses to the shortened version of the early Gold-DSI and the Toronto

Empathy Scale. They then took part in a practice block with 5 trials (using the next best-recognised clips for each emotion category) to ensure they understood the task instructions and that there were no video playback issues. Then participants watched each of the 20 movement sequences (presented in random order) and responded to the question *"What emotion do you think the dancer is trying to convey?"* from the same five options (neutral, happy, sad, angry or fearful). They also provided responses on the same intensity and certainty slider scales, as participants did in experiment 1. The whole experiment took around 15 minutes to complete and first year psychology students at the University of Glasgow were given participation credits for their time.

As it was determined that the attention check used in the first experiment was too open to interpretation, catch trials in Experiment 2 were adjusted for greater clarity. The same three scrambled dot and normal point-light display clips were presented randomly during the experiment, but this time, participants were asked *"Did you perceive a clear human form in the previous video?"*. The answer *"Yes"* was only correct in one of these three instances. It was determined that this question was less ambiguous, and therefore a more sensitive measure of attention than the check put in place in Experiment 1. Indeed, in this instance, when examining the task performance of those who responded incorrectly to the catch trials, we found that participants who failed to pass the attention checks (i.e., those who answered one or more of the questions incorrectly) performed significantly worse at the main task of identifying emotions (M= 41.74%) than participants who passed these checks (M= 48.96%): $t(43.35) = -2.18$, $p = 0.035$. Therefore, those who failed the attention checks (N= 23) in Experiment 2 were excluded from the analysis, leaving 53 participants in the final sample.

RESULTS

Perceived Emotion Results

In the Experiment 2, the average overall percentage recognition of the intended emotion, across all 53 participants was 48.96% (SD = 13.6). With a baseline of 20% established as chance level (from a 1 in 5 chance of selecting the intended emotion), after normality testing ($p > 0.05$), a one-sample t-test revealed that overall, participants identified the intended emotions at greater than chance level: $t(52) = 15.504$, $p < 0.001$, $d = 2.13$.

Examining each emotion individually, all intended emotion categories were recognised at a level significantly greater than chance. See Table 2.12 for a summary of recognition rates, and Figure 2.4 for a visualisation of the responses.

Table 2.12: A summary of recognition rates for each emotion category (means and standard deviations), and the results of inferential tests performed on the data to determine whether they were recognised at greater than chance level for the second experiment.

Emotion	Normally Distributed?	Average Recognition Rate (%)	Greater than Chance?	Test	Results
Neutral	No	58.49% *(SD= 30.99)*	**Yes**	One-Sample Wilcoxon Signed Rank Test	Z= 1381, p <0.001, d= 1.24.
Happy	No	47.17% *(SD= 28.02)*	**Yes**	One-Sample Wilcoxon Signed Rank Test	Z= 1326, p <0.001, d= 0.97.
Sad	No	58.02% *(SD= 22.88)*	**Yes**	One-Sample Wilcoxon Signed Rank Test	Z= 1431, p <0.001, d= 1.66.
Angry	No	41.98% *(SD= 22.88)*	**Yes**	One-Sample Wilcoxon Signed Rank Test	Z= 1326, p <0.001, d= 0.96.
Fearful	No	39.15% *(SD= 29.22)*	**Yes**	One-Sample Wilcoxon Signed Rank Test	Z= 1197, p <0.001, d= 0.66.

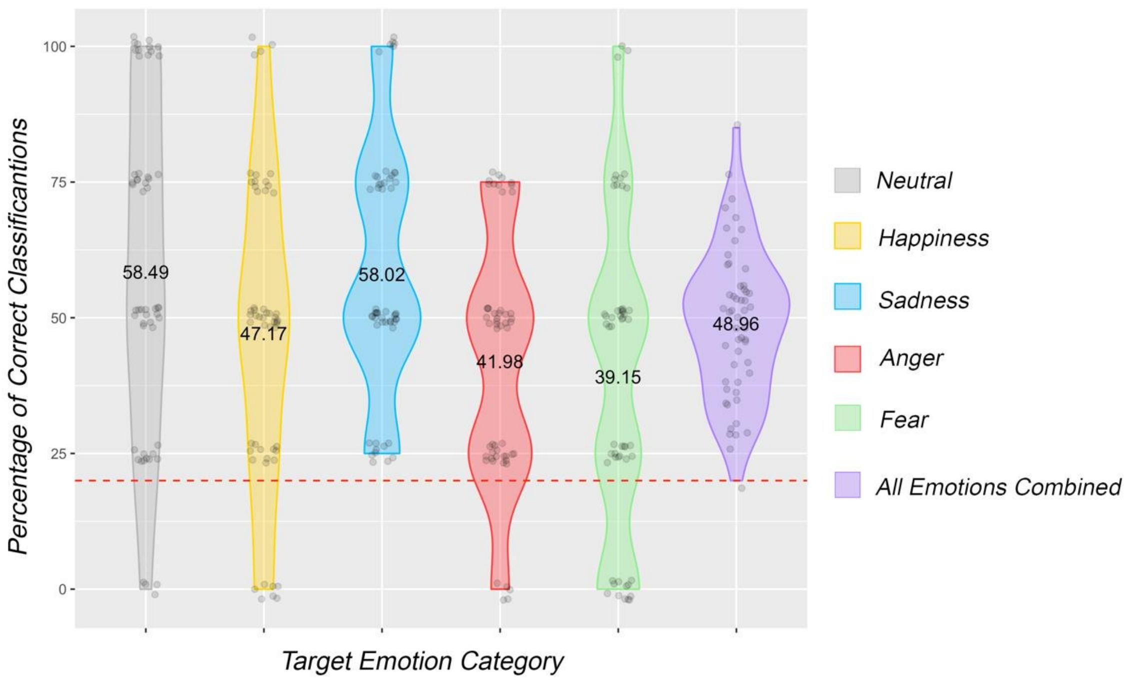


Figure 2.4: Violin plots depicting the distribution of recognition rates for each emotion category in the second experiment. Values presented in the centre of each of the violins represent the mean recognition rate, and the red dotted line indicates chance level of recognition. All emotion categories were recognised at greater than chance level.

While recognition rates in the second McNorm validation experiment were much higher than those found in the first experiment, a number of common misclassifications were still present. Participants frequently confused angry expressions with happiness (42.45%), and this occurred at a marginally greater rate than correct classification (41.98%). However, angry expressions were rarely perceived to portray any of the other emotion categories. Happy was also commonly confused with anger (24.06%), although at a lower rate than the converse. Fearful expressions were often confused with sadness in this sample (25.47%), and sad expressions were sometimes confused with fear (19.34%). A more detailed overview of classification and misclassification rates for the subset of clips from the full McNorm library can be found in Table 2.13.

Table 2.13: Correct Classification and Misclassification Rates for Each Emotion Category in Experiment 2.

		Perceived Emotion				
		Neutral	Happy	Sad	Angry	Fearful
Intended Emotion	Neutral	**58.49%**	12.74%	11.79%	4.25%	12.74%
	Happy	15.57%	**47.17%**	8.96%	24.06%	4.25%
	Sad	13.68%	7.55%	**58.02%**	1.42%	19.34%
	Angry	8.02%	42.45%	2.83%	**41.98%**	4.72%
	Fearful	15.57%	8.96%	25.47%	10.85%	**39.15%**

Summary

These results provide support for the primary hypothesis, as the intended emotion was recognised at greater than chance level for neutral, happy, sad, angry, and fearful expressions. Additional support was obtained for the secondary hypothesis, as recognition rates were found to vary across the different emotion categories; with neutral and sad expressions being recognised at the highest rate, and fearful expressions recognised with the lowest frequency.

In addition, these results suggest that the inconsistencies in recognition rates across all clips in the first McNorm validation experiment were not due to a lack of task attention, but rather, these results support the idea that stimuli clips isolated from the full library for further analysis are simply more representative of the target emotions than others in the full stimuli set.

Exploratory Analysis

Impact of Dance Experience on Recognition

As in the first experiment, participants completed a number of measures from an early unpublished version of the Gold-DSI to provide detail about their prior dance experience. A Kruskal-Wallis Test was conducted to determine whether participants' self-reported level of dance experience had an impact on recognition accuracy for the target emotions. In this experiment no participants classified themselves as a professional dancer, so there were only four levels of self-reported experience (non-dancer, beginner, intermediate, and advanced) used in this analysis. No significant differences were observed between the four groups (Chi-Square = 2.396, p = 0.494, df = 3).

As in the first experiment, a three-stage hierarchical multiple regression was conducted with average recognition rate as the dependent variable. The physical dance experience factor (a factor created as before from three measures in the early Gold-DSI) was entered at the first stage of the regression and the observational dance experience factor (a factor created as before from two measures in the early Gold-DSI) was added in the second stage. The number of styles a participant was familiar with was added at the final stage to create the maximal model. A detailed breakdown of the results can be found in Table 2.14 below.

Table 2.14: Summary of Hierarchical Multiple Regression Analysis for Factors Predicting Recognition Accuracy in the Second Experiment:

Model	Summary	Predictor	B	T	Std Err
Model 1	$F(1,51)= 0.031^{n.s.}$, $R^2_{Adj}= -0.019$, RSE = 13.73	Intercept	49.340		
		Physical Experience	$-1.154^{n.s.}$	-0.176	6.550
Model 2	$F(2,50)= 4.325^{*}$, $R^2_{Adj} = 0.113$, RSE = 12.81	Intercept	54.897		
		Physical Experience	$8.434^{n.s.}$	1.217	6.928
		Observational Experience	-24.626^{**}	-2.935	8.391
Model 3	$F(3,49)= 2.874^{*}$, $R^2_{Adj}= 0.098$, RSE = 12.92	Intercept	54.974		
		Physical Experience	$10.214^{n.s.}$	1.183	8.633
		Observational Experience	-23.909^{**}	-2.746	8.708
		Number of Styles	$-2.331^{n.s.}$	-0.351	6.636

Note: N = 50; *n.s.*p > .05, *p < .05, **p < .01, ***p< .001

Model 1: Percentage Recognition ~ Physical Dance Experience Factor

Model 2: Percentage Recognition ~ Physical Dance Experience Factor + Observational Dance Experience Factor

Model 3: Percentage Recognition ~ Physical Dance Experience Factor + Observational Dance Experience Factor + Number of Dance Styles a Participant Has Experience With

The hierarchical multiple regression revealed that at Stage one, the regression model was not a significantly better fit for the data than the null model F(1,51)= 0.031, p= 0.861, that the physical dance experience factor was not a significant predictor of recognition accuracy, and that this factor only explained 1.9% of the variation in recognition accuracy. Introducing the observational dance experience factor at Stage two contributed significantly to the regression model F(2,50)= 4.325, p= 0.019. The observational dance factor was found to be a significant predictor of recognition accuracy (p= 0.005) and explained an additional 9.4% of the variation in recognition scores. Finally, the maximal model was also found to explain the data significantly better than the null model F(3,49)= 2.874, p= 0.046. However, the number of dance styles a participant had experience with was not a significant predictor of recognition accuracy, and the inclusion of this factor decreased the amount of variance the model explained by 1.5%. Therefore, it appears that the most important predictor for task performance (as measured by average recognition rate) in this experiment was the observational dance factor (i.e., the frequency with which participants watch dance on screen (TV, phone, or computer) and in live settings). Together the three variables presented in the maximal model accounted for 9.8% of the variance in recognition rates.

Empathy Results

A Spearmans correlation was conducted to explore the relationship between self-reported trait empathy (as measured by responses on the Toronto Empathy Scale) and recognition of the intended emotion in this second McNorm validation experiment. No significant relationship was observed: r_s = -0.132, p= 0.345.

Intensity Ratings

In addition to assigning an emotion category to each movement sequence, participants also provided a measure of how intensely they felt the movements portrayed their chosen emotion on a slider scale from 0 (not intense) to 100 (very intense). The average overall intensity score provided by participants, across all emotions was 55.82 (± 11.58). The average intensity scores provided for clips in each emotion category can be found in Table 2.15.

Table 2.15: Average intensity ratings for clips in each emotion category in Experiment 2.

	Neutral	Happy	Sad	Angry	Fearful
Mean Intensity Score	46.68 *(± 16.65)*	61.48 *(± 14.06)*	53.48 *(± 15.16)*	63.26 *(± 15.13)*	54.19 *(± 14.04)*

From the mean intensity scores, it appears that portrayals of neutral expression were assigned the lowest intensity scores and portrayals of anger were assigned the highest intensity scores. A one-way ANOVA was conducted to explore the significance of these differences and several differences were observed in intensity scores across the different emotion categories: $F(5,318)= 9.18$, $p <0.001$, $d= 0.40$.

After correcting for multiple comparisons using a Tukey HSD test, angry expressions were perceived to be significantly more intense than fearful ($p= 0.016$), neutral ($p <0.001$) and sad ($p <0.01$) expressions. The difference in intensity scores for angry and happy expressions was not found to be significant ($p= 0.99$). In addition, happy expressions were perceived to be significantly more intense than expressions of neutrality ($p <0.001$). Although movements expressing happiness were perceived to be more intense than portrayals of sadness, this difference only approached significance ($p= 0.05$). No other significant differences were observed (all p values > 0.05).

Certainty Ratings

Participants also provided a measure of how certain they were of their emotion judgements on a slider scale from 0 (very uncertain) to 100 (very certain). The average certainty score for all participants across the entirety of the task was 53.61 (± 15.1). The mean certainty scores for each emotion category can be found in Table 2.16.

Table 2.16: Average certainty ratings for clips in each emotion category in Experiment 2.

	Neutral	Happy	Sad	Angry	Fearful
Mean Certainty Score	54.67 *(± 14.18)*	55.34 *(± 17.2)*	54.69 *(± 15.67)*	53.11 *(± 19.17)*	50.55 *(± 16.99)*

It appears that participants were least certain about their categorisation of fearful expressions and most certain about their perceptions of happiness. However, a one-way ANOVA revealed that these differences were not significant: $F(5,318)= 0.59$, $p= 0.71$, $d= 0.12$.

DISCUSSION

Emotion Recognition Results

Results from both validation experiments showed that participants recognised the intended emotion from movements in the McNorm library at greater than chance levels. In the first experiment, the average overall recognition rate across all participants was just above chance; at 26.9%. This rate was far lower than expected based on the results of previous work in this area (Castellano et al, 2007; Michalak et al., 2009; Wallbott, 1998; Paterson et al., 2001; Dael et al, 2012; Roether et al., 2009; Crane et al., 2013; Gross et al., 2010; Atkinson et al., 2004). It was first considered that this low recognition rate may have been the result of conducting a long duration, repetitive task in a non-laboratory setting (where environmental distractors could not be controlled for) with inadequate catch trials. However, further examination of the data revealed that this comparatively low recognition rate could perhaps be attributed to large variability in recognition rates across individual clips that made up each emotion category across the stimuli set. We suspected that some movement sequences in the full McNorm library did not convey as clearly and universally a specific emotion, compared to others. This idea was confirmed by the results from the second validation experiment. Experiment 2 examined responses from a new participant sample to a subset of the most well-recognised clips identified during Experiment 1. Recognition rates obtained in Experiment 2 were more in line with those reported in previous work, with an average recognition rate across all participants of 48.96%.

In addition to validating the McNorm library as a tool for social communication research, this result could be important to consider more broadly in the creation of future motion capture libraries. It is likely that when generating large samples of expressive movement data, there will be some specific sequences or movement clips that more effectively communicate the intended emotion better than others. These experiments present a novel solution to handling the occurrence of stimuli which do not perform as intended in the exploration of a particular research question. Perhaps in the future, researchers should examine their stimuli set as a whole and, if some individual stimuli do not serve their intended purpose, they could consider condensing the number of stimuli to create a new subset which can be exposed to further testing. This paradigm may result in more reliable materials for human movement research.

As predicted, based on recognition rates reported in the wider emotion recognition literature, fear was consistently recognised at the lowest rate. There are several relevant theories to discuss, and a wealth of potential explanations for this finding that would merit future investigation. For one, while the classic James-Lange, Cannon-Bard, and Schacter and Singer theories disagree about the specific mechanisms of the experience of emotions, these perspectives are all based on the core principle that physiology and emotions are intrinsically linked (Lange, 1885; James, 1884; Bard, 1928; Cannon, 1927; Schachter & Singer, 1962). The fear response, in particular, appears to be highly embodied; causing a uniquely recognizable set of physical manifestations within the human body. For many years, a fight-or-flight response was widely accepted as the dominant model for behavioural response to environmental stressors (Cannon, 1927). However, more recently, Barlow described an adaptive alarm model which includes, not only fight and flight, but also freeze as a potential response (Barlow, 2004; Schmidt et al., 2008). Therefore, even at the most basic physiological level, responses to fear-inducing triggers are highly idiosyncratic. It should be acknowledged that in creating the McNorm library, the method for portraying fear was decided solely by the movement performer. It is possible that this dancer's singular perspective was insufficient to capture the full scope of how fear manifests through movement, and may be different to how observers themselves experience fear – resulting in the lower levels of recognition. It would be useful for future work to examine the idiosyncrasies of fearful movement from the perspective of multiple dancers.

In addition, returning to the classic theories of emotion mentioned above, it is worth noting that they all indicate the need for some sort of trigger, or stressor, to produce the physiological arousal and elicitation of emotions (whichever order these responses appear in). It is possible that in the absence of an emotional trigger, this resulted in the dancer producing less authentic displays of the target emotions during the creation of the McNorm library. Results from facial expression research provide evidence that observers can differentiate between spontaneous (Duchenne) and posed (non-Duchenne) smiles (Krumhuber & Manstead, 2009; Trutoiu et al., 2014; Etcoff et al., 2021). Additional research in this area has found that fear is one emotion which observers are particularly sensitive to; in detecting inauthenticity from facial expressions (McLellan et al., 2010). However, to our knowledge, no research has explored observer sensitivity to the authenticity of different emotions expressed through human body movement. Based on these factors, and the low recognition of fear reported here and in previous work (Pasch & Poppe, 2007; Camurri et al., 2003; Atkinson et al., 2007; Dittrich et al., 1996; Dahl & Friberg, 2007), it is possible that observers are more sensitive to detecting inauthentic fear expressions from the body in motion. To explore this idea, it could be useful in the creation of future movement libraries to implement an emotion elicitation task prior to collection of the movement data and to more closely examine the physiological responses of the performer (e.g., heart rate, galvanic skin response) to determine how authentically 'felt' the emotions were during the movements (Val-Calvo et al., 2020). These movement sequences could then be compared to those obtained from an alternative set of movements (generated without a prior emotion elicitation session) in an emotion recognition task to explore the impact of posed versus genuine expressivity on observer judgements of emotion. Whether fear is particularly sensitive to this effect or not, it would be beneficial to social cognition research to examine the role of authenticity in recognition of emotional expressivity from the body across all of the basic emotions. Findings from such works could have a significant impact on methods for creating future libraries of emotionally expressive human movement stimuli. In sum, although largely speculative, these are two factors which may explain why fear was recognised with the lowest levels of accuracy in the McNorm experiments and in the wider emotional movement literature.

Dance Experience and Emotion Recognition Results

The data from both validation experiments did not provide support for the idea that dance experience facilitates emotion recognition abilities. Kruskal-Wallis tests performed on data from each experiment found that there were no significant differences in recognition accuracy across participants with different self-reported levels of dance experience. In addition to physical experience, increased experience of observing dance did not lead to improved recognition rates. In fact, in the second experiment, observation experience loaded negatively onto the recognition accuracy factor; thus, for the sample population in experiment two, increased amounts of observational dance experience led to a decrease in average recognition accuracy score.

Previous research has consistently reported that dance experience results in changes to behavioural outcomes (Stevens et al., 2010; Christensen et al., 2019), physiological responses (Christensen et al., 2016; Kirsch et al., 2016) and brain function and well as structure (Bläsing et al., 2012; Calvo-Merino et al., 2005; Cross et al., 2006; Cross et al., 2009; Kirsch et al., 2015; Hanggi et al., 2010; Orlandi & Proverbio, 2019). Therefore, it is surprising that participants who reported having prior dance experience did not show higher recognition of the intended emotion in the McNorm library validation studies. However, it should be noted that in previous work, experienced dance participants were typically expert performers (recruited from professional companies). In this validation study, there were only two participants who categorised themselves as professional dancers, and both were in the sample population for Experiment 1. Therefore, the majority of dancer participants in these experiments represent an often-neglected group in this area of research: amateur dancers. As so few studies explore the impact of informal or hobby dance experience on behavioural outcomes, the role of amateur experience on behavioural factors, like emotion recognition capabilities, remains unclear. It is possible that amateur dance training, regardless of duration, is not sufficient to facilitate the improved emotion recognition capabilities seen in expert performers, and this may explain the results obtained from the McNorm Experiments; at least in part. However, until future work specifically examines the role of non-professional dance training, and how this differs to professional training, this interpretation remains speculative.

Further, the role of observational dance experience on emotion recognition accuracy is unclear. While the results from the second McNorm experiment suggests that observational experience has a negative effect on recognition of the target emotion, this result should be interpreted carefully as this effect was not observed in the first experiment. Based on the previous literature and the lack of generalizability across the McNorm experiments, it is likely that this finding was specific to the sample population recruited for Experiment 2. There are a number of potential explanations for this result. First, it should be acknowledged that data for Experiment 2 was collected during the Covid-19 pandemic. The observational experience measures from the version of the Gold-DSI used in these experiments related exclusively to the frequency with which participants engaged with dance performances. During national lockdowns which took place across the globe, access to dance performances was extremely limited, and it was not clearly explained that responses to the Gold-DSI should reflect participants' behaviour during normal circumstances. Therefore, it is possible that participants were under-reporting the frequency with which they would engage with dance performances in non-Covid times, and this may have undermined the usefulness of these results. On this note, it is likely that questions relating purely to the frequency of attending performances do not fully reflect engagement with dance as an artform. There are a number of factors that may impact engagement beyond simply attendance, such as enjoyment, emotional response, and the extent to which watching dance makes the observer want to move themselves. Many of these elements of engagement can be examined using the published version of the Gold-DSI (Rose et al., 2020; which was unavailable at the time of data collection for the McNorm experiments), therefore future research should make use of the full questionnaire, and of other more general aesthetic engagement questionnaires (e.g., the Aesthetic Experience Questionnaire; Wanzer et al., 2020, and the Aesthetic Responsiveness Assessment; Sholtz et al., 2020), to examine more completely what an individual may gain from engaging with art, and dance specifically, as a spectator, and how this may influence social behaviours.

Empathy Results

Several studies have reported a relationship between individual differences in empathy and recognition of affective information. This has been observed in identification of emotion from facial expressions, voices, and body movement (Besel & Yuille, 2010; Golan et al., 2007; Fridenson-Hayo et al., 2016). However, across both experiments of this study, we found no evidence of a relationship between empathy scores and recognition of the intended emotion. In this study, we used the Toronto Empathy Scale to generate a measure of trait empathy for each participant. This measure is composed of 16 items and covers several aspects of the empathic response in everyday scenarios. However, it is possible that the reduced format of this scale did not produce a robust enough measure of empathy. In this study, the inclusion of a more detailed empathy measure may have been detrimental to the results. The first validation study, which contained all 73 items of the full McNorm library, was around 45 minutes in duration and required a high level of cognitive demand (due to the presentation of highly repetitive stimuli). Therefore, it was decided that a short measure of empathy should be included to limit interference of participant fatigue on the results. However, in future work it may be more useful to include a more detailed measure of empathy to explore this relationship further.

Intensity and Certainty Ratings

Although no significant differences in certainty scores emerged across the different expression categories, participants tended to report they were most certain of their judgements when assigning emotion labels to portrayals of happiness and anger. It is interesting to note here that participants showed the highest accuracy in recognition of sad expressions in the first study, and the greatest levels of recognition for neutrality and sadness in the second study. In addition, when examining misclassification rates across both studies, participants frequently confused happy and angry expressions with one another. It is possible that participants believed it was easier to assign expression labels to more intense emotion categories. The data do provide support for this assumption, as happy and angry expressions were perceived to communicate emotion with the greatest intensity, compared to sadness, fear and neutrality. Unsurprisingly, portrayals of neutral emotion were assigned the lowest intensity ratings. However, it is surprising that portrayals of neutrality received an intensity score of roughly 50 (on a scale of 0-100, denoting the spectrum from *not intense* to *very intense*). Neutrality in this study was described to participants as the absence of expression of any specific emotion. Therefore, it is unclear why participants did not provide an average intensity score closer to zero when asked to respond to the question of: "*How intensely is the emotion being expressed?*".

The performance of certain dance styles (classical ballet in particular) is designed to communicate a narrative or meaning to observers, and this often relies on the ability to arouse emotion from an audience. In the absence of other social cues like verbal communication, and facial expressions (which audience members at the back of a dance venue may find difficult, or impossible, to perceive), the dynamic body is the dancer's main tool for communicating expression, and formal dance training responds to this need by teaching dance students how to imbue the performance of whole-body movements with emotionally salient information. A professional ballet dancer was recruited to generate the movement sequences in the McNorm library, and as a result of her extensive experience in performative dance it was likely that the dancer found it difficult to create and perform truly neutral movement sequences. This is likely a problem for all dance-based movement libraries, as creating and performing non-expressive choreography directly contradicts training and defies the communicative nature of performative dance. This may explain why participants assigned surprisingly high intensity ratings to the neutral expressions in this set of validation studies and is an issue worth addressing when creating future dance movement libraries.

Benefits of the McNorm Library

In creating the McNorm library, several methodological issues from previous work were identified and addressed. First, the McNorm library contains movements which depict expression of four basic emotions (happy, sad, angry, and fearful) in addition to neutral, or non-expressive, movements. Recent findings from Jack and colleagues suggest that there are only four universally recognised facial expressions ('*happy*', '*sad*', '*surprise/fear*', and '*disgust/anger*') (Jack et al., 2016). However, it is currently unclear whether this framework for recognition of facial expressions applies to expression through the human body in motion. To begin answering this question, it would be beneficial for researchers working in this area to expand their research questions to explore perceptions of more than positive versus negative valence, to first tackle the basic, universal emotions (e.g., happiness, sadness, anger, fear) before exploring more socio-culturally complex emotions. This approach would result in greater methodological consistency and would improve the ease with which results can be compared across studies in this domain.

Further, the McNorm library addresses the question of how emotion is portrayed through whole-body movement while accounting for both parties involved in the dyadic nature of expressive communication (i.e., the movement performer and the movement observer). This approach confers significant advantages to prior movement libraries which focus exclusively on perceptions of observers when assigning movements to expressive categories.

In addition to the benefits addressed above, the McNorm library is specifically well suited to use in visual attention research. Previous movement libraries have a number of issues which impose unique limitations on their utility in visual attention experiments; depending on the research question. For example, several movement libraries address the limitations highlighted in the previous section (e.g., see work by Shafir and colleagues, who present an elegant series of papers which account for the intricacies of the movement observer-movement performer dyad) and are undeniably useful for social perception research. However, few libraries of this nature depict the desired movements in the form of point-light displays. It has been noted in previous visual attention research that a number of visual cues can influence perceptual and cognitive processing. Different levels of luminance, for example, have been found to specifically impair biological motion perception (Burton et al., 2016). This finding is of particular relevance to the present authors, and to other researchers exploring emotion perception from human movement. Therefore, in this regard point-light displays, like those in the McNorm Library, confer a significant advantage over full-light displays for research questions related to visual attention, as the process of creating such stimuli allows for greater control over the visual features of the display. In addition to these low-level cues, higher-order superficial cues can also significantly influence perceptual processing. For example, it has been suggested that our attentional processes have evolved to favour attractive stimuli (Dixson et al., 2014) and sexual interest has been found to drive visual attention and influence fixation patterns when observing human figures (Hall et al., 2011; Heron-Delaney et al., 2013). Although in the majority of full-light display movement libraries the face is obscured, and clothing and background can be simplified and held constant across portrayals, point-light displays (which are devoid of these features and of further morphological information about the performer) present an easier to use, and more appropriate stimulus for visual attention research.

Finally, the McNorm stimuli set confers one further advantage over previous movement libraries. In the first experiment, for the full stimuli set we reported a high miss-rate for recognition of some emotion categories. However, when placing these results in the context of the previous literature, the subsample of movement sequences validated across Experiments 1 and 2 present a competitive recognition rate.

Based on previous emotion recognition and stimuli validation experiments, average recognition accuracy scores ranged from 59.98% to 64.95% (depending on the emotion category). A more detailed breakdown of these calculations and the data used can be found in Tables S.1 and S.2 of the Supplementary Materials. The average recognition rates obtained for the McNorm Library for these same emotion categories ranged from 38.58% to 59.01%. While average hit-rates for the McNorm Library are comparatively lower, it is worth noting that the average values obtained from our results have been calculated from two different participant samples. This signifies an advantage of the McNorm Library, as other libraries are rarely, if ever, examined for test-retest reliability. Therefore, the hit-rate for recognition is not considerably reduced in comparison with existing movement libraries, and the successful test-retest values obtained across both participant samples has generated a set of stimuli that reliably invoke perceptions of the target emotions.

As such, we believe the McNorm Library presents a reliable and useful tool for the field of social perception.

Limitations of the McNorm Library

One limitation of the McNorm library is that the portrayals of each emotion category were based on the interpretation of only one dancer. Based on the type of training (i.e., style of dance, school of dance) and cultural factors there may be variations in how dancers choose to communicate different emotions through their movements. In addition, while the dancer aimed to communicate consistent portrayals of each specific expression category, it is unclear the extent to which she perceived her performance to communicate that emotion and only that emotion. It has already been noted that prior dance experience has a significant impact on a number of behavioural and neuropsychological outcomes, and therefore it is possible that unique experiences of the dancer used to create this movement library may have influenced various elements of her performance. For example, there may have been some specific movements within the sequences that the dancer personally associated with a specific emotion or mental state and this may have coloured the performance of movements with a different intended expression. To limit the impact of subjective interpretations in creating movement stimuli that are representative of emotional expressions, future movement libraries of this nature should include recordings of more than one dancer.

Finally, the McNorm library contains sequences derived from only two types of western dance (ballet and contemporary dance). Therefore, it is unlikely that these results can be applied to different, more abstract forms of western dance (e.g., modern dance, jazz) and certainly may not be applicable to non-western dance; given the previously discussed studies which emphasise cultural differences in the identification and communication of emotional expression (Archer, 1997; Jack et al., 2012; Jack et al., 2009; Van Dyck et al., 2017).

Using the McNorm Library in Future Research

Both versions of the McNorm library, (the full, original sample, and the subsample tested further in experiment 2) have a specific utility for researchers depending on the nature of the research question to be explored. For future experiments focused on emotion recognition, specifically, the present authors advise making use of only the subset library (the 20 clips subjected to further analysis in Experiment 2) rather than the full McNorm Library, as (per its intended purpose) this sample of movement clips has been identified as the most representative of the target emotions; from both a perceptual and performative perspective. However, the full stimuli set will be useful to researchers who wish to access a larger amount of data to test other hypotheses. For example, original sample would be suitable for use in biological motion and visual aesthetics research, and will also be useful for expanding the volume of data available to researchers who wish to explore the communicative value of movement through machine learning. As such, the full McNorm Library (and all 3D-position data) is also available, and can be found on the OSF (https://osf.io/458sq/).

For this validation study, the movement recordings were converted into PLD videos to explore the relationship between movement intention and movement perception, but the raw coordinate data could be used in future studies to examine the interplay of quantifiable kinematic movement parameters and observer perceptions in emotion recognition, among other constructs. Indeed, this question is currently being examined in a separate experiment by the authors of this paper; the results of which will be shared in a future manuscript. Furthermore, dance is gaining attention in aesthetics research. In this validation study, no measure of aesthetic judgements were collected. This makes the McNorm library, in its current state, unsuitable for exploring how emotion expression contributes to the aesthetic value of dance movement. However, there are plans to rectify this issue in future work by introducing measures of *beauty*, *interest,* and/or *liking* to the experimental procedure. This will allow these PLD movement sequences to be used in future research which explores the impact of performative elements of movement on emotional recognition and aesthetic evaluation.

Creation of the McNorm library is also part of a larger scale project that will be conducted by the authors of this paper. The overall aim of this project is to create a framework for the expression of emotion to others in our social environment through dynamics of the human body in motion. It is hoped that the principles of this framework could be used to inform choreographic practices (to allow dance-makers to create movement sequences which are known to inspire expressive states in observers) and could be used by dancers in their performances (by providing direction about which performative elements of motion are most important for communicating expressivity or an emotionally charged narrative to audiences). Finally, beyond the scope of aesthetics and performative dance, it is also hoped that these principles can be applied to the field of artificial intelligence. For example, a more detailed understanding how humans communicate emotion to others may help technicians, designers, and computer scientists to create more expressive, and thus likeable, artificial interaction partners (e.g., robots, virtual/augmented reality avatars) which move, and ultimately behave, in a more anthropomorphic manner.

CHAPTER 3

MOVING ME, MOVING YOU: INVESTIGATING WHETHER EMOTIONAL EXPRESSIVITY, EMPATHY, AND PRIOR EXPERIENCE SHAPE WHOLE-BODY MOVEMENT PREFERENCES

Moving me, moving you: Investigating whether emotional expressivity, empathy, and prior experience shape whole-body movement preferences

Rebecca A. Smith & Emily S. Cross

ABSTRACT

Aesthetics shape and colour almost every aspect of our daily lives, from the products we interact with, to the clothes we wear, to the design of our homes and cities. However, many people associate aesthetics with art, and an historical academic interest in the factors which shape the experience of engaging with art has yielded rich insights into our understanding of the value and ubiquity of empirical aesthetics. While the majority of such research to date has focused on music and the visual arts, there is a growing interest in the aesthetics of dance and human movement among empirical aesthetics researchers. In the present study, we sought to examine how individual differences in empathy and previous experience influence aesthetic evaluations of dance movements. Observers with varying degrees of dance experience completed a self-report measure of empathy, and rated a series of emotionally expressive whole-body movements from the McNorm dance movement library (Smith & Cross, 2022) in terms of beauty and liking. Participants demonstrated a general preference for emotionally expressive movement sequences, while specific types of emotional expressivity influenced liking, but not beauty, judgements. Additionally, both prior dance experience and higher levels of trait empathy resulted in more positive aesthetic evaluations of the McNorm Library dance clips. We consider implications of these results for empirical aesthetics and social perception research, and discuss how empirical aesthetics research in this area may be of interest, or use, to dance practitioners.

INTRODUCTION

Aesthetics have long played a key role in the natural world. For example, much of the world's flora and fauna have evolved specific morphologies designed to attract prey, mates or pollinators, or to deter predators (Rosenthal et al., 2001; Drouilly et al., 2021; Hoyle et al., 2018). Additionally, the applications of aesthetics can be found in the various technologies we interact with every day. For example, the design of our smart phones, websites and social media platforms are constantly evolving in response to aesthetic evaluations and preferences of users (Le-Hoang, 2020; Oyibo & Vassileva, 2017; Schreiber, 2017). In addition, aesthetic considerations have been found to play a significant role in the creation and implementation of renewable energy resources (Breukel et al., 2016; Johansson & Laike, 2007). Windfarms, for example, require no fuel and can contribute substantially to the energy needs of a populace, but have been heavily criticised for their visual impact on the surrounding landscape (Vlami et al., 2020). Therefore, it has been suggested that by altering the aesthetic design of these wind turbines, and reducing the extent to which locals perceive them to be "eye-sores", this may result in more positive planning progress outcomes and, therefore, less reliance on environmentally damaging energy sources. The pervasive nature and diversity of applications for aesthetics in our everyday lives demonstrates the importance of empirical aesthetics research in how we shape and understand the world around us. However, aesthetics inarguably has a special relationship with the creation and appreciation of artworks, and it has been argued that engaging with the aesthetic experience via art plays a key role in human personal development and wellbeing (Maslow, 1981).

Recent world events have highlighted the importance of engaging with the arts. The Covid-19 pandemic coincided with a dramatic increase in the number of subscriptions to popular streaming services (e.g., Netflix, Amazon Prime Video, Disney+), while book sales also soared, despite physical shops being closed for the majority of 2020 across most of the world (Faughnder, 2021; Lee, 2020; Flood, 2021). In addition to a renewed interest in visual media and reading, people also increasingly turned to home-based art activities during the peak of Covid-19 restrictions. In a recent longitudinal experiment, researchers observed that over 22 weeks during the height of the pandemic in the United Kingdom, high volumes of participants turned to home-based art activities (e.g., painting, writing, playing music) to gain a sense of escapism, and to cope with negative emotions during the most taxing periods of lockdown (Bu et al., 2022; Mak et al., 2021). Considered together, these findings, have further solidified the importance of engaging with the arts for our wellbeing, and highlight the need to develop our understanding of factors which mediate the aesthetic experience. Several factors have been found to mediate how we engage with art and aesthetic experiences, but individual differences in empathy, and previous aesthetic experience or training appear to be particularly influential factors.

Behavioural and neuroscientific evidence suggests that individual differences in empathy influence aesthetic processing of works of art. This relationship can be observed in a number of studies which have been conducted across a variety of artistic mediums, including visual artworks (e.g., paintings, photographs), music and the performing arts. In one such experiment, participants provided a measure of 'emotional contagion', reflecting their ability to "feel into" the emotions of others as an indirect measure of general empathetic ability, before evaluating a series of representational and abstract artworks while facial electromyography and skin conductance data were collected (Gernot et al., 2018). Participants who reported experiencing greater emotional contagion demonstrated greater physiological reactions, and provided more positive aesthetic appraisals of the artworks (i.e., "being moved", valence, and interest). This suggests that empathetic skill mediates aesthetic processing of visual artworks, and leads to greater emotional arousal in response to these aesthetic stimuli. Additional evidence for the role of empathy in aesthetic evaluation of visual artworks can be found in recent work by Wilkinson and colleagues (2021), who compared visual processing and aesthetic judgements of artworks in individuals with high or low self-reported levels of empathy. These authors reported that individuals recording higher empathy scores were significantly faster to respond to all works of art than those with lower scores. Additionally, these authors found that high empathy individuals showed higher aesthetic interest and emotional engagement in response to the more emotionally evocative artworks (Wilkinson et al., 2021). Taken together, these behavioural results suggest that individuals with higher trait empathy might process aesthetic material more efficiently, and provide more positive appraisals of artworks.

Finally, neuroscientific evidence for the relationship between empathy and aesthetic appreciation can be observed in a recent study conducted by Ishizu and Zeki (2017). Here, participants observed a series of images which had been classified as either beautiful and sad, beautiful and happy, neutral, or ugly. Neuroimaging data was collected across a series of brain regions, but with a particular focus on activity of the medial orbito-frontal cortex (mOFC), a brain region believed to play a key role in the experience of beauty. Greater activity of the mOFC was observed in response to both types of beautiful stimuli (sad and happy), than in response to the neutral or ugly stimuli. Through functional connectivity analyses, the authors further found that this activity was mediated by activation of the supplementary motor area and middle cingulate cortex; areas which are known to be involved in experiencing empathy when others are experiencing sadness (Ishizu & Zeki, 2017). These results suggest that emotionally salient artworks prompt coactivity in both the empathy and aesthetic circuits of the brain, indicating a biological link between these sociocognitive processes.

In addition to visual artworks, a number of researchers have explored the relationship between empathy and aesthetic processing in response to musical compositions. For example, Vuoskoski and colleagues observed that participants with higher self-reported empathy scores provided higher liking scores and reported experiencing more intense emotional responses to sad music than individuals with lower trait empathy scores (Vuoskoski et al., 2011). In addition to general empathy, a number of medium-specific types of empathy have been proposed, and these are believed to specifically impact aesthetic processing of medium-specific artworks. The concept of 'music empathy' was described by Kreutz, Schubert, and

Mitchell, who suggested that musicians and music lovers possess a series of specific empathy-related skills that facilitate expertise related differences in the cognitive processing of music (Kreutz et al., 2008). Support for the influence of domain-specific empathy, in particular for 'music empathy', on aesthetic processing can be found in a study conducted by Garrido and Schubert (2011). In this experiment, music students completed two empathy questionnaires; one which captured general trait empathy, and another which provided a measure of 'music empathy'. After completing these questionnaires, participants listened to a series of musical excerpts deemed by a separate group of participants, to be sad. A significant positive correlation was observed between general empathy, 'music empathy', and liking scores. These results suggest that both general trait empathy, and domain-specific empathic skills can influence value judgements for aesthetic stimuli.

While the majority of empirical aesthetics and neuroaesthetics research has focused on the visual arts (e.g., paintings, photographs), and music, growing evidence suggests that domain-specific empathetic skills involved in the processing of human movement stimuli are also at play. The concept of 'kinaesthetic empathy' is believed to play a specific role in aesthetic value judgements for whole-body dance movements. Calvo-Merino and colleagues, for example, were the first to obtain neuroscientific evidence for the relationship between kinaesthetic empathy and aesthetic processing. In this experiment, participants observed a series of whole-body dance movements while fMRI data were collected, and in a follow-up session, the same participants were asked to rate each of the dance clips in terms of interest and liking. They found the clips which (as a consensus) were rated more positively in terms of liking scores produced greater activation of the premotor cortex than clips with lower liking scores. The authors suggest these results indicate an inherent link between kinaesthetic empathy (as evident from activation of the motor areas) and aesthetic evaluation (Calvo-Merino et al., 2008). While the work carried out by Calvo-Merino and colleagues (2008) was conducted with dance-naïve participants, concepts related to kinaesthetic empathy (i.e., movement simulation, mirroring, embodiment) are often explored within participants who have received some form of dance training. This work reveals a wealth of evidence to support the relationship between movement experience and greater aesthetic appreciation of whole-body movements.

Even short-term dance training can influence aesthetic perception of dance sequences, as shown in a study conducted by Kirsch and colleagues (2015). Participants took part in a 4-day dance training session, where each day they physically rehearsed one set of movement sequences, passively observed another set of movements, and listened to the music accompanying a third set of movements; a fourth set of movements remained untrained to allow for comparison. fMRI data were collected prior to the training sessions, and immediately following the training period, along with measures of movement capability and aesthetic value judgements. The participants liked the movements which they had both physical and observational experience with significantly more than the music-only sequences and the untrained sequences. In addition, neurocognitive changes were observed following this dance training.

Prior to the sessions, aesthetically-relevant brain activity was centred around the dopaminergic reward system structures. However, after the training sessions, greater activity was observed in areas of the action observation network (AON), which are believed to be involved in multisensory integration, and the processing of biological motion (Kirsch et al., 2015). These results suggest that even short-term dance training can result in changes to subjective aesthetic value judgements, and can even lead to reorganisation of the corresponding brain structures responsible for aesthetic evaluation of human dance movement.

Considered together, these results suggest that enhanced kinaesthetic empathy, and other behavioural and neurocognitive changes which occur as a result of dance training, can modulate the aesthetic processing of human movement. However, it is important to note here that research directly exploring these relationships remains fairly limited and often focuses only on the impact of short-term intensive dance training. The role of long-term amateur, or hobby, dance training (which is more broadly the experience of dance across the general population than expert or empirically-driven training) has received limited research attention, and thus its role in mediating aesthetic evaluation of movement stimuli is still unclear. Therefore, to explore the role of individual differences in empathy, and previous dance experience of various levels on aesthetic value judgements for human movement stimuli, we were interested in understanding how various degrees of motor and visual dance experience, as well as empathy differences, shape people's perceptions of dance movements. To do this, we asked participants to complete a condensed version of the Goldsmith's Dance Sophistication Index (Gold-DSI; Rose et al., 2020) and the Toronto Empathy Scale (TES: Spreng et al., 2009) before rating the McNorm Library dance movement clips (Smith & Cross, 2022) in terms of beauty and liking.

An additional aim of this study was to explore the role of emotional expressivity on aesthetic value judgements. Affective arousal seems to be a particularly important facet of the aesthetic experience, as more emotionally salient works of art appear to prompt heightened physiological and behavioural responses in observers, and produce more positive appraisals. Support for this notion can be observed in a recent experiment conducted by Christensen and colleagues. In this experiment, participants rated a series of whole-body dance sequences that contained either emotionally expressive or non-expressive movements in terms of liking and expressivity while galvanic skin conductance data was collected. The authors found that the clips rated as more expressive were also liked the most by participants. In addition, emotionally expressive clips prompted greater galvanic skin responses, particularly in those with arts experience (Christensen et al., 2021). These results suggest that observers prefer emotionally expressive movement stimuli, over non-expressive movement, and that this effect may be more pronounced in arts-experienced participants.

While it seems that observers show a preference for emotionally expressive, over non-expressive, dance movement stimuli, it is unclear the extent to which portrayals of specific emotions shape these judgements. To our knowledge, no research has explored observer preferences for movements communicating different specific emotions. However, anecdotal evidence suggests that one of two patterns may emerge when examining aesthetic preferences for happy, sad, angry, and fearful dance sequences. First, we could consider the idea that "*what is beautiful is good*". This concept can be traced back to the work of early philosophers like Sappho and Plato (Dayan, 2011), but more recent evidence for this effect is provided by physical attractiveness research (Lorenzo et al., 2010; Little et al., 2006). It is possible that this principle may work in the inverse for processing of aesthetic material; in that "*what is good is beautiful*", which may lead participants to demonstrate a preference for artworks which depict the "good", or positive emotions. On the other hand, evidence from music research suggests that certain individuals derive a great amount of pleasure from sad musical compositions (Vuoskoski et al., 2011), and it should be acknowledged that many of the most popular performative ballets (e.g., *Swan Lake*, *La Sylphide*, *Giselle*) depict narrative tragedies. These factors may suggest that dance movements which depict emotions with a negative valence (and sadness in particular) may be rated more favourably in terms of beauty and/or liking by observers. Therefore, a final aim of this experiment is to explore whether the specific emotional expressivity present in movement influence observers' aesthetic evaluations of whole-body dance sequences.

Therefore, the aims of this study were threefold. First, we wished to examine whether individual differences in empathy mediate aesthetic appreciation of human dance movements. It was predicted that participants with higher trait empathy (as measured by the TES; Spreng et al., 2009) will provide more positive aesthetic value judgements than those participants with lower trait empathy. It was predicted that this would occur for both beauty and liking ratings. The second aim of this experiment was to explore whether different forms of previous dance experience influence the perceived aesthetic value of dance movements, using a shortened version of the Gold-DSI (Rose et al., 2020). In line with existing research, we predicted that participants with previous dance experience would provide higher average beauty and liking ratings than those without dance experience. We further predicted that physical dance experience would coincide with the largest increases in aesthetic ratings, while observational dance experience would result in more positive beauty and liking ratings. The final aim of this study was to explore the extent to which the communication of different basic emotions influences the aesthetic value of dance movements for observers. Finally, we predicted that participants would assign emotionally expressive dance movements more positive aesthetic appraisals than non-expressive movements. Due to a lack of research exploring the contributions of specific emotional states on aesthetic value judgements, no specific predictions were made with respect to beauty and liking ratings for happy, sad, angry and fearful movements, so analysis of these data will be exploratory in nature.

METHODS

ETHICS STATEMENT

All study procedures were approved by the College of Science and Engineering Ethics Committee (University of Glasgow, Scotland, Approval Number: 300200033).

OPEN SCIENCE STATEMENT

Consistent with open science practices widely adopted within psychological research (Open Science Collaboration, 2012), we report all manipulations and all measures in the study. In addition, following open science initiatives (Munafò, 2016), the data, stimuli, and analysis code associated with this study are freely available on the Open Science Framework. By making the data available, we enable others to pursue tests of alternative hypotheses, as well as more exploratory analyses.

PARTICIPANTS

Previous studies exploring value judgements for aesthetic stimuli have participant sample sizes which range between 21 and 155 participants (Ishizu & Zeki, 2017; Kreutz et al., 2008). Such studies exploring aesthetic perception of dance movement stimuli, more specifically, used participant samples which ranged between 6 and 41 (Calvo-Merino et al., 2008; Christensen et al., 2021). As the average number of participants tested in the prior aesthetics literature is around 81 participants (Gernot et al., 2018; Wilkinson et al., 2021; Ishizu & Zeki, 2017; Vuoskoski et al., 2011; Kreutz et al., 2008), and the average for dance-specific research is around 22 participants (Calvo-Merino et al., 2008; Cross et al., 2011; Kirsch et al., 2015; Kirsch & Cross, 2018; Christensen et al., 2021), we aimed to recruit a larger-than-average sample for this experiment to allow the results to be more comparable to the sample sizes used to explore aesthetic perceptions of other mediums. Therefore, we aimed to collect data from a minimum of 50 participants.

A total of 444 participants launched the webpage to begin the experiment. Upon excluding participants who did not complete the task in full (N=389), the final sample size was 55. The experiment was created in formr (Arslan et al., 2020) and was advertised through the University of Glasgow subject pool, and across various social media platforms.

STIMULI

The stimuli used in this experiment were selected from the McNorm Movement Library. These movement clips have been validated for emotion recognition in a series of prior experiments by the authors (Smith & Cross, 2022).

Clips from the full McNorm Library (N=20; 4 Neutral, 4 Happy, 4 Sad, 4 Angry, 5 Fearful) used in this aesthetic judgment experiment were selected as they were found to be most representative of their intended emotion category. In the creation of the McNorm Movement Library, movement sequences were devised by a professional dancer who was instructed to exclude any movements which are tied to a particular emotional state (e.g. shaking fist for anger, arms thrown in the air for joy), or movements that are associated with specific emotionally evocative narratives (e.g., "dying swan" arm movements from the traditional Swan Lake choreographic repertoire which, particularly in participants with previous dance experience, may invoke feelings of grief or sadness). These pieces of choreography were then performed five times, each time with the aim to communicate a different emotion (happy, sad, angry, fearful, and neutral/non-expressive as a control) to an imagined spectator. Stimuli considered to be most representative of these emotion categories were decided based on recognition rates for the emotion intended to be communicated by the movement performer (i.e., clips which were intended to be happy by the movement performer which received the highest proportion of 'happy' responses in the emotion validation experiments). A more detailed breakdown of the emotion recognition rates for each clip used in this experiment can be found in Table 3.1.

Table 3.1: Summary of the whole body dance movement stimuli from the McNorm Library used in this experiment. The table provides information about the name of the clip, the emotion category assigned to the clip, and the average rate of recognition for the intended emotion for the clip across the two validation experiments in Smith & Cross (2022).

Clip Name	Emotion Category	Average Emotion Recognition Rate (%)
frappe_N	Neutral	52.42
jete_N	Neutral	58.25
plie_N	Neutral	60.13
tendu_N	Neutral	61.19
balanch_H	Happy	47.59
char_H	Happy	53.30
onegin_H	Happy	49.70
pirouette_H	Happy	50.76
fondu_S	Sad	67.96
free_mov_S	Sad	55.36
plie_S	Sad	64.08
rdej_S	Sad	48.64
balanch_A	Angry	42.87
bitter_A	Angry	43.93
char_A	Angry	65.12
onegin_A	Angry	35.15

bitter_F	Fearful	30.10
ever_con_F	Fearful	43.64
ice_F	Fearful	43.47
partner_F	Fearful	30.10

The 20 movement sequences outlined in the previous section were uploaded to Vimeo and were displayed through formr using embed codes. In Vimeo, the embed codes were manipulated to remove the standard online video handles (video title, uploader information, suggested videos, and Vimeo branding). The video controllers were also removed, and the video was set to loop automatically until participants had submitted their responses for the item.

PROCEDURE

Participants were directed to the formr link, which in turn led them to the experiment page. After reading the study information and consenting to participate (in accordance with University of Glasgow and British Psychological Society guidelines), participants completed two demographics measures (age, gender) and two questionnaires: a series of questions extracted from an early, unpublished version of the Gold-DSI (Rose et al., 2020), and the TES (Spreng et al., 2009).

The selected questions from the early Gold-DSI allowed participants to provide detail about their previous dance experience, covering both formal and informal experience, and both physical and observational experience. The version of the Gold-DSI used in this experiment can be found on the OSF (https://osf.io/hvmwn/). The TES is a 16-item questionnaire which includes a variety of measures exploring emotional and cognitive empathy, used to generate an individual empathy score (ranging from 0 – 64) for each participant, where higher scores are reflective of higher empathy.

Following the completion of these measures, participants began the aesthetic judgement experiment wherein they watched each of the 20 emotionally expressive movement sequences (presented in random order) and after observing each sequence were asked to provide an answer to the question "How beautiful did you find this movement clip to be?" on a 100-point slider scale from "Not beautiful" (0) to "Very beautiful" (100), and a response on a similar 100-point slider scale to the question of "How much did you like this movement?" from "Not at all" (0) to "Very much" (100). The experiment took between 30 and 45 minutes for the volunteers to complete, and first-year psychology students at the University of Glasgow received 3-4 participation credits for taking part.

As this experiment was conducted using an online sample, three catch trials were included to test whether participants were paying sufficient attention to the stimuli and task. For this purpose, participants were shown images of animals (kitten, puppy, and toad/frog), and on the following page they were asked to type which animal they had observed a picture of on the previous page. These images were interspersed randomly within the experimental trials. No participants who completed the task in full failed these attention checks, and therefore no additional participants were excluded from the final sample or analysis.

RESULTS

Beauty Results

Participants were asked to provide beauty ratings for each of the clips on a 100-point slider scale from 'Not Beautiful' (0) to 'Very Beautiful' (100). The average beauty rating provided by participants (across all 20 clips) in this study was 60.98 (SD= 12.74).

Comparison of Beauty Ratings Across Different Emotion Categories

Each of these dance movement sequences were associated with an emotion label (happy, sad, angry, fearful, and neutral as a non-expressive control), based on both the expressive intentions of the movement performer and the emotion perceptions of movement observers (validated in Smith & Cross, 2021). A one-way ANOVA revealed significant differences in beauty ratings across clips from different emotion categories: $F(4, 270)= 20.6$, $p< 0.001$, $\eta^2 = 0.234$. This difference was explored in more detail through post-hoc Tukey testing, where it was found that clips from the neutral/non-expressive category (M= 43.1, SD=18.35) received the lowest overall beauty ratings. Neutral clips received significantly lower beauty ratings than happy clips (M= 68.13, SD= 15.48, p< 0.001), sad clips (M= 63.11, SD= 15.88, p<0.001), angry clips (M= 64.84, SD= 16.27, p< 0.001), and fearful clips (M= 65.71, SD= 16.85, p< 0.001). After controlling for multiple comparisons, no other differences were found to be significant. These results indicate that participants found movement imbued with expressivity to be significantly more beautiful than movements absent of expression, regardless of the specific emotion being portrayed. A more detailed visualisation of these results can be found in Figure 3.1 below.

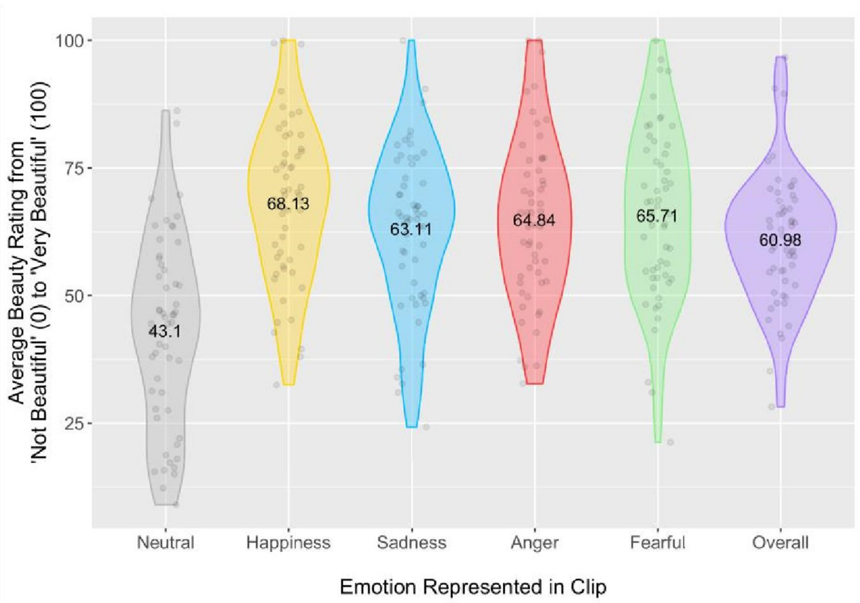


Figure 3.1: A visualisation of beauty scores provided by participants across all clips (overall), and across all clips within each specific emotion category.

Relationship Between Beauty Ratings and Overall Task Enjoyment

After completing the aesthetic judgement task, participants were asked to rate how much they enjoyed the task on a 100-point slider scale from 'Not at all' (0) to 'Very much' (100). The average task enjoyment rating provided for this experiment was 69.24 (SD = 21.87). A Spearman correlation was conducted to explore whether participants beauty ratings influenced their overall task enjoyment. A weak positive correlation was observed, but this relationship was not found to be statistically significant: $r_s = 0.214$, $p= 0.117$. Therefore, overall task enjoyment was not influenced by how beautiful participants found the movement stimuli to be.

Relationship Between Beauty and Empathy

Prior to completing the aesthetic judgement task, participants completed the Toronto Empathy Scale (TES; Spreng et al., 2009), and from their responses an individual empathy score was generated. The average empathy score across all participants was 48.56 (SD = 4.56).

A Pearson's correlation was conducted to explore whether individual differences in empathy had an impact on the overall beauty ratings provided. A significant positive relationship was observed between TES scores and average beauty rating provided: $r(53) = 0.407$, $p=0.002$. Therefore, in this experiment participants with higher self-reported levels of empathy generally found the movement sequences to be more beautiful than those with lower levels of trait empathy. A visualisation of this relationship can be observed in Supplementary Figure S.4.

Relationship Between Previous Dance Experience and Beauty Ratings

Participants also completed an early, unpublished version of the Gold-DSI (Rose et al., 2020) which provided detail about various aspects of their previous dance experience. One of these measures asked participants to classify their level of previous dance experience from one of the following five options: non-dancer, beginner, intermediate, advanced, or professional). While no participants classified themselves as a professional dancer, the majority reported having danced to some extent (15 beginners, 15 intermediate dancers, and 4 advanced dancers), and 21 participants reported having no previous dance experience.

A one-way ANOVA was conducted to explore whether different levels of prior dance experience had an impact on beauty ratings provided for these dance movement sequences, and a significant difference was observed: $F(3,51)= 3.607$, $p= 0.019$, $\eta^2= 0.175$. Further examination of this result showed that participants who classified themselves as "Intermediate" level dancers provided significantly higher beauty ratings than participants who classified themselves as "Beginners" ($p=0.047$), and marginally higher beauty ratings than participants who identified themselves as non-dancers (although this difference did not reach statistical significance after controlling for multiple comparisons; $p=0.066$). No other differences were found to be significant. Average beauty ratings provided by participants from each level of dance experience can be found in Table 3.2 and are visualised in Figure 3.2.

Table 3.2: Summary of average overall beauty ratings provided by participants across each level of previous self-reported dance experience.

	Non-Dancers (N= 21)	Beginners (N= 15)	Intermediate (N= 15)	Advanced (N= 4)
Average Beauty Rating Provided	57.25 (±13.19)	57.00 (±10.06)	68.04 (±11.02)	68.99 (±14.61)

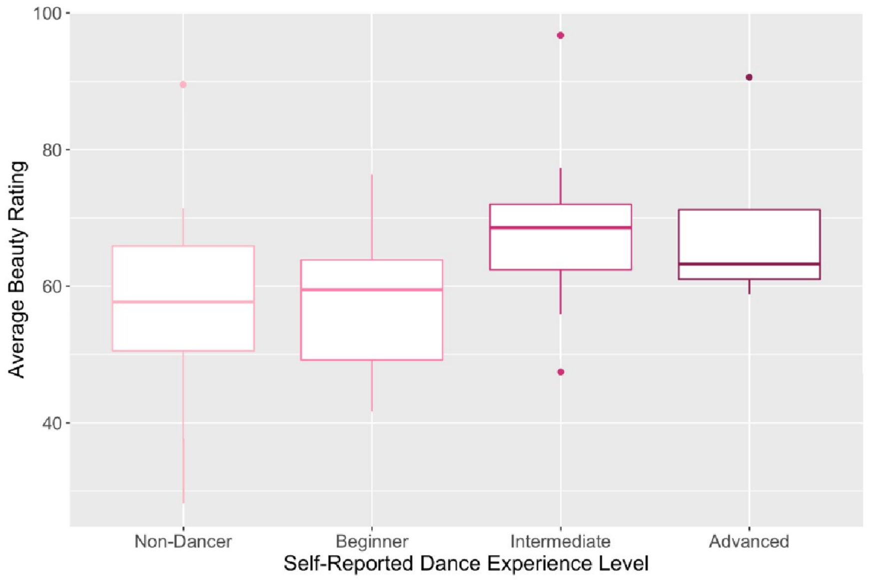


Figure 3.2: Average beauty scores provided by participants across each self-reported level of previous dance training.

The version of the Gold-DSI used in this experiment included a number of additional measures, three of which were related to physical dance experience ("How often do you currently go dancing for fun/social reasons?", "I have taken regular dance classes at least once a week for...", and "I have had formal training in any dance style for…"), and two measures which were related to observational dance experience ("How often do you watch dance performances/shows/videos on TV or internet?", and "How often do you attend live dance performances?"). These two sets of items were grouped together to create a more general score for each participants' physical dance experience and observational dance experience, respectively. One additional item from the questionnaire was not related to frequency of experience, and instead asked participants to select the number of dance styles they had experience in ('none', 'one dance style', 'two dance styles', 'three dance styles', or 'more than three dance styles') and was therefore considered separately in the analysis. As the number of possible responses to the items in this questionnaire varied from 5-7, prior to further analysis the responses were transformed using min-max normalization to account for the differences in scales.

To examine the impact of various facets of previous dance experience on beauty judgements made for the dance movement clips, a three-stage hierarchical multiple regression was conducted with average beauty score as the dependent variable. The physical dance experience factor was entered at the first stage of the regression, the observational dance experience factor was added in the second stage, and number of styles a participant was familiar with was added at the final stage to create the maximal model. The factors were added in this order to reflect the frequency with which these facets of dance experience have been explored in previous research in this field. The results can be found in Table 3.3 below.

Table 3.3: Results of a hierarchical multiple regression exploring the impact of physical dance experience, observational dance experience, and the number of trained dance styles on average beauty scores provided.

Model	Summary	Predictor	B	T	Std Error
Model 1	$F(1,53)= 4.948^{*}$, $R^2_{Adj}= 0.068$, RSE= 12.3	Intercept	56.789		
		Physical Experience	12.962	2.224^{*}	5.827
Model 2	$F(2,52)= 4.064^{*}$, $R^2_{Adj}= 0.102$, RSE= 12.08	Intercept	53.916		
		Physical Experience	7.048	$1.058^{n.s.}$	6.664
		Observational Experience	11.961	$1.730^{n.s.}$	6.914
Model 3	$F(3,51)= 3.288^{*}$, $R^2_{Adj}= 0.113$, RSE= 12	Intercept	54.370		
		Physical Experience	15.933	$1.660^{n.s.}$	9.596
		Observational Experience	12.886	$1.865^{n.s.}$	6.910
		Number of Styles	-8.215	$-1.279^{n.s.}$	6.420

Note: N= 55, $^{n.s.}p> 0.05$, $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$.

Model 1: Average Beauty Score ~ Physical Dance Experience Factor.

Model 2: Average Beauty Score ~ Physical Dance Experience Factor + Observational Dance Experience Factor.

Model 3: Average Beauty Score ~ Physical Dance Experience Factor + Observational Dance Experience Factor + Number of Dance Styles a Participant has Experience With.

The hierarchical multiple regression revealed that at Stage one, the regression model was a significantly better fit for the data than the null model $F(1,53)= 4.948$, $p= 0.030$ and that the physical dance experience factor explained 6.8% of the variation in average beauty score provided for the movement stimuli. Introducing observational dance experience as a predictor at stage two created a model which was a significantly better fit for the data than the null model $F(2,52)= 4.064$, $p= 0.023$, but did not fit the data significantly better than the stage one model $F= 2.99$, $p=0.09$ and only explained an additional 3.4% of the variance in beauty scores. Similarly, when introducing the number of dance styles a participant was familiar with as a factor in the maximal model, this was a better fit for the data than the null model $F(3,51)= 3.288$, $p= 0.028$, but did not significantly improve on the stage two model $F= 1.637$, $p= 0.207$ and only explained an additional 1.1% of the variation in beauty scores. In accordance with the Occam's Razor principle to avoid overfitting, the stage one model (including only the physical dance experience factor) is the model which best explains the data. Therefore, more extensive physical dance experience was found to significantly increase the average beauty rating provided by participants in this experiment, and this provides preliminary support for the notion that having previous dancer experience leads to increased aesthetic appreciation of expressive dance movement stimuli, even when there is limited visual information contained in these representations of the dynamic human body.

A 4*5 Factorial ANOVA was conducted to explore whether previous dance experience influences beauty ratings for movement stimuli containing different kinds of emotional expressivity. A significant main effect of self-reported dance experience level, $F(3,1355)= 52.745$, $p< 0.001$, and a significant main effect of emotion category, $F(4,1355)= 116.982$, $p< 0.001$, were observed. However, the interaction between these factors was not found to be significant; $F(12,1355)= 1.651$, $p= 0.072$. A visualisation of this interaction can be observed in Supplementary Figure S.5.

Liking Results

Participants were asked to provide liking ratings for each of the clips on a 100-point slider scale from 'not at all' (0) to 'very much' (100). The average liking rating provided by participants (across all 20 clips) in this study was 62.26 (SD= 12.06).

Comparison of Beauty Ratings Across Different Emotion Categories

A one-way ANOVA revealed significant differences in liking ratings across clips from different emotion categories: $F(4, 270)= 20.09$, $p< 0.001$, $\eta^2= 0.229$. This difference was explored in more detail through post-hoc Tukey testing, where it was found that clips from the neutral/non-expressive category (M= 45.69, SD=17.79) received the lowest liking ratings overall. Neutral clips were liked significantly less than happy clips (M= 71.57, SD= 15.65, p< 0.001), sad clips (M= 59.26, SD= 17.37, p< 0.001), angry clips (M= 69.79, SD= 17.72, p< 0.001), and fearful clips (M= 64.98, SD= 17.58, p< 0.001). In addition, sad clips were liked significantly less than angry clips (p= 0.013), and significantly less than happy clips (p= 0.002). No other differences were found to be significant. These results indicate that overall participants liked movement imbued with expressivity significantly more than movements devoid of expression. Unlike with beauty, however, liking ratings were influenced by the specific emotion being portrayed through the movements. Sad movements were liked significantly less than movements imbued with the traditionally high intensity emotions of anger and happiness. This may indicate that liking, but not beauty, judgements may be influenced by general expressivity and by the intensity of expression. A more detailed visualisation of these results can be found in Figure 3.3 below.

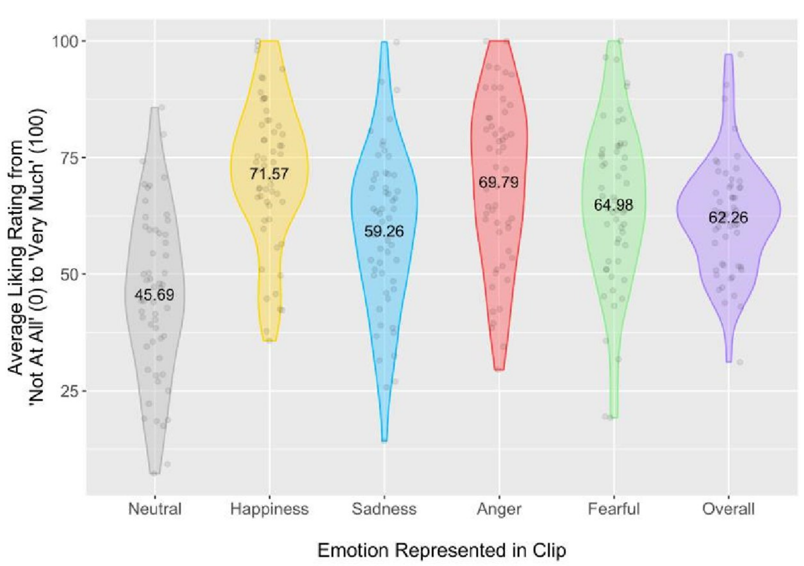


Figure 3.3: Liking scores provided by participants across all clips (overall), and across all clips within each specific emotion category.

Relationship Between Liking Ratings and Overall Task Enjoyment

A Spearman correlation was conducted to explore whether liking of the movement stimuli impacted overall enjoyment of the judgement task. A stronger positive correlation was observed for liking ratings than for beauty ratings, but this relationship was not found to be statistically significant: $r_s = 0.249$, p=

0.067. Therefore, the data does not indicate that overall task enjoyment was influenced to a significant degree by how much participants liked the movement stimuli.

Relationship Between Beauty and Empathy

A Pearson's correlation was conducted to explore whether individual differences in empathy had an impact on how much participants liked the movement stimuli. A significant positive relationship was observed between TES scores and average liking rating provided: $r(53)= 0.307$, $p= 0.022$. Therefore, in this experiment participants with higher self-reported levels of empathy generally liked the movement sequences more than those with lower levels of trait empathy. A visualisation of this relationship can be observed in Supplementary Figure S.6.

Relationship Between Previous Dance Experience and Beauty Ratings

A one-way ANOVA was conducted to explore whether level of previous dance experience had an impact on liking ratings provided for these dance movement sequences, and a significant difference was observed: $F(3,51)= 3.764$, $p= 0.016$, $\eta^2= 0.181$. As with the beauty ratings, participants who classified themselves as "Intermediate" level dancers provided significantly higher liking ratings than participants who classified themselves as "Beginners" ($p=0.046$), and marginally higher liking ratings than participants who had no previous dance experience; although this difference was not statistically significant after controlling for multiple comparisons ($p=0.053$). No other differences were found to be significant. Average liking ratings provided by participants from each level of dance experience can be found in Table 3.4 and visualised in Figure 3.4.

Table 3.4: Summary of average overall liking ratings provided by participants with each level of previous self-reported dance experience.

	Non-Dancers (N= 21)	Beginners (N= 15)	Intermediate (N= 15)	Advanced (N= 4)
Average Liking Rating Provided	58.81 (±11.24)	58.23 (±10.33)	69.02 (±11.13)	70.09 (±15.05)

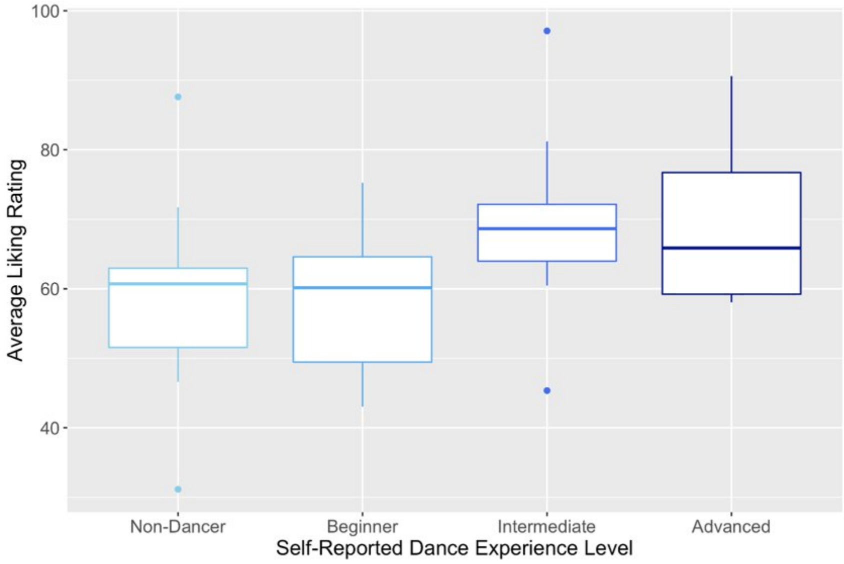


Figure 3.4: Average liking scores provided by participants across each self-reported level of previous dance training.

A three-stage hierarchical multiple regression was conducted to explore the factors obtained from the Gold-DSI as a function of how much participants liked the movement stimuli. As before, the predictors (physical experience, observational experience, and number of styles known; respectively) were added to the model in phases, with average liking score as the dependent variable. The results can be found in Table 3.5 below.

Table 3.5: Results from a hierarchical multiple regression exploring the impact of physical dance experience, observational dance experience, and the number of trained dance styles on average liking scores provided.

Model	Summary	Predictor	B	T	Std Error
Model 1	$F(1,53) = 5.845^{*}$, $R^2_{Adj} = 0.082$, RSE= 11.56	Intercept	57.979		
		Physical Experience	13.232	2.418^{*}	5.473
Model 2	$F(2,52) = 3.433^{*}$, $R^2_{Adj} = 0.083$, RSE= 11.55	Intercept	56.376		
		Physical Experience	9.932	$1.558^{n.s.}$	6.375
			6.676	$1.009^{n.s.}$	6.614

		Observational Experience			
Model 3	$F(3,51)= 2.645^{n.s.}$, $R^2_{Adj}= 0.084$, RSE= 11.55	Intercept	56.728		
		Physical Experience	16.815	$1.822^{n.s.}$	9.231
		Observational Experience	7.392	$1.112^{n.s.}$	6.646
		Number of Styles	-6.364	$-1.030^{n.s.}$	6.176

Note: N= 55, $^{n.s.}p> 0.05$, $^{*}p<0.05$, $^{**}p<0.01$, $^{***}p<0.001$.

Model 1: Average Liking Score ~ Physical Dance Experience Factor.

Model 2: Average Liking Score ~ Physical Dance Experience Factor + Observational Dance Experience Factor.

Model 3: Average Liking Score ~ Physical Dance Experience Factor + Observational Dance Experience Factor + Number of Dance Styles a Participant has Experience With.

The hierarchical multiple regression revealed that at Stage one, the regression model was a significantly better fit for the data than the null model F(1,53)= 5.845, p= 0.019 and that the physical dance experience factor explained 8.23% of the variation in average liking score provided for the movement stimuli. Introducing observational dance experience as a predictor at stage two created a model which was a significantly better fit for the data than the null model F(2,52)= 3.433, p= 0.040, but did not fit the data significantly better than the stage one model F= 1.019, p=0.317 and only explained an additional 0.1% of the variance in liking scores. Finally, when introducing the number of dance styles a participant was familiar with as a factor in the maximal model, this produced a model which was not a significantly better fit for the data than the null model F(3,51)= 2.645, p= 0.059. Adhering to the Occam's Razor principle to avoid overfitting, the stage one model (including only the physical dance experience factor) is the model which best explains the liking data. Therefore, more extensive physical dance experience was found to significantly increase the extent to which participants liked the movement clips in this experiment, and this provides further support for the idea that having previous dance experience leads to increased aesthetic appreciation of expressive dance movement stimuli, even when there is limited visual information contained in these representations of the dynamic human body.

A 4*5 Factorial ANOVA was conducted to explore whether previous dance experience influences liking ratings for movement stimuli containing different kinds of emotional expressivity. A significant main effect of self-reported dance experience level, $F(3,1355)= 44.582$, $p< 0.001$, and a significant main effect of emotion category, $F(4,1355)= 112.134$, $p< 0.001$, were observed. However, the interaction between these factors was not significant; $F(12,1355)= 1.522$, $p= 0.11$. A visualisation of this interaction can be found in Supplementary Figure S.7.

The Relationship Between Beauty and Liking

Finally, a Pearson's correlation was conducted to explore the relationship between average beauty and average liking ratings for stimuli presented to participants in this experiment. A strong positive correlation was observed between beauty and liking $r(53)= 0.866$, $p< 0.001$. This relationship can be observed in Figure 3.5. These results do show that participants provided higher beauty ratings for the clips that they liked more, but these factors were not perfectly correlated. Therefore, this result provides evidence to support the need for exploring both beauty and liking as a function of aesthetic appreciation, as these factors are related but do not measure the same thing.

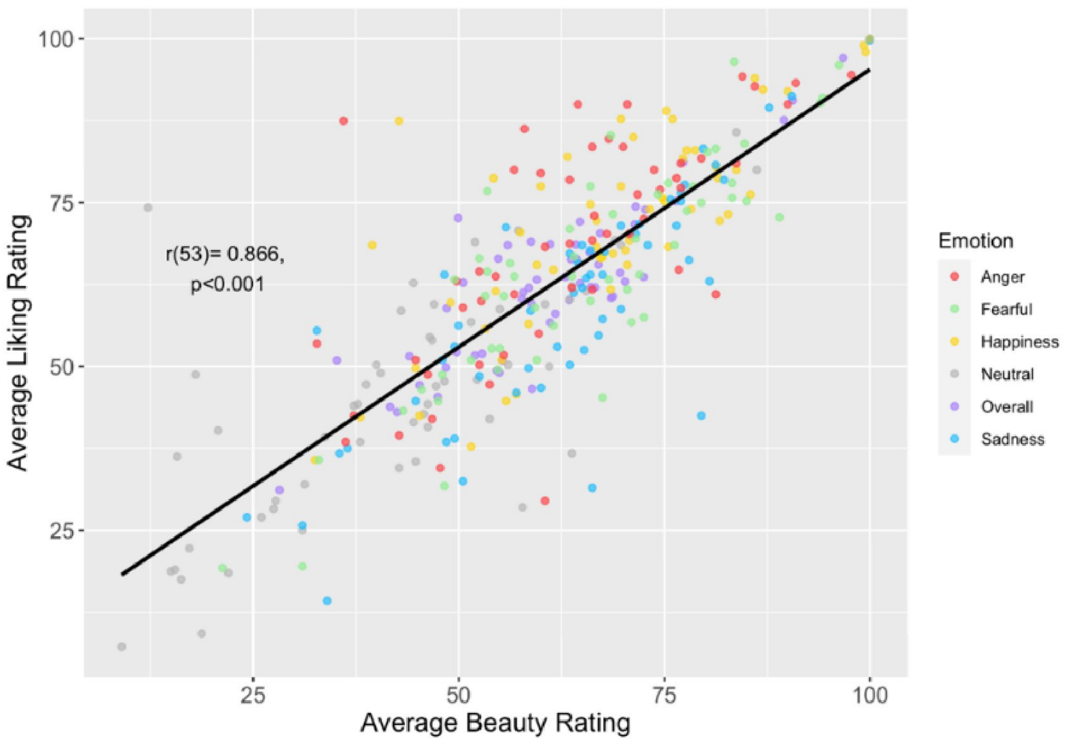


Figure 3.5: Relationship between average beauty and liking ratings provided by participants in this experiment.

DISCUSSION

In this study, we investigated whether empathy and previous dance experience influenced beauty and liking judgements beauty and liking for emotionally expressive whole-body dance movement clips (from the McNorm Library; Smith & Cross, 2022). As predicted, both dance experience and empathy had a significant impact on aesthetic evaluations, and differences in beauty and liking ratings emerged across clips communicating the expression of different emotional states. A more in-depth discussion of these results is presented in the following section.

Empathy Results

A number of studies exploring the relationship between empathy and aesthetic evaluation have provided evidence for an intrinsic link between these sociocognitive processes across multiple art mediums. For example, results from a series of behavioural and neuroimaging experiments have observed that participants who score higher on self-reported, and non-self-reported, measures of empathy provided more positive appraisals of music (Vuoskoski et al., 2011; Kreutz et al, 2008; Garrido & Schubert, 2011) and visual artworks (Gernot et al., 2018; Wilkinson et al., 2021; Ishizu & Zeki, 2017). In the current experiment, we found that participants with higher self-reported trait empathy (as measured by scores on the Toronto Empathy Scale; Spreng et al., 2009) provided higher overall beauty and liking scores for dance movement clips from the McNorm Library than individuals with lower self-reported empathy. These results add to the growing body of literature which suggests an intimate link between aesthetic processing and general empathy.

While evidence suggests the existence of medium-specific empathic skillsets, no measures of kinaesthetic empathy were collected in this experiment. It should be noted that although the concept of 'kinaesthetic empathy' is used frequently in dance movement therapy research (Rova, 2017, Castro et al., 2021), it is rarely discussed in those terms in dance aesthetics research. 'Kinaesthetic empathy' is instead typically explored through embodiment or proprioceptive capability, and activity of the AON (as measured by fMRI) is often used as a proxy for this in neuroimaging studies. To our knowledge, there is no self-report tool which provides an overview of an observer's kinaesthetic empathy-related skills. Neuroimaging research is expensive, and as a result of financial and time-related constraints, sample sizes are relatively small. Therefore, the development of such a tool would be useful for empirical aesthetics researchers who wish to explore whether this medium-specific kind of empathy exists within the wider population, and whether it plays a role in the aesthetic processing of movement stimuli.

We should note here that while empathy appears to play a significant role in aesthetic evaluation of human dance movement, it is not the sole component of an individual's personality which can influence the value we attach to dance movements. Indeed, results from studies exploring the visual arts and music suggest that Openness to Experience (one element of the OCEAN Five-Factor personality profile;

Pilarska, 2018) plays a significant role in sculpting our aesthetic experiences. It has been found that individuals with high scores along this dimension of personality reported experiencing greater "awe" in response to music and images of the natural world (Silvia et al., 2015). The impact of conscientiousness, agreeableness, and neuroticism had significantly smaller contributions to how much "awe" participants experienced than openness, and the effects of extraversion were found to be close to zero (Silvia et al., 2015). The authors therefore argue that openness to experience is essentially an aesthetic trait which comprises a substantial part of our personality. There is a wealth of additional literature which provides support for the idea that openness to experience plays a key role in determining how we view the aesthetic experience of music and the visual arts (Colver & El-Alayli, 2016; Rawlings et al., 2000; McCrae, 2007; Feist & Brady, 2004). However, the influence of this personality trait on beauty and liking judgements for dance movements is less well-understood. Future research should include some additional measures to examine the role of other personality factors, and in particular for the dimension of openness to experience, in the aesthetic appreciation of whole-body dance movements.

Dance Experience Results

Research suggests that previous dance experience can lead to changes in the perception of human movement stimuli (e.g., Calvo-Merino et al., 2005; Cross et al., 2006; Orgs et al., 2008). Studies using expert dancers have repeatedly shown that professional dance training can influence the perception of movement in a number of ways, and that even short-term dance training sessions can result in differences in the aesthetic evaluation of whole body dance movements (Kirsch et al., 2015; 2018). This has been observed with greater frequency among individuals with motor experience, but evidence also suggests that observational dance experience alone can lead to these altered perceptions (Kirsch et al., 2015; Gray et al., 1991).

Our data yielded partial support for the second hypothesis, in that individuals who classified their level of dance experience as 'intermediate' provided significantly higher overall beauty and liking ratings than beginner dancers, and marginally higher ratings than non-dancers; although these differences were no longer significant after controlling for multiple comparisons. No such differences were observed between participants with higher levels of previous experience (i.e., participants who classified themselves as advanced dancers). However, it should be noted that only four participants reported having more than an intermediate level of training, and no participants identified as professional dancers. It is possible that these differences would have emerged if a greater proportion of the participants had more substantial dance training.

In addition, we predicted that both physical dance experience – and to a lesser extent observational experience – would result in more favourable aesthetic value judgements for the dance movement stimuli. However, the results from this experiment provide only partial support for our experimental hypothesis. Hierarchical multiple regression analyses revealed that only physical dance experience had a significant impact on beauty and liking ratings. This result is surprising, given observations from the previous literature, as there are multiple studies which indicate that even observational dance experience can sculpt the way we perceive and respond to human movement. For example, in a study conducted by Jola and colleagues (2012), participants with long-term experience of watching ballet or Indian dance, and individuals with no observational dance experience ("novices"), watched three live performances of ballet, Indian dance, and a theatre play while motor-evoked potential (MEP) data was collected from the hands and arms. Movement of the arms and hands play a style-specific role in dance choreography; ballet dancing places greater emphasis on arm positioning, while Indian dance tends to involve more intricate movement of the hands and fingers. They found that participants showed heightened MEPs in response to dance movements from the style they were familiar with, in the associated body areas. More specifically, frequent observers of ballet showed higher MEPs in the arms in response to ballet movements than observers of Indian dance, or individuals with no visual experience of either style. The expected trend was also observed for frequent observers of Indian dance, as they demonstrated higher MEPs in the hands in response to the Indian dance performance than ballet spectators or novices (Jola et al., 2012). This study shows that even without motor experience in dance, physiological responses to dance are mediated by observational experience.

Few studies have specifically explored the impact of observational, rather than physical, experience on liking and beauty ratings for emotionally expressive whole-body dance movements. However, those which have tend to examine the impact of a short-term, intensive observational dance training, rather than long-term lifetime exposure to dance (Kirsch et al, 2015; 2018; Cross et al., 2009). It is possible that experience as a spectator of dance does prompt more positive aesthetic responses to movement, but that this only occurs in response to specific movement sequences which have been observed multiple times in training sessions, rather than to dance motion more generally. Therefore, going forward, it may also be useful to explore the impact of longer-term exposure to dance in real world settings on aesthetic value judgements for whole-body human movement.

Emotion Results

In support of Christensen and colleagues' (2021) findings, we also report here that observers showed a preference (in both beauty and liking ratings) for emotionally expressive dance movements, over the neutral (non-expressive) dance movement sequences. These results provide support for our final experimental hypothesis. While the potentially mediating impact of emotional expressivity on aesthetic value judgements has received relatively little research attention (in the evaluation of dance movements), this result makes sense within the context of the wider empirical aesthetics literature. For example, it has been found that observing emotionally-salient paintings result in more pronounced activation within the medial orbito-frontal cortex; an area believed to be responsible for processing beauty (Ishizu & Zeki, 2017). These results imply that affect, much like empathy, is intrinsically linked to the aesthetic experience.

This result has implications for dancers and choreographers, as it would appear from these results that movements which are perceived to contain little or no emotional expressivity may result in audiences disengaging with a performance. As this is not a desirable outcome for any dance practitioner, dance teachers and choreographers should place a specific importance on the ability to communicate emotions effectively through movement, in addition to developing technical accuracy, and dancers should ensure that when they are performing that they are using their body in ways that most effectively convey these emotional states. This poses an interesting question for future research in movement aesthetics. While some research has explored the impact of movement kinematics on emotional communication, this literature is rife with contradictory accounts of which aspects of motion are most salient for communicating basic emotional states (Atkinson et al., 2004; Crane & Gross, 2013; Gross et al., 2012; Melzer et al., 2019; Alaerts et al., 2011). In addition, it should be noted that these experiments frequently explore the success of emotion communication through measures of recognition accuracy (often as a percentage from forced-choice tasks), and it is likely that this approach only accounts for what is implicitly important for these judgements. To our knowledge, no experiments have explored what observers deem, explicitly, to be important for emotional communication through dance motion. It could be useful for future research to explore which aspects of motion contribute most substantially to emotional communication through a qualitative lens. Experiments like this may allow us to gain a deeper understanding of what is important for the communication of different emotional states from a perceptual standpoint, and this may help to generate a framework for dancers and choreographers to implement in their practice, with the hope of prompting greater audience engagement with performative dance.

Beyond differences between the presence or absence of emotional expressivity in movement, no specific predictions were made about whether the expression of specific emotions would influence aesthetic value judgements. However, in considering anecdotal evidence and research from other related fields, we proposed that observer preferences for emotional movement would emerge in one of two directions. First, in reversing the logic of "what is beautiful is good", we believed participants may have preferred movements communicating emotions with positive valence, and that the happy sequences would have received higher beauty and liking ratings than clips from the other emotion categories. Alternatively, given the pleasure certain individuals derive from sad music and the widespread popularity of performative ballets depicting tragedies, we also believed participants may demonstrate a preference for dance movements depicting sadness.

No significant differences were observed in beauty judgements across clips from the different emotion categories. However, for liking ratings, while non-expressive dance movements were still liked the least, participants demonstrated a significant preference for both happy and angry movements over sad movements. This result is surprising, as these preferences did not align with a trajectory in either of the speculated directions. Sad movements received the lowest liking ratings out of all of the emotionally expressive categories, therefore these results do not support an enjoyment of sadness in art that is universally applicable across different media. Additionally, we did not find complete support for the notion that "what is good is beautiful", as although happy movements were liked significantly more than sad movements, they were not liked more than angry movements. There are several potential explanations for these results which merit future research attention.

First, the argument has been made by dance practitioners that reducing performative dance to short clips without music, costume, stage furniture and all of the other elements involved in live performances means aesthetics researchers using such stimuli are no longer studying the artform of dance, but instead merely a form of biological motion. This may be a particular issue for experiments, like this one, which depict dance movement in the form of point-light displays. This specific presentation style is specifically designed to limit the interference of socially relevant cues (other than the movements themselves) from interfering with social judgements, but perhaps this could explain why sad movements from the McNorm Library were not found as aesthetically engaging as the sad movement sequences observed in ballets like *Swan Lake*, or *Giselle*. Perhaps it is some of the 'magic' of these different performative elements coming together, which is lost in translating movements to laboratory-based experimental paradigms, that makes observing the grief of star-crossed lovers such an aesthetically impactful experience. Future work should therefore explore how different elements of live performance contribute to overall aesthetic appraisal of human dance movement.

Additionally, it should be noted that in the traditional set of basic emotions, the only emotion with a positive valence is happiness (Jack et al., 2012). Sadness, anger, fear, and disgust all exist within the negative valence space, and surprise can be considered neutral on this continuum (as its expression can emerge in response to both positive and negative events). In this experiment, we focused on aesthetic appraisal of the most frequently studied emotions in this domain (i.e., happiness, sadness, anger, and fear). However, it is possible that "what is good is beautiful" may apply when a broader range of both negative and, in particular, positive valence emotions are included within the experimental stimuli. Alternatively, while anger and happiness exist on opposing ends along the dimension of valence, specific kinds of happiness (e.g., joy, elation) and anger (e.g., rage, fury) are both considered to be high arousal emotions. Therefore, these results could suggest that arousal plays a more important role in aesthetic preferences than valence, in terms of emotional expressivity in movement. Until future research explores aesthetic perceptions of movements communicating a broader range of emotions, differing in both valence and arousal, it would be premature to make any specific conclusions about whether the "good" is considered more beautiful.

CONCLUSIONS

The applications of empirical aesthetics research can be observed across many facets of our everyday lives, and engaging with aesthetics through the arts is important for human wellbeing and personal development. Considered together, these factors highlight the importance of research in the fields of empirical aesthetics in general. In the present study, we examined several factors which can impact the aesthetic experience of observing emotionally expressive whole-body dance movement, to add to the growing body of literature. We obtained evidence which demonstrates that, regardless of previous experience or individual differences in personality, observers show a general aesthetic preference for movement which is emotionally expressive. Additionally, while the specifics of emotional expressivity did not influence how beautiful observers found the dance movements to be, differences emerged in liking ratings as observers demonstrated a preference for movements which communicated emotions which, although differing in valence, are both higher in arousal (i.e., happiness and anger). Finally, higher levels of trait empathy and previous physical dance experience were found to mediate aesthetic perception of human movement, and prompt more positive appraisals in general.

CHAPTER 4

READ BETWEEN THE LINES: BEHAVIOURAL AND COMPUTATIONAL EXPLORATION OF FACTORS INVOLVED IN THE EMOTIONAL AND AESTHETIC EVALUATION OF MOTION-GENERATED CONTINUOUS LINE DRAWINGS

Read Between the Lines: Behavioural and Computational Exploration of Factors Involved in the Emotional and Aesthetic Evaluation of Motion-Generated Continuous Line Drawings

Rebecca, A. Smith., Andres, F. Rodriguez & Emily, S. Cross

ABSTRACT

Social perception research is an invaluable tool for identifying the factors which can mediate the success of our interactions with others. Understanding this is important for our personal development, wellbeing, and sense of identity. While a wealth of research has explored how we communicate social information to others through facial expressions and vocal cues, less is known about how the dynamic human body contributes to social information communication. Until we submit human body movement to the same rigorous empirical evaluation as other salient cues, our appreciation of social signalling as a whole will remain incomplete. Existing methods for studying human movement are fairly limited, and with this study, we aimed to develop and validate a new method for exploring human motion in social perception research. To this end, we created a series of continuous line drawings depicting movement trajectories of the wrists and ankles of a professional dancer during the performance of emotionally expressive whole-body movements. In an emotion recognition experiment we then examined whether these drawings could function as a viable proxy for human motion, but found that recognition of a series of target emotions (neutral, happy, sad, angry, and fearful) was significantly impoverished compared to existing types of movement stimuli and not significantly greater than chance. In a follow up experiment, we explored the possibility that these drawings could function as visual artworks in their own right, in isolation from their movement origins. Previous research has suggested that size, curvature, and complexity play specific roles in emotion labels and aesthetic ratings people assign to various kinds of visual art (e.g., paintings, simple polygons). When considering these trajectory drawings as visual artworks, we obtained evidence in the expected directions for both curvature and complexity. We conclude that while this stimulus creation method was not viable for its intended purpose (as an alternative stimulus for exploring the communication of emotion via human movement), the methodology presented in this manuscript nonetheless provides a new way to generate visual artworks which are suitable for use in visual empirical aesthetics research.

EXPERIMENT 1: INTRODUCTION

Developing and nurturing meaningful relationships with others plays a vital role in our health and wellbeing, as well as our sense of identity and belonging (Richardson et al., 2017; Killgore et al., 2020; Kaniušonytė et al., 2019; Haslam et al., 2022; Lim et al., 2021). A key component of successful interaction with others is our ability to read and interpret a number of verbal and non-verbal social cues, and to respond to these cues in an appropriate manner. This highlights the importance of social perception research, as research in this field allows us to develop an understanding of the factors which can mediate the success of these interactions. As a result of extensive research in this area we have made significant progress in understanding how we can effectively communicate our internal states to others through facial expressions (Jack et al., 2012; Frith, 2009; Smith et al., 2005; Chen & Jack, 2017) and vocal cues (both verbal and non-verbal; McAleer et al., 2014; Tonks et al., 2007; Rimé et al., 2002). However, much less is known about how we communicate this kind of information through the dynamic human body more broadly.

While research exploring the role of human body movement in social communication is relatively limited in comparison to other kinds of social cues, studies have identified several higher order social cues that we can interpret from observing another's body in motion. It has been found that people can rapidly attribute personality traits (e.g., dominance, trustworthiness; Levy & Duke, 2003; Heberlein et al., 2004) and affective expressivity (e.g., happiness, sadness, fear, pride; Shafir et al., 2016; Gross et al., 2010) to human movement stimuli. In such studies, human body movement has been depicted in many ways. These different kinds of movement stimuli vary widely in terms of ecological validity, in the extent to which researchers can control for confounding factors, and in the level of perceptual detail they provide to observers. In the following sections, we will provide a brief overview of existing methods, highlighting the trade-offs each existing approach offers.

Live Movement Performance

Relatively few experiments have explored responses to live movement, but the majority of the published work has centred around judgements of, and responses to, dance movements. Jola and colleagues, for example, aimed to explore the role of observational dance experience on corticospinal excitability (as a proxy for embodiment) in response to live ballet and Bharatanatyam dance. In this experiment, participants who were either experienced spectators of ballet or Bharatanatyam, or had no experience with observing dance performances of any kind observed three live solo dance performances (ballet, Bharatanatyam, and a non-dance theatrical piece) while transcranial magnetic stimulation was used to collect motor-evoked potentials (MEPs) in their hands and arms (Jola et al., 2012). They found that participants with extensive experience of watching dance produced greater MEPs in body regions associated with choreography in their familiar dance style (i.e., in the arms of ballet spectators, and in the fingers of Bharatanatyam spectators). While it may have been possible to explore this research question using recordings of dance movement, it has been argued that perceptual processing of human movement is fundamentally different when presented on a screen or when presented live (Shimada & Hiraki, 2006; Järveläinen et al., 2001). It is for this reason Jola and colleagues concluded that in their experiment, capturing genuine responses to movement (as we would encounter it in the real world) would only have been possible by using such real-time movement performance (Jola et al., 2012). Indeed, the idea that live performance is "special" is the central tenet of the 'NEUROLIVE' project led by Orgs and colleagues.

NEUROLIVE is the first large-scale project dedicated to exploring what makes 'liveness' special, by exploring the quantifiable, "behavioural, psychophysiological and neural entanglement between performers' and spectators' minds, brains and bodies" through collaboration between artists and scientists (NEUROLIVE, n.d.). As this project is the first of its kind to explore how the many facets of live movement performance coalesce to influence perceptions of observers, results obtained from this project will add significantly to our understanding of the importance and function of live experiences in the perception of human movement.

In review of the existing work, it is clear that observer responses to the live performance of whole body movements in real-time are likely to be most authentic (i.e., representative of the way these judgements are made in "real world" situations). However, a limitation of using movement in this form should be noted here. The main issue with using real-time movement performance to explore observer reactions is that it can be difficult to identify how social judgements are made in response to movement, specifically; in isolation from other confounding elements involved in performance settings. For example, it has been found that auditory cues, such as music and exertion sounds like breathing and gasping, can influence aesthetic and affective judgements of dance movements (Sievers et al., 2013; Burger et al., 2013; Christensen et al., 2014; Jola et al., 2014). In addition, it has been found that faces can draw substantial attention in social perception tasks. In one experiment, observers were presented with virtual agents who expressed either congruent or incongruent facial and bodily emotions. Through a series of emotion recognition tasks, it was found that while participants were more successful in their judgements in response to the congruent face-body agents, participants judgements were based primarily on emotional expressivity of the face (Clavel et al., 2009). These factors are just two of a number of confounds which may impact aesthetic and emotional processing of live movement stimuli. Therefore, interference from other sensory cues is an important source of artefact to consider when evaluating the overall utility and specific usefulness of using movement presented in this form to explore questions in social perception research. Further, these complications arising from the presence of other social cues may explain why research using live movement stimuli is relatively uncommon.

Full-Light Displays

Recordings of movement performances, when used as stimuli in social perception research, are referred to as full-light displays (FLDs). While it has already been noted that observation of movement in a 2-dimensional space can produce perceptions which may be inherently different to those which would be observed in response to live movement, these kinds of stimuli confer their own advantages from an empirical standpoint. One such benefit of using movement recordings is that researchers can ensure all participants are exposed to the same stimuli. In a run of live performances there will be natural fluctuations in, for example, the performer's energy level, mood, and technical accuracy across each performance, and any one of these factors may influence perceptions of the movements. By using consistent recordings of movements, researchers can be more confident that if observers differ in their perceptions, that this is not due to subtle differences in the experimental stimuli, and is instead likely to be a result of their experimental manipulation.

An additional benefit of using videos of movement, rather than real-time movement, is that FLDs allow researchers to have greater experimental control over the confounding factors outlined in the previous section. For example, while it is difficult to control for the influence of facial expressions on social judgements when observers watch real-time movement sequences, in FLD videos it is relatively simple to remove this kind of social cue (i.e., by blurring or pixelating the relevant areas). The ability to edit videos in this way allows researchers to manipulate, or remove, a number of such cues (e.g., the presence and style, or absence, of a musical accompaniment; the appearance of costumes and/or backdrops), providing greater flexibility in the research questions which can be explored. Further, some researchers argue that the use of FLDs should be encouraged because in real-world social scenarios we obtain full morphological information about the other in motion, therefore studies which limit the availability of this information may suffer in terms of ecological validity (Mazzoni et al., 2020). To this end, a number of researchers have made use of FLDs as stimuli in social perception experiments (Ross et al., 2012; Atkinson et al., 2004; Clarke et al., 2005; Manera et al., 2011).

However, it should be noted that in the ecological validity-experimental control trade-off, FLDs contain several visual confounds which cannot be removed, and depending on the specific research question to be explored, these cues may render FLDs unsuitable for use in some experiments. For example, in eye-tracking experiments, it has been found that the attractiveness of a movement performer has a significant impact on visual attention. Therefore, if a researcher wishes to explore visual fixation patterns in observers for movement, exclusively, then researchers may require stimuli without such morphological information about the movement performer.

Point-Light Displays

Out of the various movement stimuli currently used in social perception research, point-light displays (PLDs) contain the least visual information, and therefore include the fewest possible confounding factors. PLDs were originally developed by Johansson when he observed that the kinematics of biological motion could be depicted by attaching small light sources to the major joints of a model's body (Johansson, 1975; Krüger et al., 2018). The output of these displays are dot configurations that, when animated, leave the viewer with the impression that they are watching an animate being in motion. These stimuli have been used extensively in a variety of social perception experiments. For example, they have been employed in studies exploring the perception of gender (see review by Pollick et al., 2005) and identity (Mitchell & Curry, 2016), in addition to higher order social constructs like personality traits, sexuality, and affective states (Heberlein et al., 2004; Johnson et al., 2007; Atkinson et al., 2004; Gunns et al., 2002; Schneider et al., 2014).

As outlined in the previous section, the benefits of using PLDs are clear for specific research questions that require a singular focus on 'pure' motion, or the specific exclusion of other visual features like faces. However, it should also be noted here that in research aiming to explore more naturalistic experiences of motion and identification of more complex social signals, these 'benefits' may instead be detrimental, and it may be more useful to select an alternative presentation method (e.g., FLDs, or live movement performance).

Overview of Stimuli Types

In summary, the use of PLDs, FLDs, or live movement stimuli all provide a source of unique benefits and limitations. Therefore, depending on the specific research question, any of the movement stimuli discussed here may be uniquely suitable for exploring observer perceptions of the dynamic human body. Despite this, the breadth of methods available to researchers interested in the role of the dynamic human body in social perception is fairly limited (compared to those developed for facial expression research, for example).

When considering the trade-off between experimental control and ecological validity, PLDs and live performance, respectively, represent extremes at each end of this spectrum; with FLDs falling somewhere in the middle. There does appear to be significant potential for, and interest in, novel exploration of the live movement experience (e.g., NEUROLIVE), which provides the potential for new methodologies and experimental paradigms to emerge. On the other end of the scale, however, PLDs have been used extensively in social perception research since the 1970s, and, to our knowledge, few (if any) novel ways have been proposed to explore responses to reduced-format versions of movement since then. The creation of new stimuli in this category is important because this would allow researchers to explore the level of information reduction required to produce floor effects of recognition, and to examine whether there are individual differences which can facilitate more subtle improvements in recognition capabilities.

Here we present a novel stimulus presentation method for exploring emotion recognition from human dance movement. These stimuli depict the movement trajectories of the body extremities (wrists and ankles), during emotionally expressive whole body dance movement sequences, in the form of continuous line drawings. This method of using human movement to create line drawing stimuli has been employed in recent empirical aesthetics research for the visual arts. In one recent study conducted by Chamberlain and colleagues (2022), across a series of four experiments participants with various levels of art expertise provided naturalness judgements and aesthetic value judgements for graffiti-style line drawings which were either computationally generated, or produced through human drawing actions. In the first three experiments, these drawings were presented to observers dynamically (i.e., participants observed videos of the drawing's creation in real-time), and it was found that participants could distinguish between human-made, and computer-generated drawings, regardless of their previous art experience. It was also found that the human-made drawings were perceived to be more aesthetically pleasing. The authors highlight that these judgements may have been influenced by kinematic profiles of the drawing process, as human-made drawings possess several unique kinematic characteristics (including uneven velocity and jerkiness profiles). In a fourth experiment, participants were shown only the static final drawings (i.e., the end-products of the drawing process), and here only the art-experienced participants could differentiate between drawings produced by humans versus those produced computationally (Chamberlain et al., 2022). Considered together, these results suggest that human movement can be perceived from simple line drawings, and that art-experienced participants are more adept at making these judgements. However, it is unclear whether inferences about human motion can be made exclusively from visual artworks created as a by-product of drawing-specific actions. We set forth the notion that, perhaps, such stimuli created from other kinds of human movement (like whole-body dance motion) could produce similar perceptions, particularly in individuals with extensive movement experience.

If the emotional content in the line-drawing stimuli derived from dance movements presented in this manuscript can be recognised successfully by observers, this approach would provide a promising new avenue to visualise human movement for use in social perception research. Based on previous work, it appears that as visual information present in stimuli decreases, so too do recognition rates of the intended emotion. For example, recognition of the intended emotion has been found to be higher in response to FLDs, than for PLDs. To this end, we predict that emotion recognition rates for these trajectory images will likely be impoverished compared to PLDs, but still have the potential to yield result in recognition rates greater than chance level.

To further situate our results among the previous literature, we will also explore the impact of two factors known to mediate emotion recognition from human movement; individual differences in empathy, and previous dance experience.

Empathy and Emotion Recognition

Many previous studies have observed a link between empathy and emotion recognition capabilities. This has been observed across tasks using social cues of various modalities, including faces, voices, and body postures (Balconi & Bortolotti, 2012; Besel & Yuille, 2010; Holland et al., 2021; Israelashvili et al., 2020; Jospe et al., 2018; Neumann et al., 2014). This link between trait empathy and emotion recognition has also been observed for body movements (Rizzolatti & Craighero, 2005). Considering these factors together, if the continuous line drawing stimuli could be used as a proxy for representing emotionally expressive human movement, we may expect individual differences in empathy to mediate success in an emotion recognition task using these stimuli.

Therefore, we predict that participants with higher levels of trait empathy will be more accurate in their emotion judgements than participants with lower levels of empathy.

Dance Experience and Emotion Recognition

Studies have also noted that dancers appear to recognise affective expression from human movement with greater accuracy than individuals with no prior dance experience. Christensen and colleagues observed that dancers were significantly more accurate in the recognition of positive affect from movement than non-dancers (Christensen et al., 2016). Similar results were observed in a later study, wherein participants from a variety of dance backgrounds were asked to observe a series of movement sequences and identify whether these movements were expressive or non-expressive (Christensen et al., 2019). Years of dance experience was found to correlate with accuracy in affective recognition, providing further support for the idea that dance experience facilitates the ability to recognise expression from emotion whole-body movement displays. Beyond physical experience, it has also been found that purely observational dance experience can sculpt perceptions of the dynamic human body. Therefore, again, we may expect dancers and dance observers to provide more accurate emotion judgements for these trajectory line drawings if they function as an alternative way to visualise motion.

Accordingly, we predict that participants with a higher degree of previous dance experience will obtain higher emotion recognition rates than those with less previous experience. In addition, we predict that physical dance experience will have a more significant impact on recognition rates, but that observational dance experience may also result in improved recognition rates.

<u>Aesthetic Evaluation</u>

For two main reasons, we will also explore aesthetic value judgements assigned to these kinds of movement stimuli. First, although most empirical aesthetics research focuses on music and the visual arts, dance (portrayed in a variety of forms) is quickly gaining traction as a subject in this area. Most studies exploring aesthetic responses to dance have depicted the movements as full-light or point-light display video clips, but it is possible that dance movements presented in the form of these continuous line drawings hold their own aesthetic value for observers. Second, there is a wealth of evidence to suggest that aesthetic and affective evaluation of socially-relevant stimuli involve processes which are cognitively similar, and may be intrinsically linked (incl. Ishizu & Zeki, 2017; see Chapter 1 of this thesis for a more thorough discussion of this relationship). Therefore, when examining the emotional expressivity assigned to such stimuli, it may also be important to consider aesthetic appraisals in this context.

As, to our knowledge, no one has explored aesthetic value judgements through trajectory drawings, and certainly not in comparison to other types of movement stimuli, we have no set predictions about how beauty and liking ratings will compare across these stimulus presentation methods. However, it has been found that trait empathy, and art-expertise can mediate the aesthetic value judgements assigned to art across various mediums.

Empathy, Dance Experience and Aesthetic Evaluation

Behavioural and neuroscientific evidence suggests that individual differences in empathy influence aesthetic processing of works of art. This relationship can be observed in a number of studies which have been conducted across a variety of artistic mediums, including visual artworks (e.g., paintings, photographs), music and the performing arts (Gernot et al., 2018; Wilkinson et al., 2021; Crozier & Greenhalgh, 1992; Arnold et al., 2014). Additionally it has been observed that art expertise can mediate the value people assign to aesthetic stimuli from their medium of expertise. This too has been observed across multiple artforms; mostly for paintings and music (Silvia & Berg, 2011; Brattico et al., 2013; Yeh et al., 2018). However, a number of seminal experiments have provided evidence for the role of both empathy and expertise in facilitating perception of human movement stimuli, specifically.

Calvo-Merino and colleagues, for example, were the first to obtain neuroscientific evidence for the relationship between kinaesthetic empathy and aesthetic processing of human dance movement. In this experiment, participants observed a series of whole-body dance sequences while neuroimaging data were collected. In a subsequent session, the same participants were asked to provide aesthetic evaluations (interest and liking) for each of the dance clips, and it was found that the clips which (as a consensus) were liked more prompted greater activity in the premotor cortex than clips which were not as well-liked. This implies an inherent link between kinaesthetic empathy (as measured by activation of the motor areas) and aesthetic evaluation (Calvo-Merino et al., 2008). Similar results have been observed when examining the impact of short-term motor experience. In one study, participants participated in a 4-day dance training session, where each day they physically rehearsed one set of movements, passively observed another, and listened to the music accompanying a third set; a final set of movements remained untrained to allow for comparison (Kirsch et al., 2015). fMRI data were collected before and immediately after the training sessions, along with self-reported measures of movement capability and aesthetic value judgements for all four categories of movement sequences. They found that participants liked the movements which they had both physically performed and observed significantly more than the music-only sequences and the untrained sequences (Kirsch et al., 2015).

In line with these findings, individual differences in empathy and prior dance experience likely mediate the perception of these trajectory images. Therefore, we predict that individuals who obtain higher scores on the Toronto Empathy Scale (Spreng et al., 2009) will provide more positive aesthetic value judgements (beauty and liking) for the continuous line drawings. We also predicted that previous dance experience would also mediate aesthetic value judgements. More specifically, that individuals with a higher degree of previous dance training would provide more positive aesthetic appraisals of the trajectory images, and that physical dance experience would facilitate this increase in beauty and liking ratings to a greater extent than observational dance experience.

EXPERIMENT 1: METHODS

OPEN SCIENCE STATEMENT

Consistent with open science practices widely adopted within psychological research (Open Science Collaboration, 2012), we report all manipulations and all measures in the study. In addition, following open science initiatives (Munafò, 2016), the data, stimuli, and analysis code associated with this study are freely available on the Open Science Framework (https://osf.io/dj2sh/). By making the data available, we enable others to pursue tests of alternative hypotheses, as well as more exploratory analyses.

PARTICIPANTS

Previous emotion recognition experiments have reported participant samples sizes that range between 12 and 80 participants (Grezes et al., 2007; Christensen et al., 2019), and studies exploring aesthetic perception of dance movement stimuli have reported participant samples which range between 6 and 41 (Calvo-Merino et al., 2008; Christensen et al., 2021). Since the average number of participants prior studies in this domain report testing is around 35 participants, we aimed to collect data from at least 35 participants in our experiment.

68 participants were recruited through the Prolific online testing platform. 10 participants were excluded from the analysis for several reasons (including failed attention checks, technical issues during the task, nonsensical demographics responses, failing to provide the completion code, and providing the same responses across all stimuli with improbably short completion times). Therefore, after participant exclusion, the final sample size was 58. Participants ranged in the amount of previous dance experience they had, but they were assigned to one of five different dance-level groups (non-dancer, beginner, intermediate, advanced, or professional) by indicating which label they thought was most applicable to their previous experiences. The experiment was created in formr (Arslan et al., 2020).

STIMULI

The movement trajectory coordinates used to create the stimuli used in this experiment were extracted from emotionally expressive whole-body dance movement sequences from The McNorm Library (Smith & Cross, 2022). To ensure the line drawing stimuli would be most representative of the different emotion categories (neutral, happy, sad, angry, and fearful), the coordinate data was extracted from the 4 movement sequences with the highest levels of recognition for each target emotion. Further detail about the emotion validation experiments, and the raw emotion recognition data can be found on the Open Science Framework (https://osf.io/dj2sh/).

To create these trajectory line drawings, 2D movement coordinate data was extracted for the wrist and ankle markers across the timeseries of the movement sequence from a frontal view of the dancer. In

RStudio, each 2D coordinate was then plotted as an individual point on a scatterplot. These points were then joined by a continuous black line, which connected the points in order following the timeseries. Grid and axes lines, and labels were then manually removed from the resulting plots, and the scatter points under the continuous line were made fully transparent. Therefore, the final scatterplots contained only the black continuous line on a white background. These plots were then saved as png files at a consistent size of 8 x 5 (inches). This process was then repeated for the selected movements of the left wrist, right wrist, left ankle, and right ankle. A visualisation of this process can be observed in Figure 4.1, and two examples of the final line drawing stimuli can be observed in Figure 4.2.

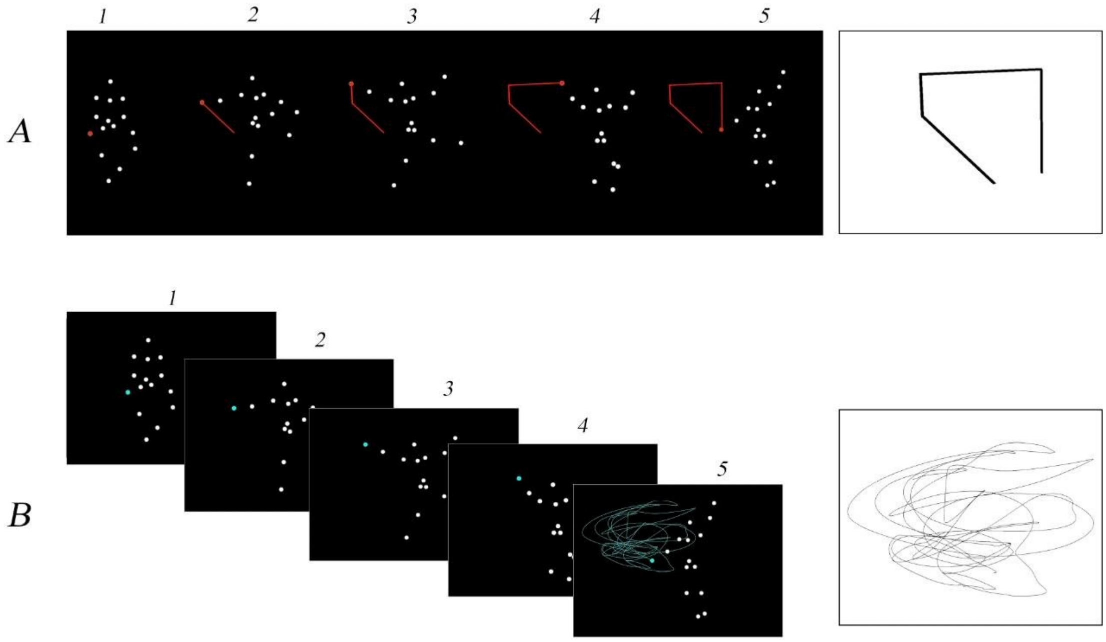


Figure 4.1: In this figure, A shows a simplified representation of the trajectory extraction process. The red lines depict the movement of the right wrist marker across 4 time points, and the rightmost image shows what that trajectory would look like as a continuous line drawing. B shows the trajectory of the right wrist (overlayed in blue) that emerged by repeating this process across all timepoints of the movement sequence, and the rightmost image shows what that isolated trajectory looked like as one of our experimental stimuli.

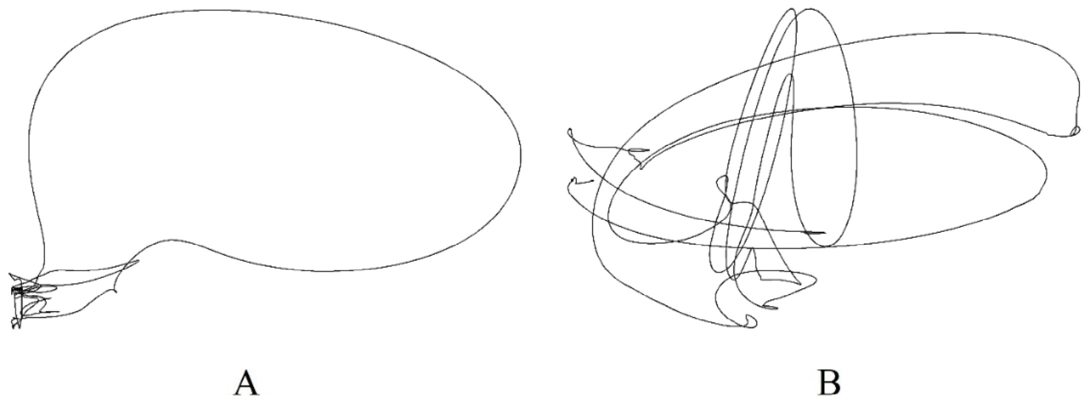


Figure 4.2: Drawing A is an example of a right ankle movement extracted from a fearful clip. Drawing B is an example of a right wrist movement extracted from a happy clip.

PROCEDURE

Participants followed the formr link to begin the experiment. After reading the information and consenting to participate (in accordance with BPS guidelines), participants completed two questionnaires: a series of questions extracted from an early, unpublished, version of the Goldsmiths Dance Sophistication Index (Gold-DSI; Rose et al., 2020), and the Toronto Empathy Scale (TES; Spreng et al., 2009).

Selected questions from the early Gold-DSI included demographics measures allowed participants to provide detail about their previous dance experience, covering both formal and informal experience, and both visual and physical experience. Participants were also asked to select which level of dance experience was most applicable to them from one of five options (non-dancer, beginner, intermediate, advanced, or professional. The TES was used to generate a trait empathy score for each participant (ranging from 0 – 64).

Participants then began the experiment wherein they observed pairs of trajectory images depicting the right and left versions of either the wrist or ankle markers presented in random order. After observing each image pair, they were asked “What emotion do you think this movement was intended to convey?” in a forced-choice paradigm (neutral, happy, sad, angry or fearful). Participants also provided aesthetic evaluations of the image pairs. They responded on a sliding scale to the questions “How beautiful did you find this image to be?” from “not beautiful” (0) to “very beautiful” (100), and “How much did you like this image?” from “not at all” (0) to “very much” (100). The whole experiment took around 30 minutes to complete, and participants received £3 for their participation.

EXPERIMENT 1: RESULTS

Emotion Recognition Results

Overall average recognition of the intended emotion from the trajectory images was 20.47% (+11.11). However, recognition differed to a certain extent across images from the different emotion categories. Average recognition rates for images from each of the intended emotion categories can be observed in Table 4.1 below.

Table 4.1: A summary of the average recognition rates (and standard deviations) for trajectory images from each emotion category.

	Neutral	Happy	Sad	Angry	Fearful
Average Recognition Rate (%)	26.17 (± 5.06)	29.30 (± 13.99)	12.50 (± 4.80)	21.09 (± 12.44)	13.28 (± 5.48)

The results of a 1 x 5 ANOVA revealed that average recognition rates were significantly different across the different emotion categories: $F(4,35)= 5.264$, $p=0.002$. A post-hoc Tukey HSD test revealed that images depicting the trajectories of happy movements were recognised with greater accuracy than trajectories of fearful movements ($p=0.012$) and trajectories of sad movements ($p=0.008$). In addition, images depicting the trajectories of neutral movements were recognised with greater accuracy than images depicting sad movements ($p=0.042$). No other differences in recognition rates across the different emotion categories were found to be significant. A visualisation of the average emotion recognition rates for images from each of the intended emotion categories can be observed in Figure 4.3.

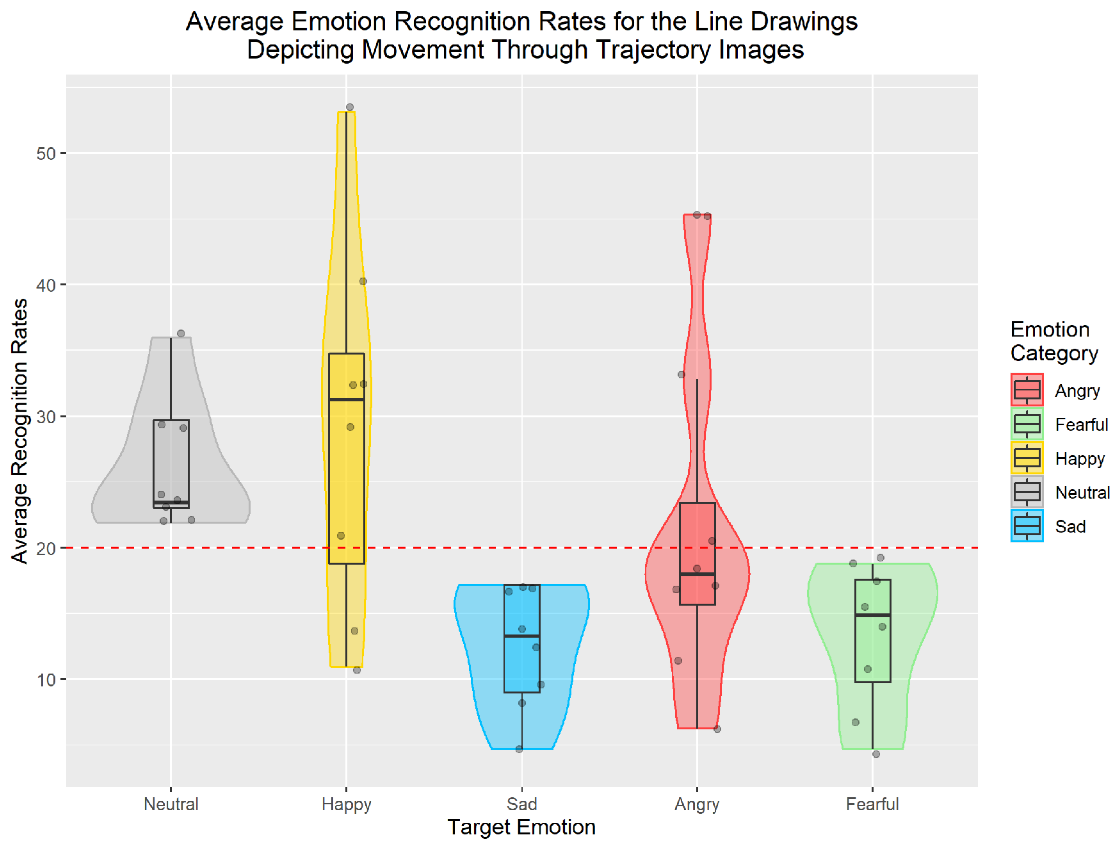


Figure 4.3: A visualisation of average recognition of the target emotion for trajectory line drawings from each of the emotion categories. The red dotted line represents the chance level of recognition (20%).

A series of one-sample t-tests were conducted to explore whether recognition of the target emotions occurred at greater than chance level. As the forced-choice recognition task included 5 emotion options, chance in this case was 20%. Images depicting the trajectory of neutral movements were recognised at greater than chance level; $t(7)= 3.448$, $p=0.005$. While the happy images were recognised with higher average accuracy than neutral images, results of the one-sample t-test revealed this recognition rate was not greater than chance: $t(7)= 1.879$, $p=0.051$. Additional t-tests exploring recognition of the intended emotion from movement trajectory images indicated than no other recognition rates were significantly greater than chance (all p values > 0.05).

As images were presented in pairs, with each pair depicting the trajectory of either the wrists or ankles of the movement performer, an additional t-test was conducted to explore whether recognition of the target emotion differed depending on movement of which part of the body the trajectory image displayed. While recognition rates were higher for the wrist images ($M= 22.66 \pm 13.79$) than ankle images ($M= 18.28 \pm 7.30$), this different was not found to be significant: $t(28.87)= -1.254$, $p= 0.2197$.

Therefore, these results do not provide support for the primary hypothesis, as only the neutral trajectory drawings extracted from movement of the wrist and ankle joints were assigned the intended emotion label at greater than chance level.

Emotion Recognition x Dance Experience Results

An ANOVA was conducted to explore whether different levels of previous dance experience influenced accuracy in the emotion recognition task. While participants at both extremes (i.e., non-dancers (M= 21.10 ± 6.77), and advanced dancers (M= 20.83 ± 6.29)) demonstrated marginally higher recognition rates than those with beginner (M= 19.35 ± 6.04) or intermediate (M= 19.29 ± 7.60) experience, these differences were not found to be significant: $F(3,54)= 0.34$, $p=0.797$.

A hierarchical multiple regression was conducted to explore whether physical dance experience, observational dance experience, and the number of dance styles known influenced recognition rates. A more detailed breakdown of the results of this analysis can be observed in Table 4.2.

Table 4.2: Results of the hierarchical multiple regression analyses exploring the role of dance experience in emotion recognition rates.

Model	Summary	Predictor	B	T	Std Error
Model 1	$F(1,56)= 0.125^{n.s.}$, $R^2_{Adj}= -0.016$, RSE= 6.503	Intercept	20.509		
		Physical Experience	-1.548	$-0.354^{n.s.}$	4.373
Model 2	$F(2,55)= 0.407^{n.s.}$, $R^2_{Adj}= -0.021$, RSE= 6.521	Intercept	21.364		
		Physical Experience	-0.516	$-0.113^{n.s.}$	4.558
		Observational Experience	-3.072	$-0.831^{n.s.}$	3.699
Model 3	$F(3,54)= 0.257^{n.s.}$, $R^2_{Adj}= -0.040$, RSE= 6.581	Intercept	21.365		
		Physical Experience	-0.539	$-0.095^{n.s.}$	5.660
		Observational Experience	-3.076	$-0.815^{n.s.}$	3.774

	Number of Styles	0.025	$0.007^{n.s.}$	3.545

N = 50; n.s. $p > 0.05$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Model 1: percentage recognition − physical dance experience factor

Model 2: percentage recognition − physical dance experience factor + observational dance experience factor

Model 3: percentage recognition − physical dance experience factor + observational dance experience factor + number of dance styles a participant has experience with

The hierarchical multiple regression revealed that at Stage one, the regression model was not a significantly better fit for the data than the null model (p> 0.05), and that physical dance experience only explained 1.6% of the variation in average recognition rate. Introducing both observational dance experience and number of styles known to the model at Stages two and three did not result in models which fit the data better than the null model and only explained an additional 0.5% and 1.9% of variance in the recognition rates, respectively. These results do not suggest that previous dance experience facilitated greater emotion recognition accuracy for the movement trajectory image stimuli used in this experiment, and therefore, we did not find evidence to support the second experimental hypothesis.

Emotion Recognition x Empathy Results

A Spearman's correlation was conducted to explore whether trait empathy (as measured by the Toronto Empathy Scale; Spreng et al., 2009) influenced recognition accuracy. While there appeared to be a very small positive relationship between empathy score and recognition accuracy, this relationship was not statistically significant: rho = 0.087, p=0.518. Therefore, in this experiment we did not find support for the prediction that individuals with higher trait empathy would provide more accurate emotion judgements.

Liking Results

The overall liking rating provided for the trajectory images was 45.26 (± 5.67). Liking ratings differed depending on which emotion was contained within the movement the trajectory image was depicting. Average liking ratings for images from each of the intended emotion categories can be observed in Table 4.3 below

Table 4.3: The average liking ratings provided by participants for trajectory images from each emotion category.

	Neutral	Happy	Sad	Angry	Fearful
Average Liking Rating	49.61 (± 6.69)	45.07 (± 6.82)	43.14 (± 4.97)	45.00 (± 3.60)	43.47 (± 4.40)

While clips depicting the trajectory of neutral movements seemed to be liked the most, and the sad and fearful movement trajectory images seemed to be liked the least, the results of a 1 x 5 ANOVA revealed that these differences were not significant: $F(4, 35)= 1.802$, $p= 0.15$. A visualisation of the difference in liking ratings across images from the different emotion categories can be observed in Figure 4.4.

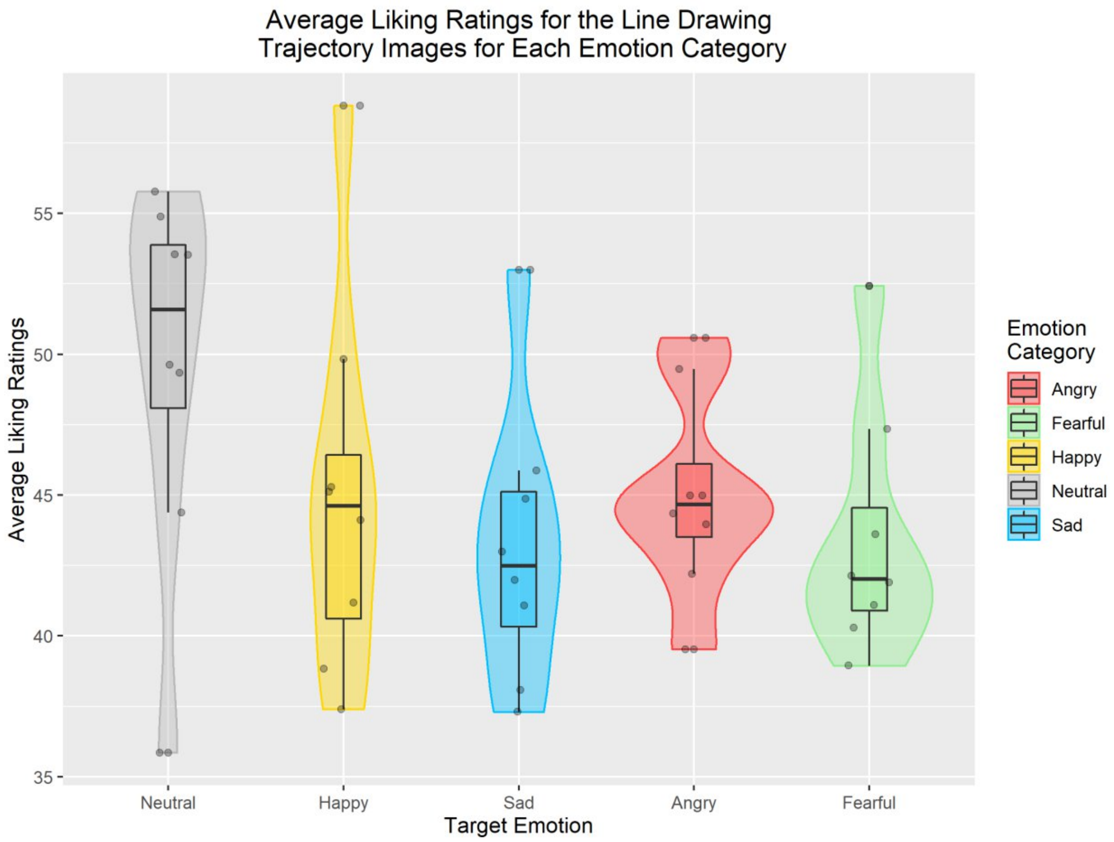


Figure 4.4: A visualisation of the average liking ratings provided by participants for trajectory images from each of the emotion categories.

Again, as the images pairs depicted the movement of either the wrists or the ankles, a further t-test was conducted to explore whether participants provided different liking ratings depending on which body part's movement they were observing. While trajectory images depicting wrist movement ($M= 46.88 \pm 5.91$) were liked marginally more than ankle movement images ($M= 43.64 \pm 5.05$), this difference was not found to be significant: $t(37.107)= -1.865$, $p= 0.070$.

Liking x Dance Experience Results

An ANOVA was conducted to explore whether different levels of previous dance experience influenced liking ratings provided for the movement trajectory images. While participants with an intermediate level of experience (M= 51.49 ± 23.81) provided higher overall liking ratings than advanced dancers (M= 48.60 ± 28.99), beginner dancers (M= 48.97 ± 22.90), and non-dancers (M= 48.88 ± 14.25), these differences were not found to be significant: $F(3,54)= 0.04$, $p= 0.991$.

A hierarchical multiple regression was conducted to explore whether physical dance experience, observational dance experience, and the number of dance styles known influenced overall liking ratings. A more detailed breakdown of the results of this analysis can be observed in Table 4.4.

Table 4.4: Results of the hierarchical multiple regression analyses exploring the role of dance experience in liking ratings provided for the trajectory drawings.

Model	Summary	Predictor	B	T	Std Error
Model 1	$F(1,56)= 0.26^{n.s.}$, $R^2_{Adj}= -0.013$, RSE= 19.56	Intercept	47.76		
		Physical Experience	6.70	$0.509^{n.s.}$	13.15
Model 2	$F(2,55)= 2.65^{n.s.}$, $R^2_{Adj}= -0.055$, RSE= 18.90	Intercept	41.08		
		Physical Experience	-1.37	-0.104	13.21
		Observational Experience	24.01	2.240^{*}	10.72
Model 3	$F(3,54)= 1.83^{n.s.}$, $R^2_{Adj}= 0.042$, RSE= 19.02	Intercept	41.24		
		Physical Experience	-6.36	-0.389	16.36
		Observational Experience	23.16	2.123^{*}	10.91
		Number of Styles	5.36	0.523	10.25

N = 50; n.s. $p > 0.05$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Model 1: liking rating − physical dance experience factor

Model 2: liking rating − physical dance experience factor + observational dance experience factor

Model 3: liking rating − physical dance experience factor + observational dance experience factor + number of dance styles a participant has experience with

None of the models explained the data significantly better than the null model (all $p > 0.05$). In the stage one model, physical dance experience only explained 1.3% of the variation in average recognition rate. Introducing observational dance experience to the model at stage two only explained an additional 4.2% of variance, and adding the number of styles known at stage three actually decreased the variance in liking ratings explained by the model to 4.2%. Therefore, these results do not provide support for our prediction that participants with previous dance experience would provide higher liking ratings for the movement trajectory images in this experiment.

Liking x Empathy Results

A Pearson's correlation was conducted to explore whether trait empathy (as measured by the Toronto Empathy Scale; Spreng et al., 2009) influenced the liking ratings provided by participants for the trajectory line drawings. While there appeared to be a very small negative relationship between empathy score and liking ratings, this relationship was not statistically significant: $r(56) = -0.181$, $p=0.174$. Therefore, in this experiment, we did not find support for the prediction that individuals with higher trait empathy would provide higher liking ratings.

Beauty Results

The overall beauty rating provided for the trajectory images was 44.53 (± 5.78). Beauty ratings differed depending on which emotion was contained within the movement the trajectory image was depicting. Average beauty ratings for images from each of the intended emotion categories can be observed in Table 4.5 below.

Table 4.5: The average beauty ratings provided by participants for trajectory images from each emotion category.

	Neutral	Happy	Sad	Angry	Fearful
Average Beauty Rating	49.04 (± 7.70)	44.68 (± 6.49)	41.76 (± 4.83)	44.34 (± 3.39)	42.86 (± 3.79)

While clips depicting the trajectory of neutral movements were deemed to be the most beautiful, and the sad movement trajectory images deemed to be the least beautiful, the results of a 1 x 5 ANOVA revealed that these differences were not significant: $F(4, 35)= 2.047$, $p= 0.109$. A visualisation of the differences in beauty ratings provided for images from each of the emotion categories can be observed in Figure 4.5.

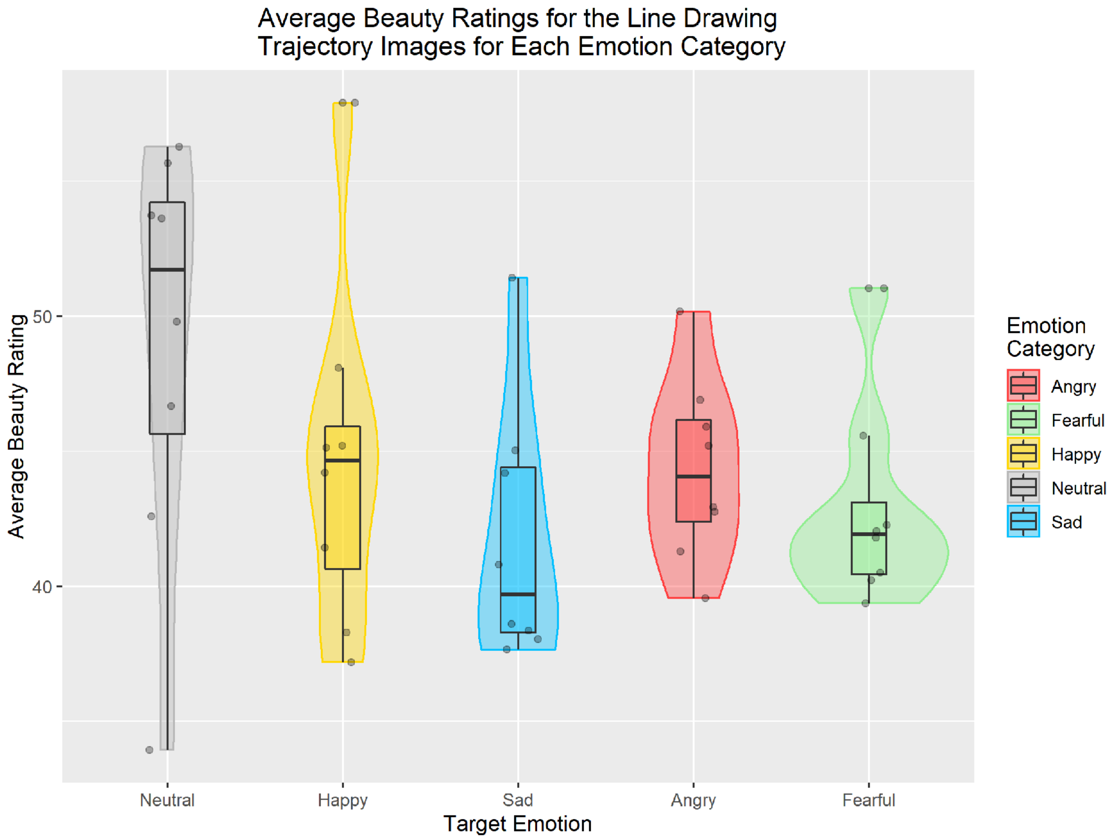


Figure 4.5: A visualisation of the average beauty ratings provided by participants for trajectory images from each of the emotion categories.

As with liking, trajectory images depicting wrist movement (M= 46.05 ± 5.71) received more positive beauty appraisals than ankle movement images (M= 43.02 ± 5.57), but this difference was not found to be significant: t(37.978)= -1.70, p= 0.098.

Beauty x Dance Experience Results

An ANOVA was conducted to explore whether different levels of previous dance experience influenced beauty ratings provided for the movement trajectory images. Advanced (M= 61.25 ± 34.76) and intermediate (M= 52.88 ± 26.23) dancers provided higher overall beauty ratings than beginner dancers (M= 47.98 ± 23.70) and non-dancers (M= 46.07 ± 14.22). However, these differences were not statistically significant: F(3,54)= 0.59, p= 0.626.

A hierarchical multiple regression was conducted to explore whether physical dance experience, observational dance experience, and the number of dance styles known influenced overall beauty ratings. A more detailed breakdown of the results of this analysis can be found in Table 4.6.

Table 4.6: Results of the hierarchical multiple regression analyses exploring the role of dance experience in beauty ratings provided for the trajectory drawings.

Model	Summary	Predictor	B	T	Std Error
Model 1	$F(1,56)= 1.01^{n.s.}$, $R^2_{Adj}= 0.0002$, RSE= 20.74	Intercept	45.39		
		Physical Experience	14.03	$1.01^{n.s.}$	13.95
Model 2	$F(2,55)= 5.56^{**}$, $R^2_{Adj}= 0.138$, RSE= 19.26	Intercept	35.81		
		Physical Experience	2.46	$0.183^{n.s.}$	13.46
		Observational Experience	34.42	3.150^{**}	10.93
Model 3	$F(3,54)= 3.68^{*}$, $R^2_{Adj}= 0.124$, RSE= 19.42	Intercept	35.92		
		Physical Experience	-0.87	$-0.052^{n.s.}$	16.70
		Observational Experience	33.85	3.039^{**}	11.14
		Number of Styles	3.58	$0.343^{n.s.}$	10.46

N = 50; n.s. $p > 0.05$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Model 1: beauty rating − physical dance experience factor

Model 2: beauty rating − physical dance experience factor + observational dance experience factor

Model 3: beauty rating − physical dance experience factor + observational dance experience factor + number of dance styles a participant has experience with

The model at stage one only explained 0.02% of the variance in beauty ratings, and was not a significantly better fit for the data than the null model: $F(1,56)= 1.101$, $p= 0.319$. However, introducing observational dance experience to the model at stage two explained an additional 13.7% of variance in the beauty ratings, and this model was found to fit the data significantly better than the null model: $F(2,55)=5.55$, $p< 0.01$. Adding the number of dance styles known as a factor in the maximal model actually decreased the amount of variance explained by the model to 12.4%, but it was a significantly better fit for the data than the null model: $F(3,54)= 3.68$, $p< 0.05$. However, the maximal model did not fit the data significantly better than the model at stage two ($p= 0.73$). Across the stage two modal and the maximal model, only observational dance experience had a significant impact on the average beauty rating provided by participants. Partial support, then, was found for the prediction that dance experience should influence beauty ratings provided for the trajectory line drawings. However, the prediction that physical dance experience would play a larger role than observational experience was not supported, as we found only observational dance experience increased how beautiful participants found the movement trajectory images to be.

Beauty x Empathy Results

A Spearman's correlation was conducted to explore whether trait empathy (as measured by the Toronto Empathy Scale; Spreng et al., 2009) influenced the beauty ratings provided by participants for the trajectory line drawings. A small negative relationship was observed between empathy score and average beauty ratings, and this relationship was found to be statistically significant: $r(56)= -0.291$, $p< 0.05$. These results show that in this experiment, participants with higher trait empathy actually provided less positive beauty ratings than participants with lower empathy. Therefore, we did not obtain support for the prediction that high empathy mediates more positive aesthetic appraisals of movement depicted in this form.

Movement Video vs Trajectory Image Analysis

Liking Results

To explore whether liking ratings aligned for the trajectory images (this experiment) and the corresponding movement videos (Smith & Cross, under review) they were created from, a Pearson's correlation analysis was conducted to explore whether there was a relationship between liking ratings for the McNorm Library movements across both formats. A weak negative relationship in liking scores was observed, as participants generally appeared to prefer trajectory images made from the least liked movement videos, and disliked the trajectory images from movement videos which were well liked. However, this relationship was not significant: $r(18)= -0.381$, $p= 0.097$. A visualisation of this relationship can be observed in Figure 4.6 below.

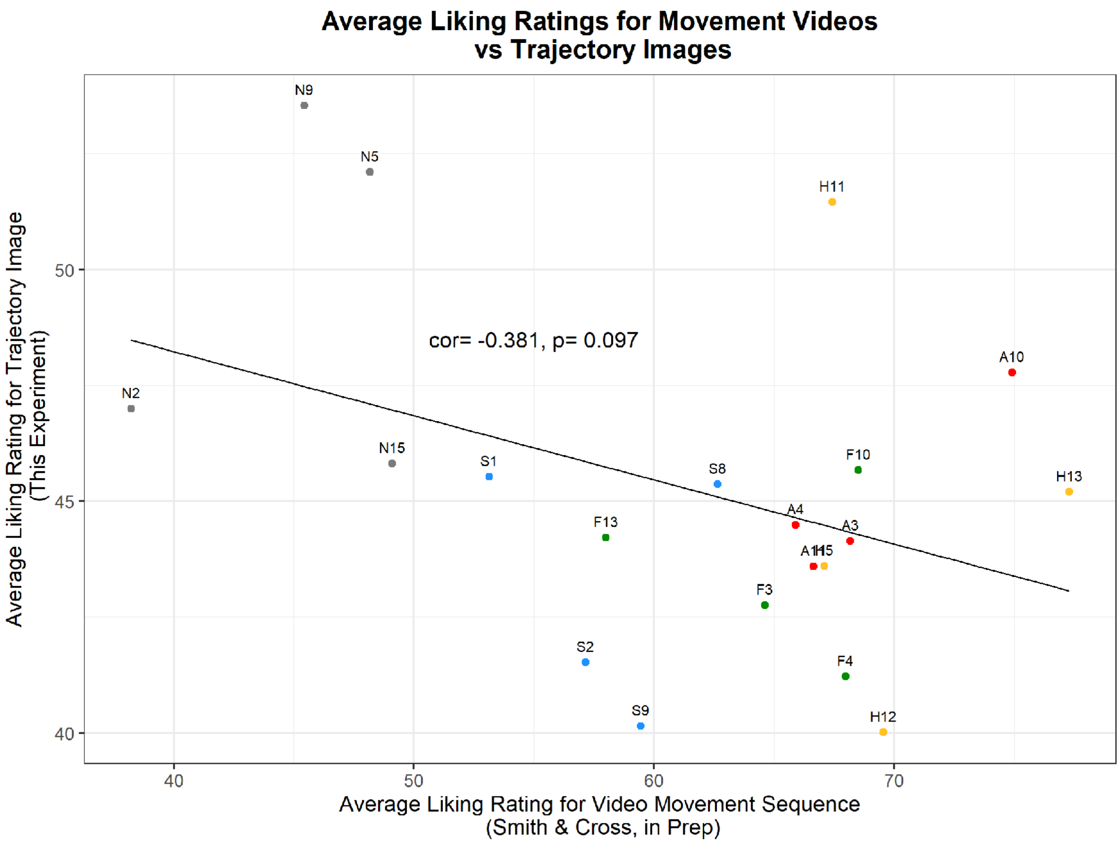


Figure 4.6: A visualisation of the relationship between liking ratings for the trajectory line drawings (this experiment), and liking ratings for the videos from which the images were created (Smith & Cross, unpublished).

Beauty Results

A Pearson's correlation analysis was also conducted to explore whether there was a relationship between beauty ratings for the McNorm Library movements across both formats (trajectory images in this experiment; movement videos from Smith & Cross, under review). Again, a non-significant weak negative relationship was observed: $r(18)= -0.387$, $p= 0.092$. A visualisation of this relationship can be observed in Figure 4.7.

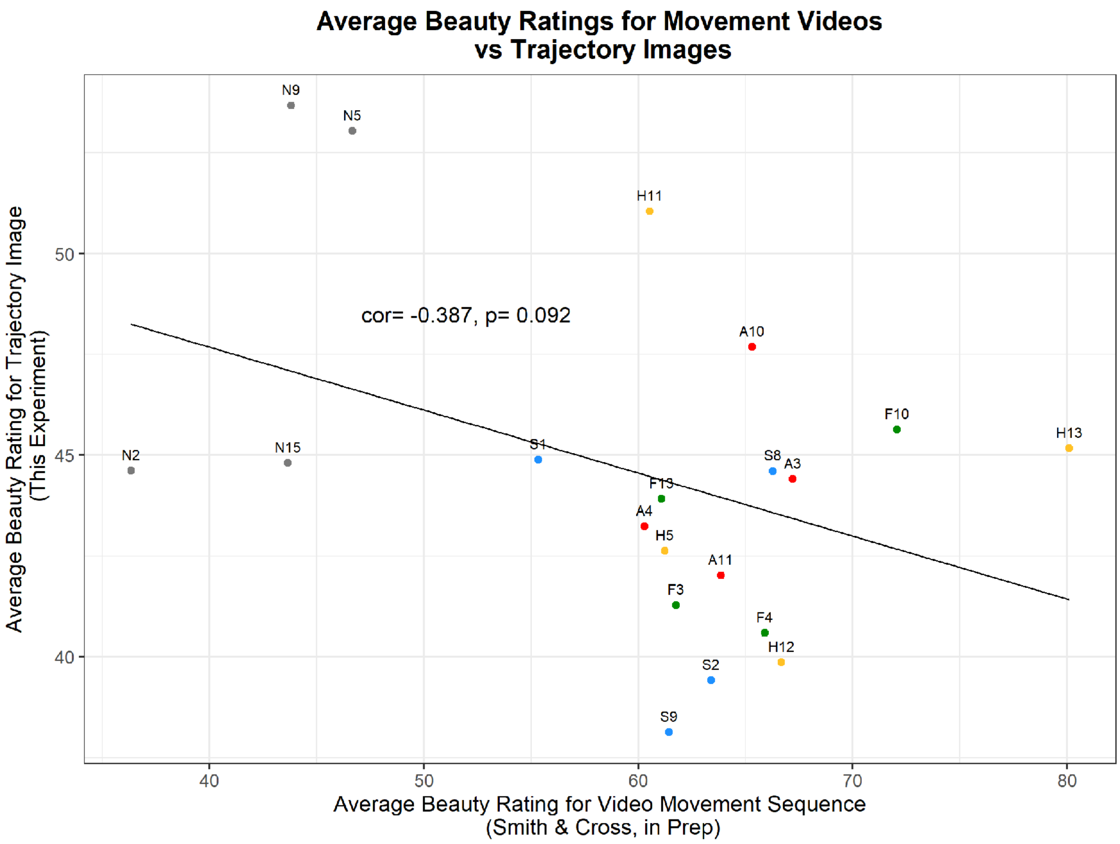


Figure 4.7: A visualisation of the relationship between beauty ratings for the trajectory line drawings (this experiment), and beauty ratings for the videos from which the images were created (Smith & Cross, under review).

EXPERIMENT 1: DISCUSSION

The primary aim of this experiment was to explore whether creating continuous line drawings from the trajectories of emotionally expressive movement sequences could provide an alternative way to visualise human motion for social perception research. More specifically, it was hoped that these line drawings would provide a novel alternative to PLDs for researchers who wish to examine human motion through stimuli with a reduced level of morphological information and confounding visual features. To test whether these drawings are a viable way to depict human movement, an emotion recognition task was conducted to explore whether intended emotion expressions could be recognised from these drawings at greater than chance level. However, results obtained from judgements of 58 observers showed that recognition was significantly greater than chance for only the neutral stimuli. Across two validation experiments presented in the McNorm Library manuscript (Smith & Cross, 2022), where movement was represented in the form of PLDs, the average recognition rate obtained across all stimuli was around 50.2%. In the present experiment, when these same movements were converted into trajectory line drawings, the average recognition rate obtained across all stimuli was only around 20.5%, which is almost exactly at chance. Therefore, we failed to obtain support for the primary experimental hypothesis, and conclude that these trajectory line drawings, at least in their current form, are not a suitable way to visualise the human body in motion.

Additional evidence to suggest that these trajectory images do not function in the same way as other movement stimuli (e.g., live movement performance, FLDs, PLDs) can be found when examining the empathy and dance experience results. A number of studies have shown that movement expertise can influence the affective and aesthetic value judgements assigned to whole-body movements (Christensen et al., 2016; Christensen et al., 2019; Kirsch et al., 2015; Calvo-Merino et al., 2008). In the present experiment, however, we did not find evidence to suggest that either physical or observational dance experience had a significant impact on recognition accuracy, or beauty and liking ratings assigned to the trajectory line drawings. Similarly, a wealth of previous research suggests there is an inherent link between trait empathy and these kinds of social judgements (Arnold et al., 2014; Crozier & Greenhalgh, 1992; Wilkinson et al., 2021; Gernot et al., 2018), but we did not find evidence of this relationship within our sample. Therefore, as we also failed to obtain support for either of the secondary experimental predictions, we conclude that creating this type of continuous line drawing from movement trajectories is not an appropriate way to visualise human movement in social perception research.

The idea that observers can recognise human movement from simple line drawings was first proposed by Chamberlain and colleagues. In a series of experiments, they observed that participants could distinguish between simple graffiti-style shapes produced from human drawing actions, and those generated computationally (Chamberlain et al., 2022). However, it should be noted that in the first experiments, participants observed videos of the drawings being produced, rather than static images of the final products. The authors concluded that participants could identify humanlike characteristics present in the kinematics of the line-drawing videos (i.e., imperfect circles, uneven speed profiles), and that these motion characteristics were important for the perception of human movement within the images. Kinematics too play a significant role in affective perception from human movement videos. It has been found that differences in speed, fluidity, and contraction (among other parameters) assist participants in deciding which emotional expressivity to assign to everyday actions (e.g., walking, knocking; Gross et al., 2012; Barliya et al., 2013; Gross et al., 2010) and more complex motion (e.g., human dance movement; Orlandi et al., 2020; White & Egermann, 2021; Van Dyck et al., 2014). With this in mind, it is possible that for participants to associate trajectory line drawings with the movements they were created from, this necessitates the use of dynamic drawings rather than static images. Therefore, while recognition rates obtained in this experiment were relatively unsuccessful (compared to those obtained for other stimulus presentation methods; i.e., FLDs, PLDs), it is possible that recognition rates may improve if participants could observe the kinematics present in the dynamic production of these drawings. Future research should be conducted to explore this possibility.

It is also possible that the ability to recognise human movement from line drawings is more pronounced in participants with art expertise. Indeed, support for this notion was obtained in the later experiments conducted in the project by Chamberlain and colleagues (2022). When art students and art-naïve observers were asked to identify the human-made drawings from static image pairs (human-made and computer-generated), only the art experienced participants could successfully identify which line drawings were created as a result of human drawing actions. In this experiment, as we were examining these continuous line drawings as an alternative way to visualise human movement, we only explored emotion recognition judgements made by participants with various degrees of movement experience, but it may have also been beneficial to explore affective judgements made by participants with different levels of previous art expertise.

On the note of expertise, it has been found that experience with a particular artform can produce a variety of skills related to that specific medium. It has been found that experienced drummers are more attuned to audio-visual synchrony than novice drummers (Petrini et al., 2009; Petrini et al., 2011). Additionally, it has been observed that both musicians and music-lovers possess a series of specific empathy-related skills that facilitate expertise related differences in the cognitive processing of music (Kreutz et al., 2008). A number of behavioural and neurocognitive changes have also been found to result from both long- and short-term dance training. For example, several papers have observed that dancers are more accurate in emotion recognition tasks for human movement stimuli than those without training (Christensen et al., 2016; Christensen et al., 2019), and it has been found that movement expertise can guide visual attention when observing others in motion (Stevens et al., 2010; Moran et al., 2002; Betram et al., 2013). In the study conducted by Chamberlain and colleagues (2022), although the experiment centred around the detection of humanlike motion from graffiti-style drawings, it is likely that the art-experienced participants relied on knowledge about the image's visual characteristics (rather than inferences about inferred motion) to assign social judgements to the stimuli. It is possible that participants in the present experiment were evaluating these images in a similar way; using visual cues from the continuous line drawings as a guide for their emotion labelling and aesthetic value judgements. Therefore, it may be useful to explore the possibility that these trajectory line drawings function as pieces of visual art, in their own right, rather than as a proxy for human motion.

In light of the findings from the first experiment, we wished to conduct a follow-up experiment to explore the impact of computational image properties on the emotion labelling and aesthetic judgements provided by participants for these continuous line drawings.

EXPERIMENT 2: INTRODUCTION

The dominant model for the role of size in aesthetic appreciation of visual arts postulates that bigger is better, and there is a wealth of evidence to support this notion. For example, in a series of experiments participants were presented with images of paintings which varied in terms of size and were asked to rate each of the paintings along a series of aesthetic dimensions (e.g., interest, awe). Participants provided significantly more positive aesthetic appraisals of the paintings presented at a larger scale than for smaller versions of those paintings (Seidel & Prinz, 2018). In another experiment, participants were presented with a series of basic 2D and 3D shapes varying in size. Small shapes were approximately one-sixteenth the size of the medium shapes, and large shapes were around four times the size of the medium shapes (in area). They found that both medium and large versions of shapes were liked significantly more than small versions of the same shapes (Chen et al., 2016). Further there is evidence to suggest that this preference for larger stimuli persists across the lifespan, as Silvera and colleagues have found a preference for larger geometric shapes and alphanumeric characters in both adults and 3-year-old children (Silvera et al., 2002). Therefore, these results add support to the idea that for visual artworks 'bigger is better', or the inverse that 'smaller is worse', in terms of aesthetic value.

Additionally, in one of the experiments conducted by Silvera and colleagues, adult participants were asked to assign a meanings to the shapes, in addition to providing aesthetic judgements (Silvera et al., 2002). These meanings were coded as positive or negative, and it was found that in addition to providing more positive aesthetic value judgements, participants assigned more positive meanings to the larger versions of shapes (Silvera et al., 2002). Evidence for the role of image size in aesthetic and affective processing has also been observed for simple drawings, as Bromgard and colleagues observed that aesthetic preference judgements were concurrently influenced by both the size and valence of these stimuli (Bromgard et al., 2013). Therefore, considered together, from these results it would appear that bigger is indeed better for the aesthetic evaluation of traditional visual artworks, and of simple 2D polygons, but also that size may be involved in mediating the affective processing of visual art stimuli.

Curvature

As is the case with size, a wealth of evidence points to a single dominant theory for the relationship between curvature and aesthetic evaluation for simple and complex visual artworks, in that there appears to be a clear preference for curvature over angularity. In one experiment, participants observed a series of curved and angular versions of simple shapes and random polygons, and were asked to rate each of these items in terms of pleasantness. After controlling for symmetry, prototypicality, and balance, they found that participants preferred the curved versions of these shapes than the more angular versions of the same shapes (Silvia & Barona, 2009). A preference for smoothed contours has been observed in participants from different cultural backgrounds, providing support for the idea that a preference for curvature in polygons extends beyond traditional western samples, and may indicate the generalisability of this preference in visual artworks (Gómez-Puerto et al., 2018).

Additionally, there is evidence to suggest that curvature, like size, mediates both aesthetic and affective judgements for basic shapes, and that these factors may interact. Leder and colleagues aimed to explore whether curvature preferences for neutral shapes and objects could be replicated in objects with positive and negative valence. In this experiment, participants observed a series of curved versus sharp versions of stimuli with positive (e.g., teddy bear, chocolate), negative (e.g., snake, missile) and neutral (e.g., stapler, lantern) valence, and provided liking ratings for each of the objects. They found that participants significantly preferred the curved versions of objects and shapes, but only for those stimuli with positive or neutral valence (Leder et al., 2011). These results suggest that curvature has simultaneous effects on both aesthetic and affective perception of objects. Further evidence for the role of curvature on the affective perception of images can be observed in the work of Lu and colleagues. Using the International Affective Picture System (IAPS; Verschuere et al., 2001) dataset, they investigated how different properties of images (including roundness, and complexity) influenced the emotions aroused in participants. They observed a distinct profile of image properties associated with several emotions. More specifically, roundness was associated with both anger and sadness, but fear-inducing shapes were jagged, rather than rounded (Lu et al., 2012).

Complexity

While the impact of size and curvature seems to be relatively clear, the role of complexity in aesthetic evaluation is highly contested. Two main accounts have been suggested to explain this relationship. First, it has been argued that there is a negative linear relationship between complexity and preference judgements. Termed the perceptual fluency account, this theory suggests that observers will prefer simpler shapes, as processing these stimuli requires significantly less cognitive demand. A number of studies have provided support for the role of perceptual fluency in aesthetic preference, through indirect measurements of complexity. For example, Reber and colleagues have observed a preference for images which participants have been repeatedly exposed to, using the mere exposure effect as a proxy for lower perceptual complexity (Reber et al., 1998). Additionally, Chen and colleagues observed that both symmetry and familiarity (argued to be indirect measures of complexity) were preferred in pattern stimuli (Chen et al., 2011). While such studies provide evidence to support a link between these facets of perceptual fluency (i.e., repeated exposure, familiarity and symmetry; as reflections of stimulus complexity) and aesthetic value judgements, it is unclear how these results translate to more direct computational measures of image complexity.

Results from recent research, taking into consideration more explicit measures of image complexity, suggest a different account for its influence on aesthetic evaluation. More specifically, results from several studies suggest that this relationship follows an inverse-U shaped trajectory. In one such series of experiments, participants provided beauty ratings for a series of random density patterns with progressively larger numbers of elements. The authors reported that patterns with higher complexity received more positive beauty ratings until a saturation point was achieved, after which the aesthetic value judgements became more negative (Friedenberg & Liby, 2016). This was also observed in their follow up experiment which included controls for patterns with fixed densities. Their data were later reanalysed by a different research team, who included some additional measures (e.g., entropy, and algorithmic complexity), and found that the U-shaped curve of preference could be explained by other features present in the images (low redundancy, balanced patterns) which were not reflected in the original measures of algorithmic complexity. These authors conclude by highlighting the problems posed by fluidity in what constitutes complexity in visual aesthetic stimuli (Gauvrit et al., 2017). Indeed, based on the breadth of methodologies employed, and variation in reported results, it would appear that the confusion around how to empirically, and computationally, define "complexity" is currently limiting our understanding of its role in the aesthetic processing of visual stimuli. This, therefore, suggests additional research attention is needed in this area, and perhaps requires the inclusion of a wide array of complexity measures to bring some clarity to the nature of this relationship.

Experimental Hypotheses

In light of the previous findings, when considering our stimuli as pieces of visual art we can make a series of predictions about the impact of these visual properties (size, curvature, and complexity) on the emotion labels and aesthetic judgements provided for our continuous line drawings.

First, it was predicted that observers would provide beauty and liking ratings which reflect a general preference for larger, over smaller, stimuli. It was also expected that size of the images may influence perceptions about their emotional content. However, the specifics of this relationship were unclear in the previous literature, so no specific directional predictions were made.

Secondly, it was predicted that participants would generally prefer curved, over more jagged, trajectory images. Therefore, we expected higher beauty and liking ratings for images which featured softer curves. Additionally, as Lu and colleagues (2012) observed that the roundness or jaggedness of a shape influenced specific emotion perceptions, we further predicted that images which are perceived to convey anger and sadness would be more rounded, while images which are considered to depict fear would be more jagged.

As the role of image complexity in aesthetic judgements is highly contested, no specific predictions were made for this property with regard to liking and beauty ratings for the trajectory images. However, one of two patterns may emerge. Participants may prefer simpler trajectory images (the perceptual fluency account), or complexity may have a bell-shaped impact on these judgements.

EXPERIMENT 2: METHODS

The emotion labels and aesthetic value judgements assigned by participants (N= 58) to the black and white line drawing images in the first experiment were reanalysed in relation to a series of computational properties of the images. These properties included 1 measure of size, 1 measure of curvature and 6 measures of image complexity. A more detailed breakdown of how each of the specific properties were computed can be found in Supplementary Section B.

EXPERIMENT 2: RESULTS

A series of visual properties were extracted from the trajectory images. The mean values for each of these properties (averaged across all images) can be observed in Table 4.7 below.

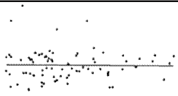
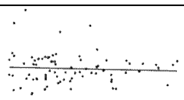
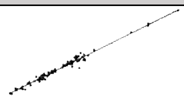
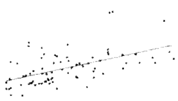
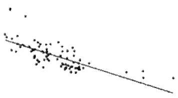
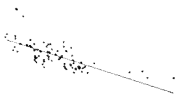
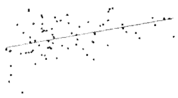
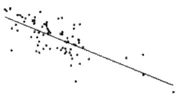
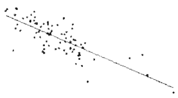
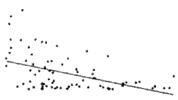
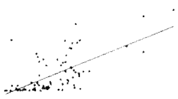
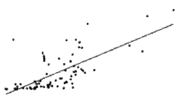
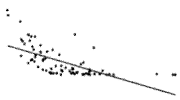
Table 4.7: An overview of the average values for each of the image properties included in this analysis.

Image Property	Average Value
Curvature	3.78 (± 3.37)
Drawing Size	0.30 (± 0.15)
Perimeter (# of pixels)	12598.46 (± 3584.14)
Ink Size (# of black pixels)	77197.49 (± 28751.01)
Jpeg Bytes	150584.30 (± 40883.21)
Perimeter to Ink Ratio	0.18 (± 0.08)
Spectral Complexity 1D	-5.65 (± 3.09)
Spectral Complexity 2D	0.0019 (± 0.0022)

Subjective Emotion Results

Based on the results of the five-option forced-choice emotion recognition task, each image pair was assigned a percentage for how often participants classified that image as neutral, happy, sad, angry or fearful. These values ranged from 0% (no assignment of that particular emotion to the image pair) to 100% (every classification of the image pair was for that particular emotion), and a specific attribution value for each emotion category was assigned to each image pair. These values were each used as the dependant variable in a multiple regression for each perceived emotion category, with the image properties included as the independent variables, to examine the impact of these properties on the perception of each particular emotion for the trajectory images. The image properties varied widely in scale, so min-max normalisation was applied across each of the values prior to the regression analyses. In addition, as it was hoped that all complexity measures would be comparable in their suitability for examining the complexity of the line drawings, as including all of these factors into the regression model may introduce the problem of multicollinearity. Indeed, we observed multiple significant correlations between the different complexity measures. A more detailed breakdown of the relationships between these different factors can be observed in Table 4.8.

Table 4.8: A scatterplot matrix which shows the relationship between each of the different measures of complexity calculated in this experiment.

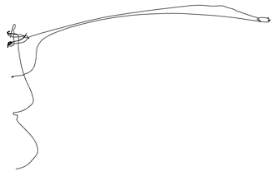
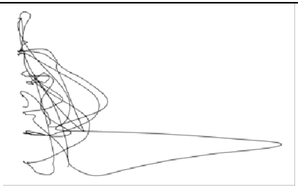
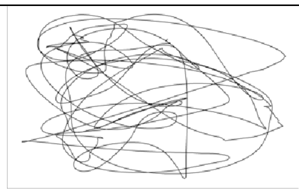
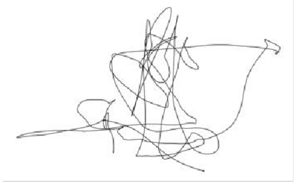
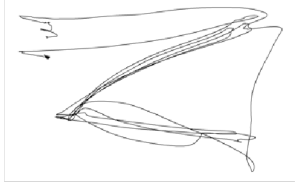
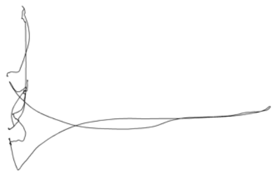
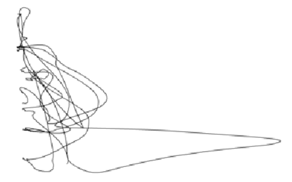
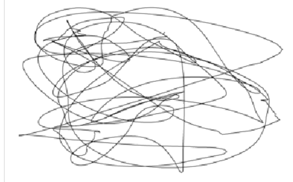
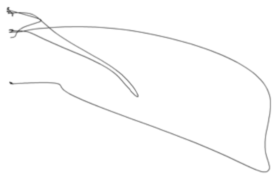
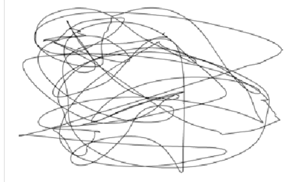
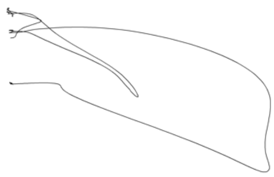
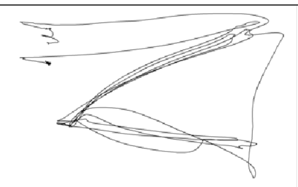
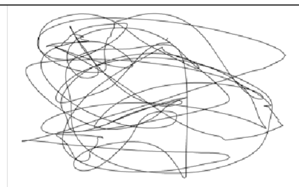
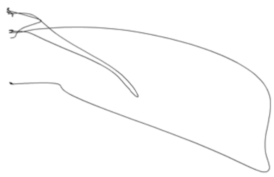
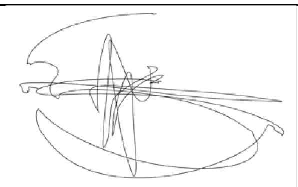
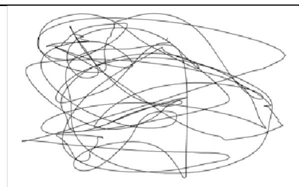
	Perimeter Size	Ink Size	Jpeg Bytes	Perimeter to Ink Ratio	1D Spectral Complexity	2D Spectral Complexity
Perimeter Size	-	r_s= 0.03$^{n.s.}$	r_s= -0.04$^{n.s.}$	r_s= 0.56***	r_s= 0.42***	r_s= -0.42***
Ink Size		-	r_s= 0.97***	r_s= -0.76***	r_s= -0.68***	r_s= 0.64***
Jpeg Bytes			-	r_s= -0.79***	r_s= -0.74***	r_s= 0.70***
Perimeter to Ink Ratio				-	r_s= 0.81***	r_s= -0.77***
1D Spectral Complexity					-	r_s= -0.97***
2D Spectral Complexity						-

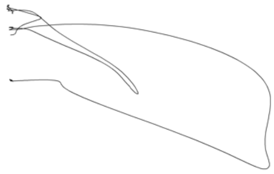
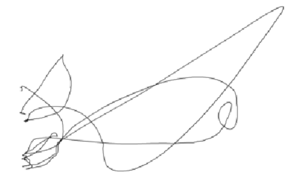
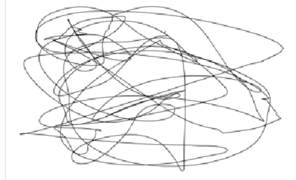
Therefore, as there were a number of significant correlations between each of these individual factors, before proceeding with the regression analyses we decided to create a general complexity factor. Perimeter size was strongly correlated with some factors (perimeter to ink ratio, 1D spectral complexity, and 2D spectral complexity), but not with others (ink size, and jpeg bytes). Due to the high degree of consistency between the other measures of complexity, which was not as clear for this factor, it was decided that the general complexity measure would be created by averaging across values for all of other normalised complexity factors, but not perimeter size. For both the perimeter to ink ratio and 1D spectral complexity, smaller values are associated with greater complexity, while all other factors represent greater complexity with higher values; this explains why these two factors are negatively correlated with all of the other measures. Therefore, when adding the former two factors into the averaging formula they were first multiplied by a factor of -1.

$$\begin{matrix} General \\ Complexity \\ Factor \end{matrix} = \frac{Ink\ Size + Jpeg\ Bytes - Perimeter\ to\ Ink\ Ratio - 1D\ Spectral\ Complexity + 2D\ Spectral\ Complexity}{(Number\ of\ Factors\ Included = 5)}$$

It is also important to note here that each of the complexity measured yielded a slightly different ordering of the continuous line drawing stimuli. A visualisation of the line drawing with the minimum, median, and maximum values for each of the original complexity factors, and the new general complexity factor, can be observed for comparison in Table 4.9.

Table 4.9: A visualisation of trajectory images from our stimuli set with the lowest, median, and highest complexity for each of the individual complexity factors, and for the general complexity factor we propose.

	Least Complexity	Median Complexity	Most Complexity
Ink Size			
Perimeter Size			
Jpeg File Size			
Perimeter to Ink Ratio			
1D Spectral Complexity			
2D Spectral Complexity			

Normalised General Complexity Factor			

To explore the contribution of each of the image properties (i.e., curvature, drawing size, general complexity factor) on the perception of each of the emotions (neutral, happy, sad, angry, and fearful) for the continuous line drawings, a series of hierarchical multiple regression analyses were conducted. As, to our knowledge, no research has explored the impact of these properties on affective and aesthetic perception of this type of stimuli, the order in which these features were added to the model was based on both the frequency and consistency with which they are reported in the previous literature using simple shapes and visual artworks. First, the impact of size appears to be the most consistently reported in the previous literature, therefore, this factor was introduced in the Stage 1 model. The impact of curvature also seems to be relatively clear from previous work in this area, therefore this factor was included in the model next at Stage 2. The relationship between perception and complexity seems to be less clear (compared to the two aforementioned factors), therefore, the general complexity factor was included at Stage 3. Finally, while there is some evidence to suggest that the tilt or orientation of a shape influences perceptions, it is unclear how this finding applies to more complex shapes (like our continuous line drawings).

Table 4.10: Summary of hierarchical multiple regression analysis for the image properties predicting the perception of neutrality from the continuous line drawings.

NEUTRAL	Summary	Predictor	B	T	Std Error
Model 1	$F(1,78)= 4.455^{*}$, $R^2_{Adj}= 0.042$, RSE= 7.764	Intercept	29.12		
		Drawing Size	-6.97	-2.11^{*}	3.30
Model 2	$F(2,77)= 2.56^{n.s}$, $R^2_{Adj}= 0.038$, RSE= 7.766	Intercept	30.08		
		Drawing Size	-7.71	-2.25^{*}	3.43
		Curvature	-5.67	$-0.82^{n.s.}$	6.87

Model 3	$F(3,76)= 10.44^{***}$, $R^2_{Adj}= 0.264$, RSE= 6.805	Intercept	34.02		
		Drawing Size	20.42	3.19^{**}	6.41
		Curvature	-6.99	$-1.16^{n.s.}$	6.02
		General Complexity Factor	-47.00	-4.96^{***}	9.47

Model 1: Percentage of Neutral Classifications ~ Drawing Size
Model 2: Percentage of Neutral Classifications ~ Drawing Size + Curvature
Model 3: Percentage of Neutral Classifications ~ Drawing Size + Curvature + Complexity Factor
$p> 0.05^{n.s.}$, $p< 0.05^{*}$, $p< 0.01^{**}$, $p< 0.001^{***}$

The hierarchical multiple regression for the perception of neutral emotion (detailed in Table 4.10) revealed that at Stage one, the regression model was a significantly better fit for the data than the null model $F(1,78) = 4.46$, $p < 0.05$, and that the drawing size factor explained 4.2% of the variation in the attribution of the neutral label to the line drawings. Introducing the curvature factor at Stage two generated a model which was not a significantly better fit for the data than the null model $F(2,77) = 2.56$, $p = 0.840$, and adding curvature decreased the variation in attribution of neutral explained by the model by around 0.4%. Finally, the maximal model added the general complexity factor, and the addition of this predictor did significantly improve the fit of the model to the data $F(3,76) = 6.81$, $p <0.001$. Additionally, after adding the general complexity factor, the model explained an additional 22.6% of the variance in the perception of neutral emotion from the line drawings (compared to the Stage two model). These results indicate that participants were more likely to attribute the neutral option to a continuous drawing if it was larger in size, and this judgement too was more likely if the drawing was less complex.

Table 4.11: Summary of hierarchical multiple regression analysis for the image properties predicting the perception of happiness from the continuous line drawings.

HAPPY	Summary	Predictor	B	T	Std Error
Model 1	$F(1,78)= 3.68^{n.s.}$, $R^2_{Adj}= 0.033$, RSE= 10.970	Intercept	24.05		
		Drawing Size	8.95	$1.92^{n.s.}$	4.67

Model 2	$F(2,77) = 2.24^{n.s}$, $R^2_{Adj} = 0.030$, RSE= 10.99	Intercept	25.52		
		Drawing Size	7.83	$1.62^{*.}$	4.84
		Curvature	-8.69	$-0.90^{n.s.}$	9.71
Model 3	$F(3,76) = 2.44^{n.s.}$ $R^2_{Adj} = 0.052$, RSE= 10.86	Intercept	27.62		
		Drawing Size	22.83	2.23^{*}	10.23
		Curvature	-9.39	$-0.98^{n.s.}$	9.61
		General Complexity Factor	-25.08	$-1.66^{n.s.}$	15.11

Model 1: Percentage of Happy Classifications ~ Drawing Size
Model 2: Percentage of Happy Classifications ~ Drawing Size + Curvature
Model 3: Percentage of Happy Classifications ~ Drawing Size + Curvature + Complexity Factor
$p > 0.05^{n.s.}$, $p < 0.05^{*}$, $p < 0.01^{**}$, $p < 0.001^{***}$

The hierarchical multiple regression for the perception of happiness (Table 4.11) revealed that at Stage one, the regression model was not a significantly better fit for the data than the null model $F(1,78) = 3.68$, $p = 0.059$, and that the drawing size factor only explained 3.3% of the variation in the attribution of the happy label to the line drawings. Introducing the curvature factor at Stage two generated a model which was also not a significantly better fit for the data than the null model $F(2,77) = 2.24$, $p = 0.114$, and adding curvature decreased the variation in attribution of happiness explained by around 0.3%. Finally, the maximal model added the general complexity factor, but the addition of this predictor also did not create a model which fit the data better than the null model $F(3,76) = 2.44$, $p = 0.071$, and only explained an additional 2.2% of the variance in the perception of happiness from the line drawings (compared to the Stage two model). These results indicate that none of the image properties (i.e., size, curvature, or complexity) had an impact on whether participants perceived happiness from the continuous line drawings.

Table 4.12: Summary of hierarchical multiple regression analysis for the image properties predicting the perception of sadness from the continuous line drawings.

SAD	Summary	Predictor	B	T	Std Error
Model 1	$F(1,78) = 7.35^{**}$, $R^2_{Adj} = 0.074$, RSE= 4.753	Intercept	16.70		
		Drawing Size	-5.48	$2.71^{**.}$	2.02
Model 2	$F(2,77) = 3.64^{*}$, $R^2_{Adj} = 0.063$, RSE= 4.783	Intercept	16.81		
		Drawing Size	-5.57	$-2.64^{*.}$	2.11
		Curvature	-0.66	$-0.16^{n.s.}$	4.23
Model 3	$F(3,76) = 9.16^{***}$, $R^2_{Adj} = 0.237$, RSE= 4.317	Intercept	18.98		
		Drawing Size	9.91	2.44^{*}	4.07
		Curvature	-1.39	$-0.37^{n.s.}$	3.82
		General Complexity Factor	-25.86	-4.31^{***}	6.01

Model 1: Percentage of Sad Classifications ~ Drawing Size

Model 2: Percentage of Sad Classifications ~ Drawing Size + Curvature

Model 3: Percentage of Sad Classifications ~ Drawing Size + Curvature + Complexity Factor

$p > 0.05^{n.s.}$, $p < 0.05^{*}$, $p < 0.01^{**}$, $p < 0.001^{***}$

The hierarchical multiple regression for the perception of sadness (Table 4.12) revealed that at Stage one, the regression model was a significantly better fit for the data than the null model $F(1,78) = 4.75$, $p < 0.01$, and that the drawing size factor explained 7.4% of the variation in the attribution of the sad label to the line drawings. Introducing the curvature factor at Stage two generated a model which was also a significantly better fit for the data than the null model $F(2,77) = 3.64$, $p < 0.05$, but decreased the variation in attribution of sadness explained by the model by around 1.1%. Finally, the maximal model added the general complexity factor, and the addition of this predictor did significantly improve the fit of the model to the data $F(3,76) = 9.16$, $p < 0.001$. After adding the general complexity factor, the model explained an additional 17.4% of the variance in the perception of sadness from the line drawings (compared to the Stage two model). As with the perception of neutral emotion, these results indicate that participants were more likely to attribute the sad option to a continuous line drawing if it was larger in size, and this judgement too was more likely if the drawing was less complex. It should be noted that for recognition of neutral, the model revealed a larger contribution of the increase in drawing size than for the recognition of sadness.

Table 4.13: Summary of hierarchical multiple regression analysis for the image properties predicting the perception of anger from the continuous line drawings.

ANGRY	Summary	Predictor	B	T	Std Error
Model 1	$F(1,78)= 5.29^{*}$, $R^2_{Adj}= 0.052$, RSE= 10.99	Intercept	13.00		
		Drawing Size	10.75	2.30^{*}	4.67
Model 2	$F(2,77)= 2.70^{n.s.}$, $R^2_{Adj}= 0.041$, RSE= 11.05	Intercept	12.32		
		Drawing Size	11.27	2.32^{*}	4.87
		Curvature	3.99	$0.41^{n.s.}$	9.76
Model 3	$F(3,76)= 16.95^{***}$, $R^2_{Adj}= 0.377$, RSE= 8.90	Intercept	5.55		
		Drawing Size	-37.08	-4.42^{***}	8.39
		Curvature	6.27	$0.80^{n.s.}$	7.88
		General Complexity Factor	80.80	6.52^{***}	12.39

Model 1: Percentage of Angry Classifications ~ Drawing Size

Model 2: Percentage of Angry Classifications ~ Drawing Size + Curvature

Model 3: Percentage of Angry Classifications ~ Drawing Size + Curvature + Complexity Factor

$p> 0.05^{n.s.}$, $p< 0.05^{*}$, $p< 0.01^{**}$, $p< 0.001^{***}$

The hierarchical multiple regression for the perception of anger (Table 4.13) revealed that at Stage one, the regression model was a significantly better fit for the data than the null model $F(1,78) = 5.29$, $p < 0.05$, and that the drawing size factor explained 5.1% of the variation in the attribution of anger to the line drawings. Introducing the curvature factor at Stage two generated a model which was not a significantly better fit for the data than the null model $F(2,77) = 2.70$, $p = 0.073$, and adding curvature decreased the variation in attribution of anger explained by the model by around 1%. Finally, the maximal model added the general complexity factor, and the addition of this predictor did significantly improve the fit of the model to the data $F(3,76) = 16.95$, $p < 0.001$. Additionally, after adding the general complexity factor, the model explained an additional 33.6% of the variance in the perception of anger from the line drawings (compared to the Stage two model). These results indicate that participants were more likely to attribute the anger option to a continuous drawing if it was smaller in size, and this judgement too was more likely if the drawing was more complex.

Table 4.14: Summary of hierarchical multiple regression analysis for the image properties predicting the perception of fear from the continuous line drawings.

FEARFUL	Summary	Predictor	B	T	Std Error
Model 1	$F(1,78)= 10.15^{**}$, $R^2_{Adj}= 0.104$, RSE= 5.35	Intercept	17.14		
		Drawing Size	-7.25	-3.19^{**}	2.28
Model 2	$F(2,77)= 8.27^{***}$, $R^2_{Adj}= 0.155$, RSE= 5.20	Intercept	15.27		
		Drawing Size	-5.82	-2.54^{*}.	2.29
		Curvature	11.03	2.40^{*}.	4.59
Model 3	$F(3,76)= 7.87^{***}$, $R^2_{Adj}= 0.207$, RSE= 5.03	Intercept	13.83		
		Drawing Size	-16.08	-3.39^{**}	4.74
		Curvature	11.51	2.59^{*}	4.45

	General Complexity Factor	17.14	2.45^{*}	7.01

Model 1: Percentage of Fearful Classifications ~ Drawing Size

Model 2: Percentage of Fearful Classifications ~ Drawing Size + Curvature

Model 3: Percentage of Fearful Classifications ~ Drawing Size + Curvature + Complexity Factor

$p > 0.05^{n.s.}$, $p < 0.05^{*}$, $p < 0.01^{**}$, $p < 0.001^{***}$

The hierarchical multiple regression for the perception of fear (Table 4.14) revealed that at Stage one, the regression model was a significantly better fit for the data than the null model $F(1,78) = 10.15$, $p < 0.01$, and that the drawing size factor explained 10.4% of the variation in the attribution of fear to the line drawings. Introducing the curvature factor at Stage two generated a model which was also a significantly better fit for the data than the null model $F(2,77) = 8.27$, $p < 0.001$, and adding curvature increased the variation in attribution of fear explained by an additional 5.1%. Finally, the maximal model added the general complexity factor, and the addition of this predictor did significantly improve the fit of the model to the data $F(3,76) = 7.87$, $p < 0.001$, and explained an additional 5.2% of the variance in the perception of fear from the line drawings (compared to the Stage two model). These results indicate that participants were more likely to attribute the fearful option to a continuous drawing if it was smaller in size, if it was more jagged, and if the drawing was more complex.

This shows that the regression analyses revealed a combination of image properties that significantly influenced the perception of all emotions apart from happiness. Both neutral and sad emotions were perceived from larger drawings, and line drawings which were less complex. An increase in the size factor seemed to play a larger role in the perception of neutral emotion than the perception of sadness. The perception of both anger and fear were also influenced by the size and complexity of the continuous line drawings, as both emotion labels were more likely to be assigned to images which were smaller, and images which were more complex. However, these factors seemed to play a larger role in the perception of anger. An additional distinction between the perception of anger and fear can be found when examining results for the curvature factor. Curvature only had a significant impact on the recognition of fear from the continuous line drawings. Therefore, jaggedness of these drawings was uniquely associated with the perception of fear.

Aesthetics Results

To explore the contribution of each of the image properties (i.e., curvature, drawing size, general complexity factor) to aesthetic perceptions (liking, and beauty) for the continuous line drawings, two further hierarchical multiple regression analyses were conducted. As with the emotion models, image property factors will be added to the model in the order which reflects the frequency and consistency with which they are reported in the previous literature. Therefore, again, drawing size will be added to the model at Stage one, curvature at Stage two, and the general complexity factor will be added to create the maximal model.

It is possible that the relationship between aesthetic evaluation and complexity is non-linear; as previous research has suggested that an inverse U-shaped distribution will be observed. Therefore, we will also explore this possibility.

Liking

Table 4.15: Summary of hierarchical multiple regression analysis for the image properties predicting liking ratings for the continuous line drawings.

LIKING	Summary	Predictor	B	T	Std Error
Model 1	$F(1,78)= 1.03^{n.s.}$, $R^2_{Adj}= 0.0003$, RSE= 5.63	Intercept	44.11		
		Drawing Size	2.43	$1.01^{n.s.}$	2.39
Model 2	$F(2,77)= 2.57^{n.s.}$, $R^2_{Adj}= 0.038$, RSE= 5.52	Intercept	45.78		
		Drawing Size	1.15	$0.47^{n.s.}$	2.43
		Curvature	-9.85	-2.02^{*}	4.88
Model 3	$F(3,76)= 1.72^{n.s.}$, $R^2_{Adj}= 0.027$, RSE= 5.56	Intercept	45.60		
		Drawing Size	-0.16	$-0.03^{n.s.}$	5.23
		Curvature	-9.79	-1.99^{*}	4.91
		General Complexity Factor	2.19	$0.28^{n.s.}$	7.73
Model 4	$F(1,78)= 4.97^{*}$, $R^2_{Adj}= 0.048$	Intercept	46.39		
		Curvature	-10.45	-2.23^{*}	4.69

Model 1: Liking Ratings ~ Drawing Size

Model 2: Liking Ratings ~ Drawing Size + Curvature

Model 3: Liking Ratings ~ Drawing Size + Curvature + Complexity Factor

Model 4: Liking Ratings ~ Curvature

$p > 0.05^{n.s.}$, $p < 0.05^{*}$, $p < 0.01^{**}$, $p < 0.001^{***}$

The hierarchical multiple regression for liking ratings (Table 4.15) revealed that at Stage one, the regression model was not a significantly better fit for the data than the null model $F(1,78) = 1.03$, $p = 0.314$, and that the drawing size factor explained less than 0.1% of the variation in liking ratings provided for the line drawings. Introducing the curvature factor at Stage two generated a model which was a better fit for the data, but not significantly better than the null model $F(2,77) = 2.57$, $p = 0.083$, and adding curvature increased the variation in liking ratings explained by an additional 3.8%. Finally, the maximal model added the general complexity factor, and the addition of this predictor did not significantly improve the fit of the model to the data $F(3,76) = 1.72$, $p = 0.170$, and adding complexity decreased the variance in liking explained by around 1% (compared to the Stage two model). As curvature was the only significant predictor in any of the liking models, a simple linear regression was used to test whether this factor (in isolation from the other image properties) was significantly better for explaining the variance in liking ratings. Indeed, the model exploring only curvature was a significantly better fit for the data than the null model $F(1,78) = 4.97$, $p < 0.05$, and that curvature alone explained 4.8% of the variation in liking ratings provided for the continuous line drawings. Therefore, these results suggest that a decrease in jaggedness of an image produced higher liking ratings.

The results of the regression analyses suggest that the relationship between liking and complexity is indeed non-linear. To more closely examine the nature of this relationship, the average liking rating provided for each clip was plotted against its general complexity factor score, and a trendline was fitted on top of these points. This visualisation can be observed in Figure 4.8 below.

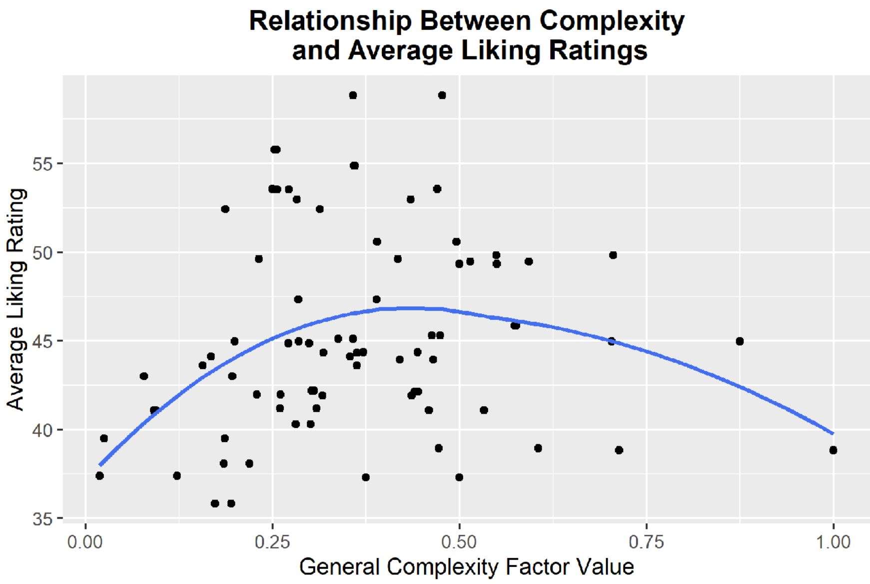


Figure 4.8: A visualisation of the relationship between our general complexity score, and the average liking rating provided for each of the continuous line drawings. The blue line represents the line of best fit through applying the 'loess' function to the scatterplot in RStudio.

It appears that this relationship follows an inverse-U shaped distribution, therefore, in this experiment we obtain additional support for the idea that there is a non-linear relationship between complexity and aesthetic evaluation.

Beauty

Table 4.16: Summary of hierarchical multiple regression analysis for the image properties predicting beauty ratings assigned to the continuous line drawings.

BEAUTY	Summary	Predictor	B	T	Std Error
Model 1	$F(1,78)= 0.78^{n.s.}$, $R^2_{Adj}= -0.003$, RSE= 5.75	Intercept	43.52		
		Drawing Size	2.14	$0.88^{n.s.}$	2.44
Model 2	$F(2,77)= 2.38^{n.s.}$, $R^2_{Adj}= 0.034$, RSE= 5.64	Intercept	45.20		
		Drawing Size	0.86	$0.35^{n.s.}$	2.48
		Curvature	-9.92	$-1.99^{n.s.}$	4.98
Model 3	$F(3,76)= 1.61^{n.s.}$, $R^2_{Adj}= 0.023$, RSE= 5.67	Intercept	44.96		
		Drawing Size	-0.84	$-0.16^{n.s.}$	5.34
		Curvature	-9.84	$-1.96^{n.s.}$	5.02
		General Complexity Factor	2.84	$0.36^{n.s.}$	7.89
Model 4	$F(1,78)= 4.97^{*}$, $R^2_{Adj}= 0.048$	Intercept			
		Curvature		-2.23^{*}	

Model 1: Beauty Ratings ~ Drawing Size

Model 2: Beauty Ratings ~ Drawing Size + Curvature

Model 3: Beauty Ratings ~ Drawing Size + Curvature + Complexity Factor

Model 4: Beauty Ratings ~ Curvature

$p > 0.05^{n.s.}$, $p < 0.05^{*}$, $p < 0.01^{**}$, $p < 0.001^{***}$

The hierarchical multiple regression for beauty ratings (Table 4.16) revealed that at Stage one, the regression model was not a significantly better fit for the data than the null model $F(1,78) = 0.77$, $p = 0.383$, and that the drawing size factor explained less than 0.1% of the variation in beauty ratings provided for the line drawings. Introducing the curvature factor at Stage two generated a model which was a better fit for the data, but not significantly better than the null model $F(2,77) = 2.38$, $p = 0.099$, and adding curvature increased the variation in liking ratings explained by an additional 3.4%. Finally, the maximal model added the general complexity factor, and the addition of this predictor did not significantly improve the fit of the model to the data $F(3,76) = 1.61$, $p = 0.194$, and adding complexity decreased the variance in liking explained by around 1% (compared to the Stage two model). As with for liking, curvature was the only significant predictor in any of the beauty models. Therefore, a simple linear regression was used to test whether curvature (in isolation from the other image properties) was also significantly better for explaining the variance in beauty ratings. Indeed, this model was a significantly better fit for the data than the null model $F(1,78) = 4.97$, $p < 0.05$, and that curvature alone explained 4.8% of the variation in beauty ratings provided for the continuous line drawings. Therefore, these results suggest that a decrease in jaggedness of an image also produced higher beauty ratings.

As with liking, the results of these regression analyses suggest that the relationship between beauty and complexity is also non-linear. To more closely examine the nature of this relationship, the average beauty rating provided for each clip was plotted against its general complexity factor score, and a trendline was fitted on top of these points. This visualisation can be observed in Figure 4.9 below.

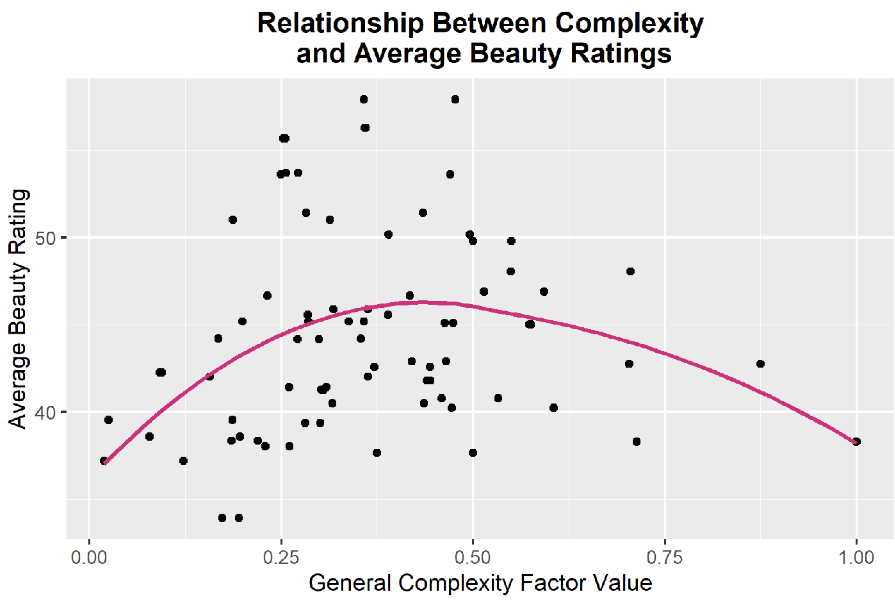


Figure 4.9: A visualisation of the relationship between our general complexity score, and the average beauty rating provided for each of the continuous line drawings. The pink line represents the line of best fit through applying the 'loess' function to the scatterplot in RStudio.

It appears that this relationship between beauty and complexity follows an almost identical inverse-U shaped distribution to that obtained for liking. Therefore, in this experiment we obtain support for the idea that there is a non-linear relationship between complexity and both kinds of aesthetic evaluation (liking and beauty).

EXPERIMENT 2: DISCUSSION

Size

In line with previous work exploring the role of size, we predicted that observers would provide beauty and liking ratings reflective of a general preference for larger over smaller stimuli. This general preference for larger stimuli has been observed across a number visual artworks, including paintings (Seidel & Prinz, 2018), basic shapes (Chen et al., 2016), and alphanumeric characters (Silvera et al., 2002). However, we did not find evidence for the "bigger is better" hypothesis in aesthetic judgements for our continuous line drawing stimuli; as hierarchical regression analyses revealed that the normed size factor did not significantly contribute to models explaining the average beauty or liking ratings provided.

However, it should be noted that in the work conducted by Chen and colleagues (2016), significant differences in preference were only observed in comparisons between small and medium, and small and large shapes, but not between medium and large shapes. Similarly, in the experiment exploring aesthetic appraisals of paintings differing in scale, the canvases used to display the paintings were drastically different in size. The continuous line drawings used in this experiment all had consistent image boundaries of 8 x 5 (inches), and the size factor only described the area of the shape contained within the perimeter of the drawings. Therefore, it is possible that the stimuli used in this experiment did not vary enough in size to produce differences in these judgements. It may be beneficial to introduce more subtle manipulations in the scaling of visual artworks to determine the level at which size becomes a key determinant of aesthetic evaluation.

An additional aim of this experiment was to explore whether the size of the line drawings would mediate perceptions about their emotional content. In the study conducted by Bromgard and colleagues, size played a concurrent role in both affective valence and aesthetic judgements for images of objects (Bromgard et al., 2013). However, to our knowledge, no experiments to date have directly explored the role of image size on emotion categorisation, and as a result no set predictions were made about the nature of this potential relationship. Through hierarchical regression analyses, we explored the possibility that differences in size may have influenced the emotion categories assigned to the continuous line drawings. We observed that size played a role in mediating the perception of neutrality, sadness, anger and fear, but not happiness, from the line drawings. More specifically, it was observed that larger shapes were assigned the neutral and sad labels, while smaller shapes were more commonly assigned the angry and fearful labels. Typically work in this area has explored the role of size on the perception of binary valence categories (i.e., positive or negative), but here we provide evidence to suggest that differences exist at the emotion-specific level; as sadness, anger, and fear are all situated within the negative valence category. Additionally, here we observed that size did not play a role in the perception of happiness, specifically, but it may play a role in the perception of other positive valence emotional states. Until future research works specifically examine the role of size and other visual properties in the perception of a broader spectrum of emotional states, rather than binary valence categories, the specifics of this relationship remain unclear.

Curvature

A number of studies, which vary in specific methodologies and participant samples have observed that participants show a general preference for curvature over jaggedness in visual stimuli (Silvia & Barona, 2009; Gómez-Puerto et al., 2018; Leder et al., 2011). Therefore, in line with these findings, we predicted that participants would provide more positive aesthetic value judgements for the continuous line drawings which featured smoother curves. The results of hierarchical regression analyses revealed that lower values on the curvature factor (an indicator of "jaggedness") were associated with higher beauty and liking ratings for the line drawings in this experiment. Therefore, we obtained full support for this experimental hypothesis. These results add to the previous literature by indicating that this curvature preference can be observed over a broader spectrum of visual stimuli (i.e., in visual stimuli that are simpler than paintings, and more complex than simple polygons).

It was also predicted that curvature would play a significant role in emotion classifications for the line drawings. Lu and colleagues explored this relationship by examining how different visual properties of images influenced the emotions aroused in participants, and found that images with smoother contours evoked both anger and sadness, while more jagged shapes induced fear in observers (Lu et al., 2012). We, therefore, predicted that line drawings with lower values on the curvature factor (i.e., smoother contours) would be associated with the perception of anger and sadness, while drawings with higher values (i.e., "jaggedness") would be associated with fear. No specific predictions were made about the relationship between curvature and happiness, as this was not explored in the study conducted by Lu and colleagues. We obtained partial support for this hypothesis, as the results of the hierarchical regression analyses revealed that the curvature factor was only a significant predictor of fearful classifications. It was, indeed, found that line drawings with higher curvature (i.e., jaggedness) values were more often assigned the fearful category. However, while curvature loaded negatively on the perception of sadness, and positively on the perception of anger, curvature did not significantly contribute to either of the models. It should be noted here that in the study conducted by Lu and colleagues, the image stimuli were extracted from the IAPS, which is comprised of a series of coloured images depicting everyday objects and scenes. It is possible then that curvature may only play a role in emotion judgements of anger and sadness for stimuli that are semantically, or socially, meaningful. To our knowledge, no experiments have explored the recognition of specific emotion categories from simpler visual stimuli (such as our continuous line drawings). With this in mind, it would be beneficial for future research to explore the role of curvature in a more diverse sample of visual stimuli, to allow us to draw more robust conclusions about the impact of curvature on emotion communication in visual aesthetics stimuli.

Complexity

Unlike with size and curvature, where there exists a single dominant theory for their impact on aesthetic evaluations, the role of image complexity is highly contested within the previous literature. Some researchers have suggested that, as a result of differences in perceptual fluency, observers will demonstrate a preference for less complex stimuli (Reber et al., 1998; Chen et al., 2011). However, it should be acknowledged that many of these experiments have explored perceptual fluency, and by extension complexity, through the lens of familiarity (e.g., mere exposure effect, using matching or non-matching primers), rather than through direct computational measures. Research works that have employed quantifiable measures of complexity suggest that the relationship between curvature and aesthetic evaluation follows an inverse-U shaped trajectory (Friedenberg & Liby, 2016). However, using computational methods to compute complexity presents its own challenges. As noted by Gauvrit and colleagues, there is perhaps too much fluidity in what constitutes complexity in visual aesthetic research, and this may be partially responsible for inconsistencies in reporting on the role of complexity on aesthetic value judgements assigned to aesthetic stimuli. As a result of these factors, in this experiment we did not make any specific predictions about the impact of complexity on the beauty and liking judgements for the continuous line drawings. However, we did aim to explore a broader range of computational measures than typically used in existing work, with the hope of generating a more robust way to quantify the complexity present within the artworks.

In our analyses we extracted six different complexity measures from the continuous line drawings. These features included a count of the black pixels making up the drawings (ink size), a count of the pixels comprising the perimeter of the shape (perimeter size), the file size of the line drawings after a consistent file compression was applied across the stimuli set (jpeg file size), a measure of pixel density (perimeter to ink ratio), and both 1D and 2D spectral complexity. When manually checking and comparing the output of these methods, we observed subtle differences in ordering of the stimuli set; particularly for line drawings with roughly median values on these factors, and less so for drawings with the minimum and maximum values. Through testing for multicollinearity between these features, we observed a number of significant linear relationships, and this prompted us to create a general complexity factor to use in the hierarchical regression analyses.

Through these regression analyses, we did not obtain support for a negative linear relationship between complexity and aesthetic evaluation, as the general complexity factor did not significantly contribute to either the beauty or liking model. However, when examining the distribution of aesthetic judgements and image complexity, an inverted-U shaped distribution emerged. Therefore, through results from this experiment, we provide additional evidence that the relationship between these factors is non-linear.

While our method for creating a general complexity factor (from an aggregated average of six specific measures) appears to be fairly successful, from a manual examination of the data and in replicating the results obtained by other researchers in this area (e.g., Fridenberg & Liby, 2016), we believe it would be beneficial in a future experiment to compare each of these specific measures, and the aggregated complexity factor, to subjective complexity ratings provided by observers. Indeed, a similar method for testing a variety of measures was employed by Chikhman and colleagues (Chikhman et al., 2012). We believe in doing so, we can begin to address the problem with defining "complexity" in visual aesthetics highlighted by Gauvrit and colleagues, and form a clearer picture of its role in deciding which visual artworks we like and find to be beautiful.

Finally, there appears to be a complete absence of any previous literature exploring the impact of complexity on emotion classifications for visual images. As a result of this, we did not make any specific predictions about how complexity would impact the emotion categories assigned to our continuous line drawings. However, the results of several hierarchical regression analyses revealed that images with lower complexity were perceived to be neutral and sad, whereas line drawings with greater complexity were more often associated with anger and fear. As with both size and curvature, complexity was not found to play a significant role in the perception of happiness from the continuous line drawings. These results provide some tentative first steps in elucidating the relationship between image complexity and the emotional expressivities we assign to visual art stimuli. However, as these results are the first to report on these relationships, we encourage further research in this area to examine the robustness of these findings.

OVERALL CONCLUSIONS

In the first experiment presented in this paper, we aimed to examine whether creating continuous line drawings from the trajectories of wrist and ankle movements of a dancer during emotionally expressive whole-body movements could provide a novel, alternative way to visualise human motion for social perception research. To explore this idea, we asked participants to complete a shortened measure of the Goldsmith's Dance Sophistication Index (Rose et al., 2020), and the Toronto Empathy Scale (Spreng et al., 2009), before taking part in an emotion recognition task. We proposed that if recognition rates were impoverished compared to those obtained for the PLD movement videos they were created from (Smith & Cross, 2022), but were still significantly greater than chance level, they may be useful for future emotion recognition experiments. However, we failed to obtain evidence in support of this notion. Additionally, we explored whether these continuous line drawings could function as a proxy for human dance motion, indirectly, by examining recognition rates obtained from individuals with higher levels of trait empathy and those with prior dance experience (two factors known to mediate the success of these judgements). Again, however, we failed to find support for these predictions. Therefore, we conclude that, in their current form, these trajectory line drawings do not provide a viable, or empirically robust, way to visualise movement of the human body in motion.

However, we also explored the possibility that these drawings could function as pieces of visual art in their own right, in isolation from their movement origins. In the second experiment, we explored the role of size, curvature, and complexity on the aesthetic value judgements and emotion labels assigned to these images; to determine whether these visual properties would produce similar social judgements to those obtained for other visual art stimuli. We, for the most part, observed evidence to support our experimental predictions; as both curvature and complexity facilitated similar beauty and liking judgements to those obtained for paintings and simple shapes. We also provided new evidence for the role of these factors in mediating emotion judgements, and explored a novel way to quantify "complexity" in visual aesthetics research.

Therefore, in summary, while this stimulus creation method was not suitable for its intended purpose (as an alternative to existing human movement stimuli), we believe the methodology presented in this manuscript provides a new way to generate visual artworks which are suitable for use in visual empirical aesthetics research.

CHAPTER 5

“A LITTLE MORE CONVERSATION”: A QUALITATIVE ASSESSMENT OF DECISION-MAKING BY DANCERS AND NON-DANCERS DURING A WHOLE-BODY MOVEMENT EMOTION RECOGNITION TASK

“A Little More Conversation”: A Qualitative Assessment of Decision-Making by Dancers and Non-Dancers During a Whole-Body Movement Emotion Recognition Task

Rebecca A. Smith & Emily S. Cross

ABSTRACT

The dyadic nature of communicating and interpreting social cues from others in our environment plays a key role in forming and maintaining relationships with others, which in turn has an impact on our mental wellbeing, sense of identity and belonging, and ability to successfully navigate our social world. A number of researchers have explored how we perceive and interpret the emotional expressivity present in human movement (through emotion recognition tasks), and the factors that can facilitate or impair these abilities (e.g., neurodevelopmental disorders like Autism Spectrum Disorder, individual differences in trait empathy, and previous movement experience). Forced-choice tasks have dominated the bulk of the previous emotion research, and exploration of the role of individual differences, such as previous dance experience and kinaesthetic empathy, is also typically examined through a quantitative lens. However, an over-reliance on these kinds of tasks limits the scope of information we can obtain about whole-body social signalling. The primary aim of this study was to begin to overcome some of these limitations through exploration of how observers assign judgements of happiness, sadness, anger and fear (in addition to an absence of emotional expressivity – i.e., neutrality) to whole-body movement sequences, through open-ended questions asked during semi-structured interviews. To this end, 10 dancers and 10 non-dancers observed emotionally expressive whole-body dance movement sequences from the McNorm Library (Smith & Cross, 2022), assigned emotion labels to each of the videos, and were asked to explain the reasons for their decisions. Through these interviews, we observed that a number of factors can influence decision-making processes in emotion recognition tasks. These factors included the imagination of contextual cues (i.e., the visualisation of environmental triggers and interaction partners), movement speed, aesthetic considerations (e.g., beauty, liking, interest), kinaesthetic empathy, and learned associations from previous experiences. This hybrid approach, combining standard quantitative measures of emotion perception with deeper qualitative insights, provides novel perspectives on the individual differences that can sculpt decision-making in emotion recognition tasks, and in so doing has also provided several avenues for future research in this area.

INTRODUCTION

Throughout history, people have been deeply interested in understanding emotion. The nature of emotional experiences has been of long-standing discussion within philosophy (Knuuttila, 2014; Leigton, 1982; Konstan, 2006), and debate about how to empirically define emotions can be observed through the chronicles of prominent psychologists, neuroscientists, and physiologists (Scheff, 2015; Hofmann & Doan, 2018; Scherer, 2005; Celeghin et al., 2017; Adolphs & Anderson, 2018; Kleinginna & Kleinginna, 1981). Within this discourse, a number of emotion theories have been proposed. Darwin's evolutionary perspective highlights the adaptive and functional purpose of emotions (Darwin, 1872/1965; Ekman, 2009). Within this framework, he proposes that we experience certain negative emotions, like fear, to evade danger or harm, and we experience emotions like love and affection for the purpose of finding and retaining a mate to procreate with (Hess & Thibault, 2009). Beyond such theories about the function of emotional experiences, a number of theoretical accounts have been proposed to explain how emotions are produced within an individual.

The James-Lange (James, 1884; Lange, 1885), Cannon-Bard (Cannon, 1927; Bard, 1928), and Schachter and Singer (Schachter & Singer, 1962) theories of emotion diverge slightly in the specific mechanisms by which emotional experiences occur within the individual, but broadly agree that this process involves physiological responses to and cognitive processing of an emotionally-salient stimulus in the environment. Through successive refinements of these theories, we have made substantial advances in our understanding of how emotions occur at an individual level. However, these theories do not account for how emotional experiences occur in our social world, more specifically, they do not provide an overview of how these affective states can be successfully communicated and interpreted by both parties in the interaction dyad. In fact, to our knowledge, only in facial expression research has any substantial conceptual framework for the recognition of emotional expressivity been proposed.

Although the Universality Hypothesis for facial expressions (Elfenbein & Ambady, 2002; Ekman & Friesen, 1978; Andrew, 1963; Susskind et al., 2008) has received criticism (Jack et al., 2016), through research in this area we are starting to develop an understanding of the communication potential of the face in social interactions. However, we can obtain a wealth of socially-relevant information from other non-verbal cues, such as movement of the human body (Dekeyser et al., 2002; Ma et al., 2006; Vanrie & Verfaillie, 2004; Pollick et al., 2005; Mitchell & Curry, 2016), but to our knowledge there does not currently exist such a framework for the recognition of emotional expressivity through such bodily cues.

Despite the lack of a unified theory, a wealth of research has explored specific aspects of social perception of the human body in motion through a variety of quantitative methods. For example, the success of emotional communication is typically measured through forced-choice emotion recognition experiments (Johnson et al., 1968; Demenescu et al., 2014; Atkinson et al., 2004; Ross et al., 2012). Atkinson and colleagues are among a number of researchers who have employed such methods to examine the recognition of emotion from whole body movements. In their 2004 paper, Atkinson and colleagues developed a new set of dynamic and static whole-body stimuli depicting five emotions (happiness, sadness, anger, fear and disgust) with various degrees of exaggeration, in the form of full-light and point-light displays. Participants were asked to identify the target emotion for each of these stimuli from the five forced-choice emotion options. Results from this study showed that recognition rates as high as 92% could be obtained from such stimuli (Atkinson et al., 2004). Similar methods were employed by Ross and colleagues, who explored developmental changes in the recognition of happiness, sadness, anger and fear from human movement. In this experiment, children aged 4-17 and a group of adults observed the movement sequences and were asked to make forced-choice emotion judgements for whole-body movements depicted in the form of full-light and point-light displays. They observed that children recognised the intended emotion with less accuracy than the adults for both movement presentation styles, but these recognition rates could still be as high as 90% in teenage participants (Ross et al., 2012).

These results suggest that observers can recognise target emotions from whole-body movements with high levels of accuracy. However, it has been argued that such simple "button-pressing" tasks are critically limited for enabling us to build a rich understanding of emotion perception (Russell, 1993; Tracy & Robins, 2004; see Limbrecht-Ecklund et al., 2013 for an evaluation). For example, it has been argued that the presence of several options may lead observers to provide relative judgements, whereby observer may be selecting an option that seems most likely within the constraints of the experiment rather than providing a response that reflects their genuine beliefs (Russell, 1993). It has also been suggested that the nature of forced-choice tasks can lead to the presence of ceiling effects, which may explain why recognition rates of more than 90% (and sometimes as high as 100%; Gross, Crane & Fredrickson, 2010) have been reported in the previous literature (Atkinson et al., 2004; Ross et al., 2012; Montepare et al., 1999; Bernhardt & Robinson, 2007; Dael et al., 2012). This may be of greatest concern in studies examining the recognition of happiness. Happiness or Joy is frequently the only positively-valenced option present within such forced-choice formats (Tracy & Robins, 2004), and this could explain why stimuli depicting happiness often receive the highest levels of recognition in emotion recognition tasks (Atkinson et al., 2004; Pasch & Poppe, 2007; Melzer et al., 2019; Dahl & Friberg, 2007).

We argue that perhaps a larger concern with such simple "button-pressing" tasks, however, is that they necessarily reduce the myriad complexities of emotion communication to a single percentage value of recognition. Forced-choice tasks have dominated the bulk of the previous emotion research, setting out a standard operating procedure for affective science, but an over-reliance on these kinds of tasks limit the scope of information we can obtain about whole-body social signalling. For example, while such experiments have allowed us to develop an understanding of the accuracy with which observers can assign emotions to socially-salient stimuli, they do not provide any indication of *how* observers make these decisions. Indeed, the reasons for why observers assign specific kinds of emotional expressivity to whole-body human movement represents a significant gap in our understanding of affective cognition. Therefore, for researchers aiming to wade into these uncharted waters, it may also be useful to turn to more open empirical methodologies. Qualitative research methods can be invaluable for exploring novel, complex, or relatively unaddressed areas of research (Clarke & Jack, 1998), as the open-format of such experimental procedures can be used to identify new perspectives on existing findings and to generate new avenues for future research.

Therefore, the primary aim of this study was to explore how observers assign judgements of happiness, sadness, anger and fear (in addition to an absence of emotional expressivity – i.e., neutrality) to whole-body movement sequences, through a combination of more standard forced-choice questionnaires and open-ended questions asked during semi-structured interviews. The choice to explore only these four kinds of emotional expressivity, excluding disgust or surprise, is rooted in both the uncertainty surrounding their universality in recognition from facial expressions (Jack et al., 2016), and the lack of research exploring observer recognition of these kinds of expressivity from human motion, more specifically. In this experiment, participants were asked to observe a series of emotionally expressive dance movement sequences from the McNorm Dance Movement Library (Smith & Cross, 2022), and provide evaluations of the movements in a forced-choice emotion recognition task. During this task, participants were also asked to provide justification for their choices and detail about anything present within the movements that assisted them with their decisions. Our aim was that combining in-depth examination of *how* observers assign emotional expressivity to the body in motion with objective measures should provide new insights into the mechanisms of whole-body emotion communication, while also generating new research questions which can be submitted to rigorous quantitative assessment in future work.

We are not the first to address the question of *how* emotion judgements are made. A number of researchers have identified individual differences that can mediate success in emotion recognition tasks, as a way to isolate the mechanisms underlying these judgements. The role of previous dance training is one such factor which has been found to sculpt the way individuals perceive human motion, and can result in improved ability to read social cues from the movement of others. Numerous studies conducted by Christensen and colleagues have observed that dancers are more accurate than non-dancers in identifying emotional expressivity from human movement stimuli (Christensen et al., 2016; Christensen et al., 2019; Christensen et al., 2021). However, the reason for these improvements following extensive dance training are unclear. It is common for these results to be explained through the concepts of kinaesthetic empathy, simulation or embodiment. The idea being that dancers are better able to put themselves in another's shoes, prompting them to imagine themselves performing those movements and ultimately helping them to identify how doing those movements would make them feel. Neuroscientific evidence provides some support for this claim. For example, in a study conducted by

Cross and colleagues, expert dancers were found to showed heightened activity in areas associated with the "Action Observation Network" (AON; including the ventral premotor, supplementary and primary motor cortices) when observing the movement of others, and imagining themselves performing those movements. Additionally, they found that self-reported physical ability mediated activity within the inferior parietal lobule (IPL) and ventral premotor cortex, suggesting that motor simulation may occur more strongly in response to movements that observer feel they can successfully execute themselves (Cross et al., 2006). In a separate experiment, the IPL was found to play a causal role in processing affective cues from the human body in motion (Engelen et al., 2015). Considered together, these results show that motor expertise prompts greater activation in the IPL (which is involved in both movement simulation and processing of social cues from the dynamic human body) and suggests that movement expertise, motor simulation and affective recognition are related.

However, the conclusions drawn about how dance experience and empathy may interact to influence emotion recognition from human movement are rooted in a quantitative approach, and are often inferred from quantitative data; obtained from studies exploring similar, but empirically distinct, research questions. Few experiments have explored these elements of social processing in tandem within the context of an emotion recognition task specifically and, to our knowledge, no research has examined these factors using qualitative research methods. Therefore, an additional aim of this study was to explore how dancers draw upon previous physical experience to assist them in this emotion recognition task, and to probe the extent to which kinaesthetic imagery and embodiment sculpt their decision-making process, by asking them directly. Additionally, although non-dancers do not have the same rich history of motor experience to draw upon, it is possible that they too make use of kinaesthetic imagery to assist them with their emotion judgements. Therefore, questions about the use of kinaesthetic empathy will be posed to all participants in the interview sessions, regardless of their dance background.

Furthermore, the concept of 'previous experience' in movement perception experiments typically refers to motor experience, specifically. It is unclear whether other types of previous experiences or personal associations assist participants with making these emotion judgements. While dance is a key component of the lives of professional dancers, they do not exist within a vacuum, and movement experts will have a wealth of other previous life experiences which they may draw upon to assist them with social decision-making tasks. Therefore, it is possible that non-dance related experiences may be a tool used by both dancers and non-dancers in such tasks. However, the question of whether, and the extent to which, these experiences contribute to affective decision-making remains so far unexplored. Therefore, bridging this gap in the previous literature was the final aim of this project. To address this, we will explore the influence of various kinds of previous experiences in assigning emotional expressivity to whole-body human movement.

It should be noted here that this study is exploratory in nature, and for that reason we do not evaluate specific, pre-registered hypotheses. Instead, we hope to show the value of employing more open-format, qualitative research methods and hope that this work can be used to generate future hypotheses, and research questions, which can be submitted to further rigorous empirical evaluation in future work.

METHODS

PARTICIPANTS

Two groups of participants were recruited to take part in this experiment: experienced dancers, and non-dancers. Ten experienced dancers with three or more years of formal training at a dance institution, and/or years of experience dancing with a professional company participated in this experiment. Ten non-dancers with no previous physical dance experience, including hobby dance classes were recruited to take part in this experiment. However, sound quality of the audio recording from one non-dancer participant was too poor to allow for successful transcription, therefore this data was discarded. One additional non-dancer's data was excluded from the analysis, as technical issues resulted in an incomplete recording. As a result of these exclusions, to allow for equal groups, a further two non-dancer participants were recruited to take part in this experiment. Therefore, data from 20 participants (10 experienced dancers, 10 non-dancers) are included in this manuscript. A more detailed breakdown of each participant's demographics information can be found in Table 5.1.

Table 5.1: A summary of the age and gender of each participant, and their total years of dance experience, and the main dance style they had the most experience with. It should be noted that one participants in the non-dancer group provided a numeric value for years of experience that was not zero (i.e., GM161853), as they mentioned dancing for social reasons and felt that the response of "2" most accurately described their informal dance experience. Additionally, a number of non-dancers provided a dance style they were most familiar with (e.g., Waltz, Ceilidh, Disco, Formal Dance) based on their lifetime exposure to dance in informal settings.

Subject ID	Group	Age	Gender	Years of Dance Experience	Main Dance Style
AM031508	Non-Dancer	27	Male	0	N/A
CB241104	Dancer	26	Male	8	Jazz
CC051508	Dancer	20	Female	18	Ballet
CR201309	Dancer	24	Female	15	Contemporary
CT041605	Dancer	24	Male	7	Breakdancing
EH301811	Non-Dancer	34	Female	0	N/A
ET270909	Dancer	29	Male	15	Ballet
GM161853	Non-Dancer	56	Male	2	Waltz
GW011752	Non-Dancer	27	Male	0	Ceilidh
KM161203	Dancer	21	Female	15	Ballet
KR281508	Dancer	19	Female	15	Jazz

LJ181312	Non-Dancer	24	Female	0	N/A
LM1111	Dancer	31	Female	12	Ballet
PS291837	Non-Dancer	44	Male	0	N/A
RT231409	Non-Dancer	65	Male	0	Disco
SD081104	Non-Dancer	20	Female	0	N/A
SK161007	Dancer	28	Female	8	Cabaret
SM031603	Non-Dancer	25	Female	0	Formal Dance
SMW021605	Dancer	22	Male	16	Ballet
TC081402	Non-Dancer	23	Gender Variant/ Non-Conforming	0	N/A

The experiment was advertised via word of mouth, through correspondence with UK-based dance and performing art companies (e.g., Scottish Dance Theatre, National Theatre of Scotland), and on a variety of social media channels (including Twitter, Facebook, and Instagram).

Participant samples included in qualitative research are typically smaller than those used in quantitative research, to facilitate the kind of in-depth analysis that is fundamental to this kind of empirical work (Vasileiou et al., 2018; Sandelowski, 1996). Small scale quantitative works often fall victim to issues such as low statistical power, inflated false discovery rates, and a low rate of reproducibility (Button et al., 2013; Colquhoun, 2014; Lakens & Albers, 2017). It should be acknowledged here that quantitative data analysis was not the primary focus of this manuscript, and was only included for the sake of completeness. Using results from the original McNorm Movement Library validation experiments (i.e., where testing whether overall recognition rates were greater than chance level produced an effect size of 1.54 in Experiment 1, and 2.13 in Experiment 2), we calculated the smallest sample size required to achieve adequate study power in this emotion recognition task. Using the smaller effect size of 1.54, we found that a minimum of 12 participants were required to make up the total sample size, and the suggested group size was 6 participants. Therefore, it is likely from a quantitative perspective that a sample size of 20 may be adequately powered. However, including only 10 dancers and 10 non-dancers to take part in this experiment may not have been sufficient to capture more subtle group differences in the recognition of specific expressivity, and as such, quantitative conclusions drawn from this work should be treated with caution.

STIMULI

The stimuli used in this experiment were selected from the McNorm Movement Library. These movement clips have been validated for emotion recognition in a series of prior experiments by the authors (Smith & Cross, 2022).

Clips from the full McNorm Movement Library (N=10; 2 Neutral, 2 Happy, 2 Sad, 2 Angry, 2 Fearful) used in this aesthetic judgment experiment were selected as they were found to be most representative of their intended emotion category. In the creation of the McNorm Movement Library, movement sequences were devised by a professional dancer who was instructed to exclude any movements which are tied to a particular emotional state (e.g. shaking fist for anger, arms thrown in the air for joy), or movements that are associated with specific emotionally evocative narratives. These pieces of choreography were then performed five times, each time with the aim to communicate a different emotion (happy, sad, angry, fearful, and neutral/non-expressive as a control) to an imagined spectator. Stimuli considered to be most representative of these emotion categories were decided based on recognition rates for the emotion intended to be communicated by the movement performer (i.e., clips which were intended to be happy by the movement performer which received the highest proportion of 'happy' responses in the emotion validation experiments). A more detailed breakdown of the emotion recognition rates for each clip used in this experiment can be found in Table 5.2.

Table 5.2: Summary of the whole-body dance movement stimuli from the McNorm Library used in this experiment. The table provides information about the name of the clip, the emotion category assigned to the clip, and the average rate of recognition for the intended emotion for the clip across the two validation experiments in Smith & Cross (2022).

Clip Name	Emotion Category	Average Emotion Recognition Rate (%)
plie_N	Neutral	60.13
tendu_N	Neutral	61.19
char_H	Happy	53.30
pirouette_H	Happy	50.76
fondu_S	Sad	67.96
plie_S	Sad	64.08
bitter_A	Angry	43.93
char_A	Angry	65.12
ever_con_F	Fearful	43.64
ice_F	Fearful	43.47

The 10 movement sequences outlined in the previous section were uploaded to Vimeo and were displayed through formr using embed codes. In Vimeo, the embed codes were manipulated to remove the standard online video handles (video title, uploader information, suggested videos, and Vimeo branding). The video controllers were also removed, and the video was set to loop automatically until participants had submitted their responses for the item.

PROCEDURE

Participants who had agreed to take part in this experiment after reading the information sheet (available on the OSF: https://osf.io/wbm6s/), received an invitation to join a password-protected Zoom meeting, where the interviews would be conducted. After an opportunity to ask questions, participants were provided with a formr link to begin the experiment. After reading the task information, consenting to participate, and consenting to the recording of the Zoom meeting for the purpose of transcription (in accordance with BPS guidelines), participants completed several questions from the Goldsmith's Dance Sophistication Index (Gold-DSI; Rose et al., 2020). Selected questions from the early Gold-DSI included demographics measures which allowed participants to provide detail about their previous dance experience, covering both formal and informal experience, and both visual and physical experience. Participants were also asked to select which level of dance experience was most applicable to them from five options (non-dancer, beginner, intermediate, advanced, or professional), and to quantify the number of years of dance experience they had.

After completing the Gold-DSI measures, recording of the session began and participants started the emotion recognition task. At this point, participants were also invited to share their screen to allow the researcher to guide them through the experiment. During the emotion recognition task, participants observed the selected clips from the McNorm Movement Library (10 clips presented in a randomised order; 2 neutral, 2 happy, 2 sad, 2 angry, and 2 fearful), and were asked to select which emotion they believed the dancer was conveying through their movements from five forced-choice options; neutral, happy, sad, angry or fearful. After making their decision, all participants were asked the following two questions;

1. *"What made this clip seem X (i.e., the emotion they selected) to you?"*
2. *"Is there a different emotion, outside of these options, that you think is a better description for what the dancer is communicating?"*

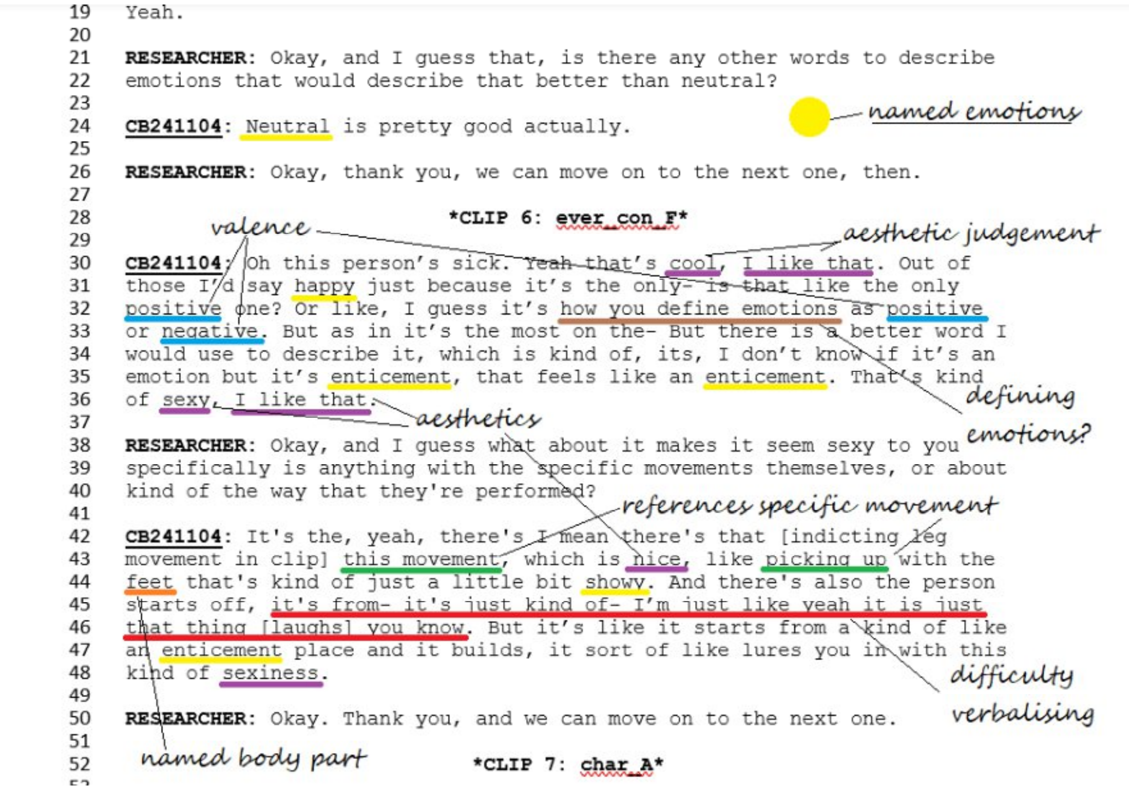
This process was repeated for each of the 10 McNorm Movement Library clips. Additional unscripted questions were sometimes asked during the task to allow participants to elaborate on how they made their decisions. After completing the emotion labelling task, participants were asked to stop sharing their screen, and were asked a series of final questions about their decision-making processes. These questions centred largely around the topics of emotion regulation, visual attention, kinaesthetic empathy, the use of imagery, and the role of specific movements and performative quality on decision-making in an emotion-labelling task. Further questions were asked when required to allow the participant to provide clarification or to expand upon their responses. A full list of these questions is available on the OSF page for this project (https://osf.io/wbm6s/). After completing the experiment, the session recording was terminated. Participants were then informed of the purpose of the experiment, given a final opportunity to ask any questions, and thanked for their time.

Qualitative Data Processing

The video conferencing software (Zoom) produced automated transcripts which were manually checked and edited by the experimenter to correct for errors. The text files were then merged to create a document which contained all of the interview data. In conducting the analysis of this interview data, we followed the guidance proposed by Braun and Clarke (2002) for thematic analysis.

1. *Initial Coding*

This file was then printed and coded by hand by the lead investigator. The first stage of coding the data involved identifying any words or phrases that related to any of the research questions and those which appeared to be relevant to decision-making more broadly. Therefore, it should be noted here that this process involved both deductive and inductive methods of coding. A worked example of this initial coding process can be observed in Figure 5.1. This process was carried out across all 210 pages of the collated transcripts, and a number of common codes emerged.


Yeah.

RESEARCHER: Okay, and I guess that, is there any other words to describe emotions that would describe that better than neutral?

CB241104: Neutral is pretty good actually.

RESEARCHER: Okay, thank you, we can move on to the next one, then.

CLIP 6: ever_con_F

CB241104: Oh this person's sick. Yeah that's cool, I like that. Out of those I'd say happy just because it's the only- is that like the only positive one? Or like, I guess it's how you define emotions as positive or negative. But as in it's the most on the- But there is a better word I would use to describe it, which is kind of, its, I don't know if it's an emotion but it's enticement, that feels like an enticement. That's kind of sexy, I like that.

RESEARCHER: Okay, and I guess what about it makes it seem sexy to you specifically is anything with the specific movements themselves, or about kind of the way that they're performed?

CB241104: It's the, yeah, there's I mean there's that [indicting leg movement in clip] this movement, which is nice, like picking up with the feet that's kind of just a little bit showy. And there's also the person starts off, it's from- it's just kind of of- I'm just like yeah it is just that thing [laughs] you know. But it's like it starts from a kind of like an enticement place and it builds, it sort of like lures you in with this kind of sexiness.

RESEARCHER: Okay. Thank you, and we can move on to the next one.

CLIP 7: char_A

Figure 5.1: An example of the initial coding process carried out on an excerpt of the transcription data. During the first pass of the transcription data, words and phrases relevant to the research questions were underlined with code-specific colours and annotated. In this visualisation, words underlined in yellow indicated any references made to a specific emotion, purple was used to highlight any references to aesthetic parameters, valence was highlighted in blue, and references to specific body parts were

underlined in orange. Phrases highlighted with brown or red were associated with the concepts of "*defining emotions*" and "*difficulty verbalising*", respectively.

2. *Identifying Initial Codes*

Over time, a number of common codes emerged, and segments of the transcripts (words and phrases) were collated under each of these headings. For example, Figure 5.2 shows how words and phrases were collated under the code "knowledge of performative dance". A number of such codes were identified, and participants' comments were grouped under the relevant headings.

RESEARCHER: Okay, what about this movement is it that makes it seem happy to you?

CC051508: It's like, very like big and it reminds me of, oh what's the one in Nutcracker with like the two guys and they're like tumbling and stuff? It reminds me of that. It's like very big and like energetic and stuff.

LM1111: Um, I think it would be- I think what it was about, now I'm like thinking about general like classical dance, which is quite specific that there aren't that many fearful. I can't think of ballets where that happens a lot, and when it happens, the I just I saw Giselle the other day, basically, the guy just keeps running away right [laughs] and keeps jumping and running away right. It's like, I feel like fearful it's like in its physicality it would have to be quite… Like someone who is a bit unsure what they're doing in a way which I think, now that I think about it, doesn't happen a lot in classical ballet because it feels like very clear where you going if I'm making sense, so I feel like the fearful sort of the slightly [leans back] what would have been sort of slightly a movement not necessarily walking backwards [arm movement], but something that's not as front [punches arm forward] from the focus and not as clear in direction, maybe. Maybe that's it. I'm not sure.

RT231409: Probably no- in some respects aye, you know, you're looking at it, you're thinking that's reminiscent of Swan Lake or, you know, a movement in Swan Lake or something like that. But other than that I can't say that I was, you know, constructing any great narrative.

Figure 5.2: A visualisation of words and phrases assigned to the code "knowledge of performative dance". These excerpts were extracted from the transcripts of three participants, all of whom provided separate references to specific performative ballets. Over time, any such references to specific ballet stories were added under this code.

3. *Identifying Themes*

After identifying the initial codes, thematic analysis was used to group these codes into broader "themes". Several of these themes were created through deductive coding (Braun & Clarke, 2002; Joffe, 2012), as they were identified to be of interest before beginning the experiment. These centred around the concepts of kinaesthetic empathy, kinematics, and previous dance experience. For example, the code "*knowledge of performative dance*" (Figure 5.2) was identified to be a subset of the theme "**previous experience**"; along with the codes "*learned associations in dance*", "*learned associations from real-world experiences*", and "*learned associations from popular culture*". An additional code related to visual attention was also identified to be of interest prior to the collection of this data, but results pertaining to visual attention were ultimately outwith the scope of this chapter.

A number of additional codes and themes emerged through inductive coding (Braun & Clarke 2002; Joffe, 2012). These themes were not identified to be of interest prior to the beginning of data collection, but appeared frequently within the transcripts and seemed to be relevant to decision-making, and perception of dance more broadly. These centred around *difficulties in defining emotions*, *difficulties with verbalising thought processes*, *aesthetic considerations*, and the *influence of other imagined social cues*. A full list of the themes and codes identified in these interviews can be found in a collated document on the OSF (https://osf.io/wbm6s/).

REFLEXIVITY

In this section, we acknowledge the role of the researcher, and their potential assumptions or biases, within the research process. All qualitative research is subject to the influence of subjectivity, therefore this is not intended to be a criticism of the researcher or the work discussed in this manuscript. However, we believe that the role of the experimenter's previous knowledge and beliefs should be addressed before presenting the results of the experimental process. The primary experimenter (RAS) has a background in Psychology, with a specific interest in Social Psychology and Affective Science. As a result, the research process may have been influenced by their previous knowledge and experiences, and it is to be expected that findings will be interpreted with a strong emphasis on concepts identified in the previous literature. In addition, the primary experimenter has around 20 years of informal (hobby) ballet training, therefore it is likely that an understanding of dance-specific concepts and dance-related vernacular will be greater than in an experimenter without such training. This experience also creates a shared point of reference between the experimenter and the dancer participants which may have unintentionally influenced certain aspects of the discussion.

QUANTITATIVE RESULTS

Again, it is worth reiterating here that quantitative data analysis was only included for the sake of completeness, and was not the primary focus of this manuscript. Therefore, we should be careful about drawing conclusions from this data, given the small sample size and the potential lack of statistical power for more complex group differences.

Emotion Recognition Results

The average percentage recognition of the intended emotion, across all participants and all emotions was 52.00% (SD = 10.05). With a baseline of 20% established as chance level (from a 1 in 5 chance of selecting the intended emotion), after normality testing ($p > 0.05$), a one-sample t-test revealed that overall participants identified the intended emotions at greater than chance level: $t(19) = 14.24$, $p < 0.001$, $d = 3.18$.

Examining each emotion individually, all emotions were recognised at significantly greater than chance level. See Table 5.3 for a summary of the recognition rates for clips from each emotion category, and a summary of the inferential tests performed on the data. Further, the mean recognition rates for each expression category, and the distribution of responses can be observed in Figure 5.3.

Table 5.3: A summary of recognition rates for each emotion category (means and standard deviations), and the results of inferential tests performed on the data to determine whether they were recognised at greater than chance level.

Emotion	Average Recognition Rate (%)	Greater than Chance Recognition?	Test	Results
Neutral	60.00 (± 34.79)	Yes	One-Sample Wilcoxon Signed Rank Test	Z = 204, $p < 0.001$, $d = 1.15$
Happy	75.00 (± 30.35)	Yes	One-Sample Wilcoxon Signed Rank Test	Z = 209, $p < 0.001$, $d = 1.81$
Sad	62.50 (± 39.32)	Yes	One-Sample Wilcoxon Signed Rank Test	Z = 200, $p < 0.001$, $d = 1.08$

Angry	32.50 (± 29.36)	Yes	One-Sample Wilcoxon Signed Rank Test	Z = 174, p< 0.01, *d* = 0.43
Fearful	30.00 (± 34.03)	Yes	One-Sample Wilcoxon Signed Rank Test	Z = 155, p< 0.05, *d* = 0.29

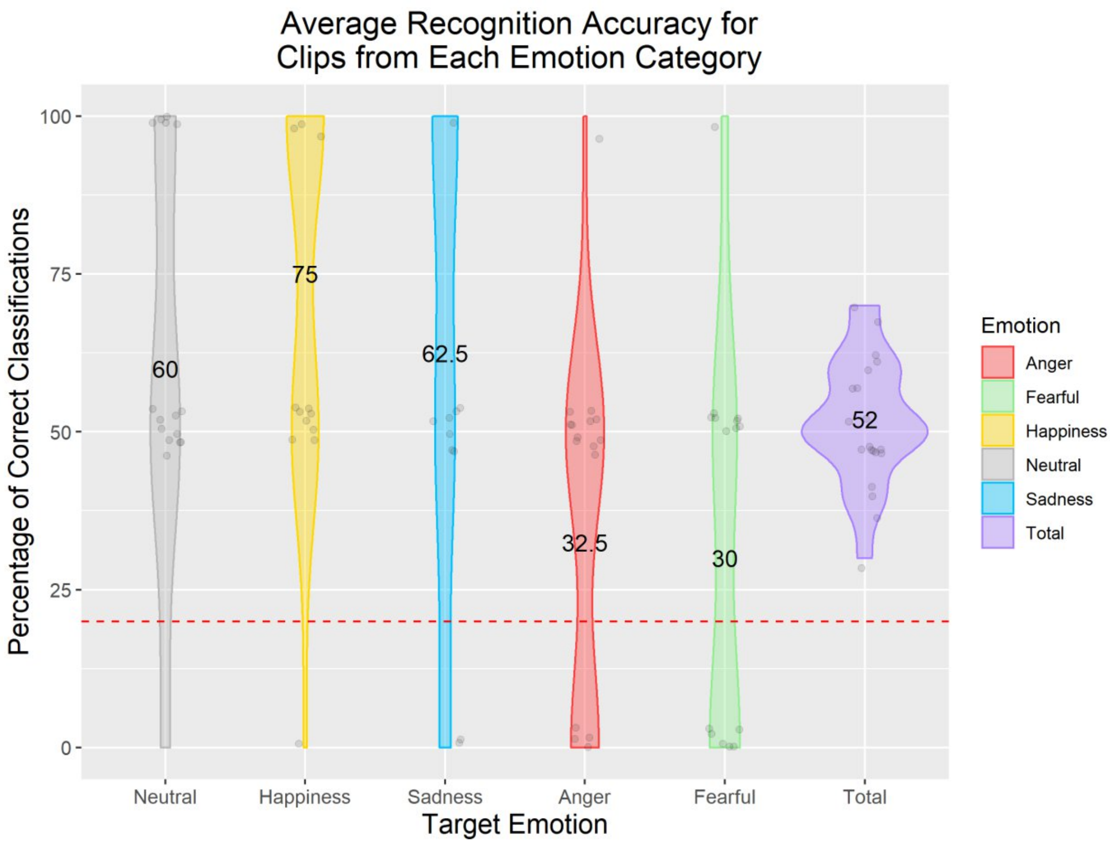


Figure 5.3: Violin plots depicting the distribution of recognition rates for each emotion category. Values presented in the centre of each of the violins represent the mean recognition rate, and the red dotted line indicates chance level of recognition. All emotions were recognised at significantly greater than chance level.

Although all emotions were recognised at greater than chance level, a number of frequent misclassifications still occurred. Movements intended to communicate sadness were frequently perceived to communicate neutral expression (25%). Movements intended to express anger were very frequently mistaken for happiness (47.5%), even more often than they were labelled correctly (32.5%). Fearful movements were frequently confused with all other emotion categories; neutral expression (15%), happiness (17.5%), sadness (17.5%), and anger (20%), but not more often than it was correctly labelled (30%). A more detailed breakdown of the correct classification and misclassification rates can be found in Table 5.4.

Table 5.4: Correct classification and misclassification rates for clips from each emotion category.

		Perceived Emotion				
		Neutral	Happy	Sad	Angry	Fearful
Intended Emotion	Neutral	**60.00%**	17.50%	7.50%	7.50%	7.50%
	Happy	5.00%	**75.00%**	10.00%	7.50%	2.50%
	Sad	25.00%	10.00%	**62.50%**	0.00%	2.50%
	Angry	5.00%	47.50%	7.50%	**32.50%**	7.50%
	Fearful	15.00%	17.50%	17.50%	20.00%	**30.00%**

These results align with those obtained in the original validation experiments conducted on the McNorm Library (Smith & Cross, 2022), as clips from all emotion categories were recognised at greater than chance level, and the lowest recognition rates were obtained for the fearful clips.

Impact of Dance Experience on Recognition

Participants were recruited to form two different groups, each containing observers with different levels of previous dance experience. One group comprised participants who had received three or more years of formal dance training, and the other group comprised participants who had no previous dance training (including informal or hobby dance lessons). A series of independent samples t-tests were conducted to explore whether differences in previous dance experience had an impact on performance in the emotion recognition task in general, and whether any differences would emerge at the emotion-specific level.

While experienced dancers ($M = 53.00\%$, $SD = 9.49$) performed slightly better in the emotion recognition task, overall, than the non-dancers ($M = 51.82\%$, $SD = 10.79$), this difference was not statistically significant: $t(19) = 0.267$, $p = 0.792$. At the emotion-specific level, the only significant difference in recognition between experienced dancers and non-dancers occurred for recognition of the sad movement clips. In this experiment, non-dancers ($M = 86.36\%$, $SD = 23.35$) recognised sadness from the McNorm Library clips with greater accuracy than the experienced dancers ($M = 40.00\%$, $SD = 39.44$): $t(14.3) = -3.237$, $p < 0.01$. No other group differences emerged (all $p > 0.05$). A detailed breakdown of these group differences in recognition accuracy can be found in Table 5.5.

Table 5.5: A summary of recognition rates for each emotion category (means and standard deviations) for each participant group (i.e., dancers and non-dancers), and the results of inferential tests performed on the data to determine whether group differences were significant.

	Average Recognition Rate (%)		Inferential Test Results
	Dancers (N= 10)	Non-Dancers (N= 11)	
Overall	53.00 (± 9.49)	51.00 (± 11.00)	t(17.6) = 0.44, p = 0.67, *d* = 0.19
Neutral	75.00 (± 35.36)	45.00 (± 28.38)	t(17.2) = 2.09, p = 0.05, *d* = 0.94
Happy	80.00 (± 25.82)	70.00 (± 34.96)	t(16.6) = 0.73, p = 0.48, *d* = 0.33
Sad	40.00 (± 39.44)	85.00 (± 24.15)	t(14.9) = -3.08, p < 0.01, *d* = 1.38
Angry	30.00 (± 34.96)	35.00 (± 24.)	t(16) = -0.37, p = 0.71, *d* = 0.17
Fearful	40.00 (± 39.44)	20.00 (± 25.82)	t(15.5) = 1.34, p = 0.20, *d* = 0.60

Therefore, these results do not provide support for the prediction that dancers would perform better on the emotion recognition task than non-dancers, as the difference in overall recognition between these two groups was not significant. In addition, when examining recognition rates at the emotion-specific level, non-dancers were significantly more accurate in assigning the correct emotion label to the sad clips from the McNorm Library than dancers.

Impact of Kinaesthetic Empathy on Recognition

During the interviews, participants were asked the following question: "Did you find you were imagining yourself performing the movements while you were watching them?". Answers were coded as "Yes" or "No" depending on their response to this question, and this was used as an indicator of whether the participants made use of kinaesthetic empathy during the emotion recognition task.

A series of independent samples t-tests were conducted to explore whether recognition accuracy differed between participants who did and did not report using kinaesthetic empathy during the recognition task. For overall performance on the task, the participants who did report using kinaesthetic empathy obtained higher recognition rates (M= 53.08%, SD= 10.32) than those who did not (M= 50%, SD= 10), but this difference was not found to be significant: $t(12.8)= -0.649$, $p= 0.528$.

At the emotion-specific level, significant differences emerged in recognition of neutrality (non-expression) from the McNorm Movement Library clips. Participants who reported using kinaesthetic empathy (M= 76.92%, SD= 25.94) recognised neutral expression with greater accuracy than those participants who reported not using kinaesthetic empathy (M= 28.57%, SD= 26.73): $t(12.1)= -3.897$, $p< 0.01$. At the emotion-specific level, no other differences were found to be significant (all $p> 0.05$). A detailed breakdown of these comparisons can be found in Table 5.6.

Table 5.6: A summary of recognition rates for each emotion category (means and standard deviations) for participants who did, and did not, report using kinaesthetic empathy, and the results of inferential tests performed on the data to determine whether differences between these participants were significant.

	Average Recognition Rate (%)		Inferential Test Results
	Used Kinaesthetic Empathy (N= 13)	Did Not Use Kinaesthetic Empathy (N= 8)	
Overall	53.08 (± 10.32)	50.00 (± 10.0)	$t(12.8)= -0.65$, $p = 0.53$, $d = 0.30$
Neutral	76.92 (± 25.92)	28.57 (± 26.73)	$t(12.1)= -3.90$, $p < 0.01$, $d = 1.84$
Happy	69.23 (± 32.52)	85.71 (± 24.40)	$t(15.8)= 1.28$, $p = 0.22$, $d = 0.55$
Sad	61.24 (± 46.34)	64.29 (± 24.40)	$t(18)= 0.17$, $p = 0.86$, $d = 0.07$

Angry	34.62 (± 31.52)	28.57 (± 26.73)	t(14.3)= -0.45, p = 0.66, *d* = 0.20
Fearful	23.08 (± 33.01)	42.86 (± 33.01)	t(11.9)= 1.24, p = 0.24, *d* = 0.59

These results provide partial support for the idea that kinaesthetic empathy facilitates improved performance on emotion recognition tasks. Specifically, overall recognition did not differ between those who did and did not report using kinaesthetic empathy, but recognition of neutrality (non-expression) in the McNorm Library clips was significantly greater among participants who reported using kinaesthetic empathy. Therefore, these results provide at least partial support for the idea that imagining how the movements would feel within their own bodies aided participants in recognising a lack of emotional expressivity within human movement, but not the recognition of specific emotional states.

QUALITATIVE RESULTS

An overview of the most frequently occurring words in the transcripts can be found in Figure 5.4 below. This visualisation provides a summary of the responses provided by participants during the emotion recognition task and semi-structured interviews.

Figure 5.4: An overview of the 50 most frequently occurring words in the transcripts. For the purpose of this visualisation common words were excluded, and duplicates of words (e.g., with different tenses) were removed.

Through the use of Thematic Analysis, we identified a series of initial codes which were then collated under broader themes. These themes included: 1) Difficulty Defining Emotions, 2) Links to Basic Emotion Theories, 3) Kinematics in Emotion Recognition, 4) Role of Aesthetics in Emotion Judgements, 5) Kinaesthetic Empathy, and 6) Role of Previous Experience. In the following section we will discuss each of these themes in more detail, and provide relevant excerpts from the interview transcripts.

Difficulty Defining Emotions

Whilst coding the data, it became apparent that several participants experienced difficulties with defining what constitutes an emotion. Five out of 20 (25%) participants noted that they struggled to understand whether certain descriptive labels (*pride, enticement, pensive, introverted, hesitant, cautious*) for the movements would "count" as an emotion for the purpose of task. It should be noted that these comments were mainly made by participants when they were asked to consider emotions beyond the five basic emotions included in the forced-choice options, in response to the free-labelling question: "*Is there any other emotion outside of that list of five [neutral, happy, sad, angry, fearful] that you think is a better description for what the dancer is communicating to you?*".

> CB241104-DANCER (P14, L56): *"Is pride an emotion? I'd say pride."*
>
> CB241104-DANCER (P17, L34-35): *"...I don't know if it's an emotion but it's enticement, that feels like an enticement."*
>
> LM1111-DANCER (P79, L37-38): *"I don't know if that's an emotion, but it's quite sort of pensive like introverted…"*
>
> SMW021605-DANCER (P105, L11-14): *"I [laughs] found kind of overall sometimes I questioned if I knew exactly what an emotion was. Like some of the examples I gave, then I'd think is this- does that actually classify as an emotion technically, if you know what I mean?"*

These results suggest these difficulties with defining emotions occur when discussing more complex (and potentially culturally-specific – e.g., pride) emotions.

These excerpts suggest that the conceptual understanding of what actually constitutes an emotion may pose difficulties for some emotion recognition tasks, particularly for free-labelling tasks. They also suggest that these difficulties with defining emotions are most salient when discussing more complex, and potentially culturally-specific emotions (e.g., pride).

Links to Basic Emotion Theories

Whilst coding participant responses to the emotion recognition task, several aspects appeared to relate to components, or limitations, of the basic emotion theories (i.e., James-Lange, Cannon-Bard, Schachter & Singer). These included difficulty defining emotions, the impact of a trigger in the emotional experience, and the social nature of emotion communication. These codes will be discussed in the following section.

<u>Impact of a Trigger in the Emotional Experience</u>

It has already been noted that the basic emotion theories (James-Lange, Cannon-Bard, Schachter & Singer) disagree about the specific mechanisms of emotion production, but these theories do share a number of common characteristics. For example, these theories agree that some kind of external trigger or stressor is required initiate an emotional response.

Although the McNorm Movement Library clips were presented in the form of PLDs, which contain simply a number of white dots moving against a black background (i.e., stripped of any other contextual visual cues), 7 participants (35%) noted imagining the dancer responding to some environmental trigger or stressor when they were assigning an emotion label to the clips. One participant, for example, imagined the presence of a monster that the dancer was trying to avoid, and this prompted the participant to assign the fearful label to the movement sequence.

> CT041605-DANCER (P27, L14-17): *"...I say more fearful... They're really slow movements like sort of like, checking if there is like a monster there…"*

> CT041605-DANCER (P32, L32-36): *"I remember the one uh, there was one that I picked up as fearful when it started like really, uh, still and sort of checking out, I imagine sort of his, the dancer was checking out the room, if there was like anybody or like sort of monster I don't know- a fearful figure in there…"*

Indeed, a total of 5 participants (25%) reported imagined the dancer in the McNorm Library clips to be avoiding, or hiding from, a threat.

> SMW021605-DANCER (P98, L46-50): *"...I kind of feel like it's almost like they're trying not to like be noticed in some way… They don't really seem to travel much it's almost as if they're moving in and out of hiding places, I guess, that type of thing. Yeah it seems like they're trying to hide from something."*

> EH301811-NON DANCER (P123, L38-41): *"...it kind of felt like you were like tiptoeing through a hallway or something. Okay, and then kind of how they lunged really low like they were trying to like sneak by or avoid something."*

Another participant discussed imagining oppressive weather conditions which were placing physical strain on the dancer, and this imagery sculpted their emotion classification.

> SK161007-DANCER (P213, L15-18): *"Hmm. Yes, a texture, or temperature, or a weather-like something like a weather phenomenon. Like I have seconds where I see them struggling with heat, or rain, an external factor making them not comfortable in their body."*

While these examples do not provide direct support for the necessity of an environmental trigger for producing an emotional response in an (observed) individual, it does suggest that an imagined trigger can guide responses to human movement in an emotion recognition task.

It should also be noted that while environmental imagery was described in response to clips perceived to be angry (N=1), and happy (N=1), the majority of these examples were provided as justification for assigning the fearful emotion label to the movement sequences (N=5). In the vignette often used to illustrate the basic emotion theories (i.e., encountering a bear in the woods), the emotion typically discussed is the production of fear. It may be that the fear response, in particular, requires (at least the imagination of) some kind of environmental trigger to appear genuine or obvious to observer.

<u>Social Nature of Emotion Communication</u>

It has also been noted that these basic emotion theories do not account for the social nature of emotional communication. However, a wealth of evidence from previous literature highlights the social functions of emotion (Polo et al., 2016; Van Kleef et al., 2016; Morris & Keltner, 2000; Algoe et al., 2020). Through these interviews, we obtained further evidence to support this notion, as 11 participants (55%) reported imagining the movements in an interpersonal context to assist them with their emotion classifications.

Some of these imagined social contexts were negative. For example, imagined confrontations frequently prompted the selection of anger from the forced-choice options.

> KM161203-DANCER (P70, L33-35): *"I don't know it just kind of looks like a fight, if you imagine like another person. I feel like I can almost see another person fitting into it, doing similar things in it looking like a fight."*

> TC081402-NON DANCER (P207, L38-41): *"...I think there should be another person besides this dancer and it's kind of part of a conversation. And this dancer's trying to show maybe some angriness, but specifically towards another person if he's there [laughs]."*

> TC081402-NON DANCER (P207, L47-50): *"...because, I think, the person is walking towards one direction, and then is immediately turning back. So from that kind of movement, I think he's getting angry with someone else."*

> GW011752-NON DANCER (P148, L30-32): *"...I imagine like another person being there as well in that one. Like that felt very, like, sort of like a workplace kind of squabble almost."*

Others, however, were positive and lead to the selection of happiness or joy; or emotions not included in the forced-choice option list, such as seduction. For example, several participants noted that it looked like the dancer in the McNorm clips was "showing off" to others.

> SK161007-DANCER (P90, L33-36): *"...And I see something seductive or like trying to impress, like I would assume it was something that someone does to impress someone, to be seen, and to convince someone to dance with them. It's like inviting."*

> SK161007-DANCER (P90, L41-43): *"Yes, I imagine like I could totally see this person telling me, "Why don't you want to dance? Look!" as if they're waiting for me to go and dance with them."*

> CB241104-DANCER (P18, L58 – P19, L2): *"Yeah I'd put happy for this in terms of there's an, well I feel there's an enjoyment to it. Like if, like there was someone who got tipsy at a party but they were still in control. But they were really cool, and they just started dancing which is like "watch this"."*

CB241104-DANCER (P22, L23-24): *"…Like here's this person, like I said, was at a party and got slightly drunk but they were like, they were showing off…"*

Others suggested that the dancer in the McNorm sequences looked excited to see their friends.

GW011752-NON DANCER (P146, L16-19): *"It's coming across as someone who's walked into a room of friends, and then like when the knees go, it's like when you see your best friend in the room, if that makes sense."*

GW011752-NON DANCER (P148, L12-16): *"…as soon as I saw that I was like, to me, that's someone who's walked into a pub. Like in a weird way, that's just like something I clocked in my head. I was just like someone's walked into a pub, seen their friends, then seen like a very-like, you know, the person I'm going to sit next to the whole night."*

These specific excerpts, and the frequency with which such comments were made throughout the interviews, provide evidence to support the idea that the social nature of emotions plays a significant role in the recognition of emotional expressivity.

<u>Social Communication Styles</u>

Further evidence for the social nature of emotions can be found when examining the communication styles participants used when they struggled to verbalise their thought processes. Thirteen out of the 20 participants (65%) noted that they struggled to verbalise their thought processes. Of these 13 participants, 6 dancers reported finding it particularly difficult to discuss movement verbally.

CR201309-DANCER (P35, L51-56): *"I do, like it's hard to describe it sort of, like I just know the feeling of that sort of… Oh it's so hard to put into words. When you do like those expansive sort of movements, for me it feels like I'm sort of dancing through space, and I think there's a lot more joy in that. Sorry, it's probably not helping them in terms of describing things. It's really, it's quite hard."*

CB241104-DANCER (P14, L8-11): *"…I'm thinking of like, yeah, there's kind of- it's- I don't know I'm going off on one here. Yeah, God, Jesus I just can't- I knew this was gonna happen. I have no idea how to communicate dance in verbal form."*

Often when such communication difficulties arose, participants would resort to "acting out" a specific scenario to illustrate their ideas. Out of the 20 participants, 8 dancers (80%) and 3 non-dancers (30%) communicated with the researcher in this way.

PS291837-NON DANCER (P168, L54 – P169, L4): *"…But then, it's like sometimes you can get a person who can watch a movie, and you see the same movie, and you're like "oh did you see that bit and that bit?", "oh no I didn't notice that bit or that bit, oh I'll need to watch it* *again to see that". Whereas like sometimes people go to see the same movie but they see different movies, if that makes sense. They see different bits, different clips, different- say for example, sometimes a person says "oh such and such is in that movie", "are they? Is she? I didn't see them, no, but who are they playing?"…"*

PS291837-NON DANCER (P170, L33-40): *"…I mean who knows, they could be having a bad day but they're smiling all the time, and you're like "are you okay?", "yeah I'm fine, aye, no problem" [laughs], you're like "are you really fine?", "yeah, yeah". You know, you sometimes get that, they can keep it hidden. So I find that sometimes with people, it's hard to kind of read them sometimes, and sometimes people have got a poker face and you don't know. "How are you feeling?", "Aye, aye, I'm fine, I'm okay" [laughs], but you don't know that, or they could be really good and really bad, you know."*

KR281508-DANCER (P58, L23-27): *"I can see this one being an angry one. But not in a loud, shout in your face like I'm so angry. But, especially this bit it's almost like hurt rather than angry? Like turn back and almost look back as if it's like "you've done this to me". Here's me expressing how I am hurt."*

KM161203-DANCER (P64, L53-54): *"I don't know how to explain it, but it's almost like a very "look at what I'm doing", if that makes sense."*

It is possible that drawing attention to a social scenario that is familiar in everyday life and 'acting out' the response they would typically provide was an easier way for participants to communicate their understanding of emotional states demonstrated by the dancer in the McNorm Library clips. In other words, their descriptions were based on a perceived mutual understanding of social communication scenarios. It may be that, rather than having to verbalise a thought or idea, participants relied on a theoretical understanding (on the part of the researcher) to communicate their perception, thus providing further evidence for the need to capture the social aspects of affective communication.

Kinematics in Emotion Recognition

Many researchers have tried to identify a unique profile of movement kinematics associated with specific emotions, but this has produced mixed results so far (Sawada et al., 2003; Wallbott, 1998; Gross et al., 2010; Brownlow et al., 1997). By a substantial margin, the most consistent finding appears to be the role of speed in differentiating between different emotions, more specifically between high and low intensity emotions.

We obtained substantial support for the importance of speed in emotion recognition, as in these interviews all of the participants (100%) noted that speed played a role in their decision-making.

TC081402-NON DANCER (P206, L43-45): *"So I sort of think this is quite similar to the one I've select neutral for, but this time is much slower and it kind of have a sad vibe to it because, just because it's slower."*

TC081402-NON DANCER (P206, L55-56): *"Yeah, because I think the movement is the same it's just the speed that makes a difference."*

CB241104-DANCER (P17, L54 – P17, L4): *"We've seen this one before haven't we?" [RESEARCHER: "I guess that's for you to say."] "Or is this a combination of the two? Is this the other one but faster? God, now I'm confused. I think this is the other one but faster isn't it? Yeah it is the other one but faster. And that is slightly closer to anger."*

EH301811-NON DANCER (P124, L42-43): *"...I probably wasn't looking for emotions at first, I think I was looking for speed at first."*

SM031603-DANCER (P187, L36): *"High speed movement tends to me to speak of heightened emotion."*

These frequency with which speed was discussed during the emotion recognition task, and the apparent importance of speed to participants' judgements noted in these excerpts, adds to the literature suggesting that movement speed plays a key role in whether participants assign high or low intensity emotion labels to movement sequences (Smith & Pollick, forthcoming; Halvoic & Kroos, 2018; Gross et al., 2012; Roether et al., 2009). Speed as a kinematic parameter therefore appears to be important from a computational standpoint (evident from previous kinematics research), but also from a perceptual standpoint as participants in this experiment were aware of how speed sculpted their emotion decisions.

Role of Aesthetics in Emotion Judgements

Previous research has implied that affective and aesthetic judgements are intrinsically linked (Gernot et al., 2018; Marković, 2010; Jankowski et al., 2020; Sachs et al., 2015; Brattico et al., 2016; Van den Tol et al., 2016; Christensen et al., 2021). Additional evidence for this relationship can be observed the similar mechanisms involved in producing aesthetic and emotional experiences. In these interviews, we obtained further evidence that these processes may be linked, at least for the processing of abstract, whole-body dance movements.

Although the forced-choice task and interviews did not involve any specific mention of aesthetics, nor were participants asked to provide any aesthetic judgements for the clips, 17 of the participants (85%) provided spontaneous aesthetic evaluations of the movements in the McNorm clips. These judgements included the standard dimensions of *liking* (7 participants), *interest* (2 participants), and *beauty* or *prettiness* (each by 1 participant). Other aesthetic-related judgements were also provided by participants (*nice* and *graceful* each noted by 4 participants, *elegant* and *impressive* each by 3 participants, *sexy* and *cool* each by 2 participants, and 1 participant noted each of the following characteristics; *awesome*, *amazing*, *enjoyable*, *eye-catching*, *forgettable*, *funky*, *good*, *lovely*, *poised*, *satisfying*, *surprising*).

> *SK161007-DANCER (P89, L26-27): "I really like this [indicates to clip] the lifting of the leg, this is impressive and really interesting."*

> *CB241104-DANCER (P17, L30-36): "Oh this person's sick. Yeah that's cool, I like that. Out of those I'd say happy just because it's the only- is that like the only positive one? ...But there is a better word I would use to describe it, which is kind of, its, I don't know if it's an emotion but it's enticement, that feels like an enticement. That's kind of sexy, I like that."*

These examples provide evidence to suggest that aesthetic appraisals occurred spontaneously (i.e., unprompted) during the emotion recognition task, and that these judgements were sometimes used to inform the emotion judgements. These examples also provide evidence in support of the "what is beautiful is good" hypothesis, as participants tended to provide a positive aesthetic appraisal (only 2 negative characteristics were reported; *disliking* and *forgettable*, each by 1 participant), which then typically prompted the selection of a positive emotion label (i.e., happy).

Indeed, even more specific evidence emerged from our investigation that suggests that aesthetic evaluation influenced participants' emotion judgements for the movement sequences. One dancer participant noted that they felt they could not assign the label of anger to the movement sequence, despite the fact they felt it may be most accurate, because they found it difficult to reconcile something beautiful with that particular negative emotional state.

> *CR201309-DANCER: "This one's quite hard to..."*

> *RESEARCHER: "Struggling a bit more with this one?"*

> *CR201309-DANCER: "Yeah. I find it hard because I want to say angry, but see because the ballet movements are so beautiful it's so hard to- [laughs] Because for me, angry is something really that's not beautiful in a sense. But I think from the structure of the movement, especially the very beginning which was really destructive, like you weren't expecting. And the angular positions. I think I'm going to go for angry again."*

RESEARCHER: "Okay, would you say you find this one beautiful?"

CR201309-DANCER: "Yeah, yeah."

RESEARCHER: "So is that kind of where that struggle for angry comes from? Would you say you, kind of for you, it's hard to match up something that's beautiful with something that's angry?"

CR201309-DANCER: "Yeah. I think so."

In review, as aesthetic evaluations were provided spontaneously (i.e., unprompted) during the emotion recognition task, and considering the specific difficulty one participant experienced in assigning a negative emotion label to stimuli perceived to be beautiful, this provides crucial support for the notion that an intrinsic link may exist between affective and aesthetic processing.

Kinaesthetic Empathy

Previous research has suggested that domain-specific empathy skills can help with processing of the domain-specific stimuli (e.g., musicians possess a set of music-specific empathy skills; Peters, 2015; Van der Schyff & Krueger, 2019; De Bruyn et al., 2011; Garrido, 2017). In dance, the greater use of kinaesthetic empathy is one reason why dancers are thought to perform better on emotion recognition tasks (Reason & Reynolds, 2010; Behrends et al., 2016; Seifert, 2018). This is believed to be linked to the concept of embodiment and, in turn, to empathy. It is argued that the ability to emulate the movements of others within our own body allows us to better place ourselves in another's shoes. However, these claims tend to be inferred from neuroimaging data of dancers (i.e., greater activity in motor areas for familiar movements), rather than by asking dancers specifically. As a result of this gap in the literature, it is unclear the extent to which kinaesthetic empathy is universally used across dancers, whether this is exclusive to dancers or movement experts, and the extent to which those who do use kinaesthetic empathy believe this aids their ability to recognise emotion from the movement of others.

During these interviews, 7 out of the 10 dancer participants (70%) reported making use of kinaesthetic empathy, and explained that doing so assisted them with assigning emotions to the McNorm Library stimuli.

LM1111-DANCER (P82, L25-27): "...I think it's also interesting, I find it easier to think about the emotion not looking at it, but trying to do it [hand movements] like a little bit."

CC051508-DANCER (P9, L46-50): "… I think it was the first one I was quite unsure about, and I went to say neutral. And then, based on looking at it and thinking about how I would feel doing those movements I started changing my answer to more happy. Like I can picture doing those movements and like smiling [laughs]."

CB241104-DANCER (P21, L20-41): "…I am putting myself in their position and I would imagine what I would be feeling. That's how I, that's how I judged it. …that's basically what I was doing. I was like, I'm empathizing with them… that was essentially what I was doing I was just looking at it from like, if I was performing this, this is what I'd perform it with. This is what I'd give it so yeah. … Yeah, it helped firm up the decisions, because I would go off my initial imprint and then like, my initial watch it, and then I'd go yeah that's how I'd be performing that piece. Like, it was like see someone else, be like yeah I'd do the same as that so yeah."

KR281508-DANCER (P59, L44-47): "It's almost, I could feel myself as I watched it like multiple times then, [arm and torso movement] not doing this [laughs], I'd like- I'd start like picturing how it would feel to do it, and then I'd just think how like, my emotions would be as I was dancing it."

These excerpts provide support for the idea that dancers make use of kinaesthetic empathy, or motor simulation, when observing other bodies in motion, and that imagining how those movements would feel in their own body helped them to assign an emotion label to the movement sequences.

It is also of note that one of the dancer participants used a combination of both visual and kinaesthetic imagery to make their decisions. They discussed making their emotion judgements based on how doing the movements would feel within their own body, but they also noted that for some of the clips they envisioned a visual image of themselves performing the movements from the McNorm clips. They noted that which method they used differed depending on the perceived difficulty of assigning an emotion label to a particular clip.

KR281508-DANCER (P60, L9-16): "The ones that I kind of imagined myself watching almost were like clearer emotions. The ones that I could picture myself dancing were much harder to like, decipher what I'd then be feeling. I think it was just strange because it shouldn't like. In my head I think if I was what it's done and it like, in my head I'd think if I was the one that was dancing it I'd know how I feel. But then it was almost then like imagining yourself dancing it, and then imagine watching myself dance it, and then that almost felt different as well."

However, through these discussions we observed that kinaesthetic imagery was not used universally by dancers, as 3 of the dancer participants (30%) explicitly stated that they did not imagine simulating the movements within their own bodies.

> *CC051508-DANCER: "Like I didn't really, yeah, didn't really like relate it to my own life, or like picture it being, yeah, myself particularly."*

In addition, while 4 non-dancer participants (40%) did not report using kinaesthetic empathy or motor simulation to make their emotion decisions, 6 non-dancers (60%) did report imagining themselves performing the movements to assist them with their emotion decisions.

> *[RESEARCHER: "Okay. Did you find you at any point imagined yourself doing the movements?]* EH301811-NON DANCER (P124, L52): *"No." [RESEARCHER: "No. If you had do you think that would have helped with your decisions?"]* EH301811-NON DANCER (P124, L57): *"No, probably not."*
>
> SM031603-NON DANCER (P192, L35-36): *"I think I might have tried to do one or two of the movements as well, yes."*
>
> SM031603-NON DANCER (P192, L48-50): *"I think I sort of, yeah I did the movement and I sort of correlated it in my head with the emotion that would make me do the movement, as it were."*
>
> TC081402-NON DANCER (P209, L25-27): *Yeah because I first trying to do all the movement myself and then trying to imagine what will I feel if I act like that, so yeah. That's how I judge the emotion going on.*

These excerpts suggest that domain-specific empathy skills (i.e., kinaesthetic empathy) may be of importance to some observers, but appears not to be universal, nor unique, to movement experts. Therefore, this is not the sole reason why previous experience may lead to improvements in performance in emotion recognition tasks, and should not be considered as such going forward. There may be other facets of previous experience that facilitate the improvements in emotion recognition capabilities observed in dance-experienced samples.

Role of Previous Experience

Extensive evidence suggests that dancers perform better in emotion recognition tasks than non-dancers. But, if kinaesthetic empathy alone does not distinguish between movement experts and non-dancers, through these interviews we uncovered a number of additional aspects of previous experience which may facilitate these improvements. Findings in this area related to the themes of learned associations in dance, knowledge of performative dance, learned associations from real-world experiences, and learned associations from popular culture. These themes are discussed in more detail in the following section.

Learned Associations in Dance

In these interviews, 9 out of the 10 dancer participants (90%) indicated that they had learned associations between specific movements and particular kinds of emotional expressivity from their own dancing experiences (as performers and as choreographers).

> LM1111-DANCER (P80, L24-29): *"Yeah. I would associate that with… Like especially quicker turns. Like anything fast I would associate with something either like anger or, or like, but it was essentially faster movements and turns and sort of the turning away from the audience but looking I would very much associate with like joy and happiness or something flirtatious or yeah quite out- it feels quite outgoing also this one."*

> LM1111-DANCER (P82, L48-53): *"Yeah I think it's similar with, with jumps I tend to think more joyful, more happy also more angry, maybe. Whereas, I think anything slow, especially… Hm, I think slow can be, can be fearful or sad or sad, especially. But like it can also be intimate like sort of for me is very slow movements or like développés or things like that they also happen in pas de deuxs."*

> SK161007-DANCER (P92, L42-53): *"…when I see a dancer being on one leg and lifting all other limbs, this is associated with freedom and joy. Absolutely, when I see someone doing this comfortably. When I see someone lifting their leg, but in a straight, strict position this is associated with discipline to me, and concentration. And then I really like seeing dancers- I really like when they bend their back because this is something that I find really- I use a lot in my work, in cabaret- like as I said, cabaret is the dance that I have the most experience with, it's important to have the lower back arced so this is something charming, this is something that I find charming when I see a dancer lifting their head and curving their back, I would say seduction. It conveys seduction to me."*

> KM161203-DANCER (P73, L2-14): *"… I feel like we would always be told like [claps] sharp, you need to make it sharp [claps] and that's just stuff that's like almost stuck in my head. When I was doing it, I would like try [hand movements] and do it like as sharp as possible. I think because that was one of the things I could actually do, like it was one of the things I think I would get kind of complimented on. It would be like "oh [KM161203]'s good at doing that*

sharp", so I think it's one of those things that I would really focus on, because I know like maybe someone else would be better at pirouettes, maybe someone else would be better at this. So I think that kind of like… you're almost told to [hand movements – imitating feet movements from clip 10] kind of put anger into it. So I think that's maybe why that's stuck in my mind for that. Whereas someone else maybe wouldn't have that association."

Another dancer participant indicated that they held associations between particular movements and the expression of a particular emotion based on their experience of choreographing a performative dance sequence.

> SMW021605-DANCER (P103, L6-12): *"…One of the ones, when they bring your arms [arm movement] round I think they turned and brought their arms to the other side, the fourth. I choreographed a dance like June in my last year of my undergrad for one of the competition teams and that was in my dance. So that dance was very upbeat, it was to Physical by Dua Lipa, so I kind of think of joy when I see that, and happiness if that makes sense, and it reminds me of a happy time."*

These personal associations seemed to be of particular relevance for the identification of neutral (or non-expressive) movements, as 7 out of the 10 dancers (70%) reported associating the specific movements included in the neutral stimuli with movements they had performed as warm ups, or technique exercises in a training setting (i.e., classes), and this helped to guide them to provide the correct response.

> CB241104-DANCER (P16, L56-57): *"I mean I'm just going to be putting neutral, because I know they're at the barre or they're in the centre doing stuff…"*

> CB241104-DANCER (P17, L15-19): *"…yeah it's kind of, that's an exercise, so at least- so the purpose of that is for exercise so yeah, you sort of do it neutrally. …I just thought that's, I know what that person is doing."*

> CR201309-DANCER (P18, L25-30): *"I think for this one I'm going to go for neutral. Maybe because I recognize the exercise is like a warm up. Something you would do in class. Which I think I associate more in terms of getting into the body, rather than expressing something. Say it's in like, I don't think the dancer's in like a performance mode, it's more just internally, for themselves. Yeah and it's very, it's functional movement."*

> SMW021605-DANCER (P98, L16-18): *"It does look very much kind of just like a standard ballet barre exercise or like a centre plié exercise. So I think because of that, because it's kind of standardized, it just seems neutral overall."*

SMW021605-DANCER (P102, L47-54): *"…I think, also because I do ballet, it's one of like my main styles, I'm just very used to associating them more with technique, rather than like performance. So I think it's probably ingrained in me to just see them and think of them as being quite neutral because part of the whole kind of essence of those exercises, I guess, is about being controlled, and keeping your composure, and making sure your technique is as good as possible. So I think when I see them, I just instantly think of that."*

These results suggest that learned associations with specific movements can influence the emotional expressivity observers assign to those same movements. This may be particularly salient for experiments which involve presenting dance movement stimuli to dance-experienced samples.

Knowledge of Performative Dance

The dancer participants in our study sample also drew upon their professional knowledge and observational experience with dance in general, and performative ballets specifically. Four out of the 10 dancer participants (40%) used these associations to draw comparisons between the experimental stimuli and specific ballet stories.

CB241104-DANCER (P21, L58 – P22, L7): *"Like the first one, because they weren't really going that far above like the legs were- their attitudes and arabesque weren't going above, but this is like a technical thing. So I imagined it's a romantic ballet because back then the skirts didn't allow- they stay below 90 degrees, they didn't go above. So I was thinking like Giselle and that. And then later on, like the Swans you- there's just a way- yeah, so I did associate it with stuff, I was like "I imagine that's kind of like Swan Lake?". That's kind of Swan like movement there…"*

CB241104-DANCER (P23, L22-25): *"In a state of sorrow it would be the head would be down, like I'm thinking Swan Lake, but like there's it's a kind of- a sort of need to go down, like go to the floor kind of thing. Which is what happens in Swan Lake…"*

CC051508-DANCER (P2, L53-55): *"It's like, very like big and it reminds me of, oh what's the one in Nutcracker with like the two guys and they're like tumbling and stuff? It reminds me of that."*

However, we also found that three out of the 10 non-dancers (30%) also related the movements they were observing in the McNorm Library stimuli set with ballet performances they were familiar with. It has been suggested that even visual experience (without motor experience) of dance leads to changes in movement perception. It may be that both dancers and non-dancers who have observed performative ballet may draw on previous knowledge to associate particular movements with narratives, to aid them in assigning emotion categories to expressive motion stimuli.

> RT231409-NON DANCER (P181, L17-19): *"…you know, you're looking at it, you're thinking that's reminiscent of Swan Lake or, you know, a movement in Swan Lake or something like that."*
>
> SD081104-NON DANCER (P201, L4-6): *…I was sort of visualizing like trees and forests and things. Just sort of that just must be like my natural association with like dance performances."*

These results provide additional support for the idea that observational experience may sculpt the perception of emotionally expressive human movement, regardless of motor experience.

<u>Learned Associations from Real-World Experiences</u>

Drawing upon learned associations and knowledge of performative dance was utilised much less by the non-dancer participants. Instead they tended to base their emotion judgments for the movements on their memories of experiencing those same emotions in their everyday life. In these interviews, 6 out of the 10 non-dancers (60%) discussed specific circumstances where they had performed similar movements, physically, in response to events or experiences in their own lives. For example, one participant discussed a very vivid example, where they related movements in the McNorm Library clips to their own experiences of dancing in a bar when on holiday, and used this previous experience to assign the happy label to the clip.

> PS291837-NON DANCER (P161, L26-30): *"Yeah. Very happy, like the blue sky in the summer, you're in the bar, you've got a few beers, you know, like going out and having a bit of a cha-cha-cha kind of thing, you know [laughs]. It kind of puts you in the mind of holidays, if you were in Spain, Portugal or wherever, kind of, that's the impression I get."*
>
> PS291837-NON DANCER (P168, L19-27): *"…like the salsa one, if you put on like that song, what was the holiday song years ago, you might remember it if you're young enough, Las Ketchup, that was the one where you kind of [laughs] do that dance [hand movements]. That kind of brought that back in my memory, and it's just like you know, you've got that memory of the sun setting, you've got the kind of pina coladas in the bar and all that, that's playing in the background and you think "oh cool", you know. But it's amazing, you know, how one thought can trigger a whole lot of memories…"*

Another non-dancer participant mentioned finding some of the movements in the McNorm clips to reflect the way they celebrate when their favourite football team wins.

RT231409-NON DANCER (P180, L45-48): *"Again, just perhaps the first one, the fact that, you know, it seemed to be carefree and I think we've all experienced that sort of feeling at times. Particularly when Celtic are gubbing Rangers 3-0 on a Wednesday night [laughs]."*

Similar anecdotes were provided by non-dancers who noted that some movements in the McNorm Library clips reminded them of negative previous experiences, and that they based their judgement of anger on interpersonal confrontations they have had in the past.

TC081402-NON DANCER (P208, L44-47): *"I think there's one I told you about there should be another dancer beside this dancer. Yeah, for that one because that picture kind of appeared to me quite vividly, I'm guessing that's from one of my life experience arguing with another person. Yeah [laughs]."*

These results suggest that various types of previous experiences, not just those related to dance specifically, may influence the affective perception of the human body in motion.

<u>Learned Associations from Popular Culture</u>

In addition to real-world experiences, 6 out of the 10 non-dancer participants (60%) made a number of references to popular culture when explaining the associations they had with some of the movements in the McNorm clips. There were references made to films and television programmes, some of which were specifically related to dance (e.g., Strictly Ballroom, Strictly Come Dancing), and others which were not obviously related to dance (e.g., Toy Story, The Titanic). Musicals, such as West Side Story, were also discussed.

PS291837-NON DANCER (P171, L22-30): *"…But the one where it was like, the sad one or like the kind of the lamenting one, I was imagining a bit of The Titanic soundtrack, or the sad soundtracks, or even like in that movie, you might have seen it years ago, Toy Story. You've got the sad bit where it's like Woody is left on his own, and like in Toy Story 2 when Jessie's left down the back of the bed and like years go past, and it's playing a sad song and she's sitting with her wee head down, and that kind of brings it back to you, you've got that in your mind, that image, of making them sad…"*

SM031603-NON DANCER (P191, L38-39): *"…I don't watch a lot of dance, videos and things in general but Strictly Ballroom, that influenced my decisions…"*

RT231409-NON DANCER (P181, L56 – P182, L4): *"You know, just think perhaps maybe West Side Story, you know that there's you know, obviously between the Jets and the Sharks, and that's quite dynamic, and rapid action if you know what I mean, it's sudden. Jerky movements almost would tend to sort of lead you to think, aye there was anger in that."*

LJ181312-NON DANCER (P157, L18-20): *"…most like have the happy dance is like in some movies, like cartoons, when you see them dancing they usually do almost the same as in the moves as those…"*

These excerpts provide some suggestion that previous motor and observation experiences which are not specific to performative dance, may also influence success in emotion recognition tasks. Overall, these comments provide some clue as to how previous experience can influence emotion judgements assigned to emotionally expressive movement, beyond the role of domain-specific empathy skills or motor simulation.

Other Factors

Several other interesting factors emerged during these interviews that are beyond the scope of the current manuscript. These included references to visual attention strategies (e.g., holistic visual processing, movement-driven visual attention, experience-driven fixations, and body region-specific attention), and the role of other imagined social cues (e.g., music, facial expressions, morphological body information). However, all interview data is publicly available on the OSF (https://osf.io/wbm6s/), and we would encourage other researchers to make use of this rich source of data to pursue new research questions.

DISCUSSION

1. QUANTITATIVE FINDINGS

1.1. Emotion Recognition Results

In comparing the emotion recognition results obtained in this study, we observed a number of similarities to those obtained in the original McNorm validation experiments (Smith & Cross, 2022). First, in the present experiment, participants obtained an average recognition rate of around 52% for the selected clips from the McNorm Library. These results align extremely well with recognition rates reported in the original validation experiments (52.2% in Experiment 1, 49% in Experiment 2). Additionally, as in the original validation experiments, all emotion categories (neutral, happy, sad, angry, and fearful) were recognised at greater than chance level. Finally, similarities can be found when examining recognition rates at the emotion-specific level. In the original validation experiments movements depicting the expression of fear were recognised with the lowest level of accuracy (38% in Experiment 1, 39% in Experiment 2), and in the present experiment fear was also recognised with the lowest accuracy (30%). Despite the small sample size for a replication study, these results do provide compelling evidence for the high consistency in recognition of emotional expressivity from whole-body dance movement clips in the McNorm Library, and provide further evidence of its utility in social cognition research.

1.2. Dance Experience Results

In the current study, participants with extensive dance experience recognised the intended emotions with slightly greater accuracy (53%) than non-dancers (51%), but this difference was not found to be significant. The only significant difference that emerged between these groups was that non-dancer participants recognised the expression of sadness with greater accuracy than dancers. This result is surprising considering the previous literature. Researchers have observed on multiple occasions that dance experience sculpts the perception of human movement (Kirsch et al., 2015; Cross et al., 2011; Bläsing et al., 2012; Orgs et al., 2018), and that dancers perform better in emotion recognition tasks than non-dancers (Christensen et al., 2016; Christensen et al., 2019; Christensen et al., 2021). However, it should be noted here that in the original McNorm validation experiments (which examined responses from dance-experienced and non-dancer observers) neither physical, nor observational, dance experience had a significantly positive effect on the recognition of emotional expressivity. It should also be noted that in the experiment conducted by Christensen and colleagues, movement was typically depicted in the form of full-light displays, while clips in the McNorm Library depict such movements in a reduced format (i.e., as point-light displays). It is possible that greater morphological information is required for dancers to successfully identify emotional states from the dynamic human body, particularly when considering the widely ranging choreographic repertoires associated with different dance styles.

The movement repertoires of certain non-Western dance styles (e.g., Bharatanatyam, Mohiniyattam) and some Western dance styles (e.g., Scottish Country Dancing) rely on specific movements or positioning of the hands and fingers, and Western ballet repertoires are centred around five specific positions of the feet (‘*cinq positions des pieds*’). However, due to the reduced visual format of the McNorm Library’s PLD movement representations, fine motions of specific body parts (e.g., fingers) and specific limb positions are not clearly visible to observers (as the feet and hands are represented as singular markers). This lack of visual detail may be particularly important to take into consideration in future emotion recognition tasks using dance-experienced samples. We suggest that it may be useful to conduct a similar task exploring emotion recognition in dancers and non-dancers using both PLD and FLD representations of emotionally expressive whole-body dance movement sequences.

1.3. Kinaesthetic Empathy

As part of the semi-structured interviews, participants provided yes or no answers when asked whether they imagined themselves performing the movements while they observed the McNorm Library clips. These responses were coded to provide a binary measure of the use of kinaesthetic empathy; which was entered into a two-sample t-test to explore group differences in average recognition of emotional expressivity in those who did and did not imagine simulating the movements within their own body. Although no significant differences were observed in overall task performance, we observed that participants who reported using kinaesthetic empathy were more accurate in identifying neutral clips from McNorm Library than those who did not. These results suggest that using kinaesthetic empathy in this emotion recognition task may not have assisted with identifying specific target emotions, but did assist participants in discriminating between expressive and non-expressive whole-body movement sequences more generally.

It has already been noted in this manuscript that evidence supporting a relationship between kinaesthetic empathy and improved performance in emotion recognition tasks is often inferred from the results of related, but disparate, quantitative experiments. We aimed to probe the extent of this relationship through our semi-structured interviews, therefore additional discussion of qualitative kinaesthetic empathy results can be found in section 2.5. of this chapter.

2. QUALITATIVE FINDINGS

In the following, we discuss results from the semi-structured interviews pertaining to the following themes: 1) Difficulty Defining Emotions, 2) Links to Basic Emotion Theories, 3) Kinematics in Emotion Recognition, 4) Role of Aesthetics in Emotion Judgements, 5) Kinaesthetic Empathy, and 6) Role of Previous Experience.

2.1. Difficulty Defining Emotions

During the experimental proceedings, a quarter of the participants specifically noted being unsure of whether the descriptive label they had assigned to the McNorm Movement Library clips "counted" as an emotion. This ambiguity in what exactly constitutes an emotion is mirrored in the range of dictionary definitions provided to summarise the term. In the Merriam-Webster Dictionary, the first entry under the definition of emotion is, "a conscious mental reaction (such as anger or fear) subjectively experienced as strong feeling, usually directed towards a specific object and typically accompanied by physiological and behavioural changes in the body" (Merriam-Webster(b), n.d.). This definition brings attention to a number of key identifying components of emotion; including the need for conscious awareness of the affective response and the experience of strong feelings, and the typical requirement of both a trigger and a bodily response. However, previous behavioural and neuroscientific research highlights some issues with the specificity of this description. In a study conducted by Sato and Aoki (2006), participants were asked to provide preference judgements for a series of ideograms which were preceded by 25ms positive or negative valence primes (happy and angry faces) on the left or right of the target ideogram. They found that the participants were not consciously aware of the presence of the primes, but also found that the presence of a negative prime significantly reduced liking ratings assigned to the target ideogram, compared to the positive or neutral primes. These results suggest that the experience of emotions, and production of subsequent behavioural responses, do not always require conscious attention.

Returning to the definitions of emotion in the Merriam-Webster dictionary, the second and third entries are simply, "a state of feeling" and "the affective aspect of consciousness", respectively. While these latter entries may be more appealing than the first due to their simplicity, they leave room for ambiguity in conceptually and operationally defining emotions for social perception research. This is perhaps best illustrated through evaluation of the experience of "calmness". We choose to discuss this example, as two participants noted difficulty in deciding whether "calmness" would count as an emotion when describing the movement clips in this experiment (CR201309, pg. 34, L52-53; EH301811, pg. 117, L41). Taking the first definition, calmness would likely not be classified as an emotion, given where it is situated on the spectrum of arousal, as it does not meet the proposed requirement of producing strong feeling. However, calmness would be classified as an emotion under the second definition, as there is a wealth of evidence highlighting the physiological and cognitive effects associated with the 'feeling' of being calm.

We should note, that neither we nor the participants in this experiment are the first to draw attention to this issue. In a recent essay, Scheff (2015) provides an exceptional summary of the uncertainties in this area when he says that the, "meaning of words that refer to emotions are so confused that we hardly know what we are talking about" (Scheff, 2015, p111). Moving forward, it would be beneficial to develop a definition for what we mean by "emotions" in the context of social perception research, and it may be useful to provide this definition to individuals who participate in emotion recognition tasks; particularly in cases that involve free-labelling.

2.2. Basic Emotion Theories

It has already been noted that few existing theoretical frameworks capture the real-world human experience of emotions. However, the seminal emotion theories proposed by James-Lange (James, 1884; Lange, 1885), Cannon-Bard (Cannon, 1927; Bard, 1928), and Schachter and Singer (Schachter & Singer, 1962) have allowed us to develop a thorough understanding of how physiological and cognitive processes interact to produce emotional experiences. A number of participants in these interviews alluded to certain aspects, or limitations, of these theories during the emotion recognition task.

2.2.1. Impact of an External Trigger

While the James-Lange, Cannon-Bard and Schachter and Singer theories of emotion disagree about the specific mechanisms underlying the elicitation of emotions, they do share a number of common features. First, they broadly agree that physiological responses and cognitive processing are required for an emotion to be felt. Secondly, they all indicate the need for some external trigger or stressor for the experience of an emotion to be produced. In these semi-structured interviews, a number of participants reported imagining the presence of some kind of trigger in the environment of the McNorm Library movement clips, that the dancer was reacting to, when they were making their emotion decisions.

It should be noted again here that the clips from the McNorm Library represented the movement of the dancer in the form of point-light displays. By reiterating this point, we draw attention to the fact that the clips contained the movements of only one dancer (depicted as a series of moving white dots without morphological body information) presented on a black background. Therefore, the clips contained no visual information that could indicate features of the environment, highlighting the fact that all triggers were completely imagined by the observers during the emotion recognition task. While these comments do not provide support for the necessity of an environmental trigger (as is suggested in the traditional emotion theories), they do provides support for the idea that an imagined trigger may strongly invoke the perception of certain emotional states.

It is unclear how employing the use of such imagery (e.g., a general threat, a monster, or extreme weather conditions) contributed to the participants' emotion decisions. However, it should be acknowledged that in the majority of these cases, these examples were provided as justification for assigning the fearful emotion label to the movement sequences and this is interesting when we consider the vignette typically used to illustrate the basic emotion theories. Typically, to illustrate differences in the emotion elicitation processes proposed in the James-Lange, Cannon-Bard, and Schacter and Singer theories, we are encouraged to consider the following situation: we encounter a bear in the woods, then in some order we experience a physiological response (e.g., sweating, increased heart rate) and process cognitively (i.e., we associate growling with aggressive animals and know this means danger), which produces the experience of fear. Therefore, fear is often used as the exemplar emotion in such illustrations. When considering both this, and the fact that the fear response was most frequently associated with the imagery of an environmental trigger during this experiment, we suggest that perhaps the communication of fear, in particular, is most dependent on some external trigger (be it real or imagined).

This may provide some tentative explanation for why fear is often recognised with the lowest frequency in emotion recognition tasks (in this manuscript, and in the previous literature; Smith & Cross, 2022; Pasch & Poppe, 2007; Camurri et al., 2003; Dittrich et al., 1996; Grèzes et al., 2007; Dahl & Friberg, 2007). When creating emotionally salient stimuli for social perception research it is not possible, for obvious reasons, to introduce any genuine threat or trigger to the experimental proceedings. However, it is possible that introducing an emotion elicitation task to the stimulus creation methodology may be useful for the emotional expressivity captured in such stimuli to appear more genuinely authentic, and may translate to greater recognition among observers. Indeed, there is evidence from facial expression research to suggest that observers are sensitive to the authenticity of emotional expressivity. A number of studies report that observers can differentiate between spontaneous (Duchenne) and posed (non-Duchenne) smiles (Etcoff et al., 2021; Krumhuber & Manstead, 2009; Trutoiu et al., 2014). Additionally, it has been found that observers can be specifically sensitive to the detection of inauthenticity in facial expressions of fear (McLellan et al., 2010).

It is unclear the extent to which these findings from facial expression research extend to the communication of genuine or inauthentic portrayals of emotional expressivity, and of fearful expression in particular, for whole-body human movement. Again, for obvious ethical reasons, it is not possible to generate authentic emotional reactions through exposure to potentially harmful triggers in the creation of stimuli sets for social perception research, but emotion elicitation tasks could present a viable alternative (Gross & Levenson, 1995; Clark, 1983; Sun & Yin, 2008; Cowie & Cornelius, 2003). To our knowledge, no research to date has explored differences in the perception of expressivity from a set of movements recorded following an emotion elicitation session, versus a matched set of movements that were not preceded by such a session. Such an experimental approach has the potential to reveal the importance of authenticity in expression for the successful recognition of emotion, and may shed some light on whether this is of specific importance for the successful recognition of fear.

2.2.2. Emotions in a Social World

It has already been noted that a significant limitation of existing emotion theories lies in the fact that they only provide an account of how emotions are produced within the individual, and do not account for the social nature of emotional experiences as they occur in the real-world. We obtained evidence from these interviews which suggests that the social aspects of emotional expressivity play a vital role in the recognition of emotion from whole-body movement.

Various examples of what we term *interpersonal imagery* emerged throughout the interview transcripts, as many of the participants reported imagining the dancer in the McNorm Library clips within a social context to assist them with their emotion judgements. Unlike with the environmental imagery, instances of interpersonal imagery were discussed evenly in both positive and negative contexts. For example, several participants reported believing the dancer to be engaged in a confrontation with an imagined other (e.g., a workplace squabble, argument with a friend), and others mentioned that they imagined the dancer to be present at some sort of gathering or celebration (e.g., dancing at a party, seeing their friends in a pub). These comments provide some evidence to suggest that imagining social contexts may be important for affective recognition more broadly, rather than for recognition of specific kinds of emotional expressivity.

The importance of social aspects of emotional communication were also reflected within the nature of the discussions themselves. Many of the participants made comments suggesting that they were experiencing difficulty in communicating their ideas with the interviewer (e.g., "how do I say this?", "it's hard to put into words"). Such instances appeared to be more prevalent among the dancer participants, but several non-dancers also noted similar struggles. On a number of occasions, when struggling to verbalise their thought processes, participants would resort to "acting out" certain scenarios to clarify the meaning of their comments. For example, when struggling to communicate why they were labelling a movement sequence as angry, one participant adopted the role of two imagined others navigating a situation where one of these others felt "hurt" during a confrontation (i.e., "you've done this, you've hurt me"). In linguistics research, this has been referred to as providing quotations (Sams, 2010; Clark & Gerrig, 1990). It has been suggested that these quotations function as thought experiments, allowing the speaker to present something they believe in the hope that the listener shares their understanding, and can offer their own interpretation or thoughts on the relevant assumption or question (Myers, 1999). In the context of this experiment, it is possible that the participants were highlighting these social examples as a way to resolve difficulties in verbalising their ideas through the presentation of a scenario through which the interviewer and interviewee can evaluate a shared point of reference. If this notion holds weight, it implies a perceived mutual understanding of social factors involved in the experience of emotion, and provides further evidence for the importance of considering emotions within a social context. This idea, however, is rooted in linguistics and philosophy, so it is unclear how quotations function within the context of social psychology research. Therefore, future research in this area may be beneficial for developing a picture of the extent to which mutual understanding of social situations contributes to the success of social communication.

2.3. Kinematics in Emotion Recognition

It has already been noted that in the study of facial expressions, researchers have identified unique patterns of muscle activity ('action units') that correspond to the expression of particular emotional states. This approach is mirrored in social perception research exploring affective expressivity through whole body movement cues, as a number of researchers have attempted to isolate unique patterns of movement kinematics that correspond to the expression of different emotional states. However, in review of the existing literature, compared to the consensus in which action units corresponds to which emotions in facial expressions (e.g., happiness involves raising the orbicularis oculi and pars orbitalis, and the zygomaticus major, to scrunch the eyes and smile, respectively), researchers provide contradictory accounts of which human movement kinematics result in the communication of these same emotional states. For example, sadness has been associated with both contracted (Shikanai et al., 2013) and limp (Shafir et al., 2016) motion, and happiness has been associated with both relaxed (Gross et al., 2010) and high energy (Brownlow & Dixon, 1997) movement. These inconsistencies are further exacerbated by the lack of unity in which emotions are studied within this context. For example, some studies explore the emotion of 'grief' or 'despair' in place of 'sadness' when studying the basic emotions (Camurri et al., 2002), while others have considered sadness and such emotions separately (Wallbott, 1998). This is even more pronounced in the exploration of 'happiness', perhaps to accommodate for the lack of positively-valenced emotions typically included in these experiments, as many researchers choose to forgo the inclusion of 'happiness' completely, in favour of 'contentedness', 'pleasure', or 'joy' instead (Shikanai et al., 2013; Masuda & Kato, 2010; Castellano et al., 2007; Crane & Gross, 2013).

By a clear margin, the most consistently reported result in this field is that speed allows observers to differentiate between high and low intensity emotions (Smith & Pollick, forthcoming; Halvoic & Kroos, 2018; Gross et al., 2012; Roether et al., 2009). Therefore, from a quantitative perspective, it would appear that speed of movement plays a pivotal role in emotion recognition. However, it is unclear whether this finding is of conscious perceptual importance to observers as few researchers have directly asked observers what aspects of movement are important for discriminating between different emotional states. We addressed this gap in the literature in the present experiment and indeed, we found that speed was important; as all of the participants mentioned speed of motion at some point during the emotion recognition task. These results suggest that speed is important from both a quantitative and perceptual perspective in discriminating between different emotional states communicated through whole-body movement of the human body. It may then be important for any framework for whole-body affective signalling to include movement speed as a key component of emotion recognition. The importance of accounting for movement speed in such a framework may be of additional importance when we consider the potential role of aesthetic evaluation in affective decision-making (a relationship that will be given further consideration in the following section). In a recent study conducted by Orlandi and colleagues (2020), participants observed a series of dance sequences, which varied in speed but not choreographic content, before providing judgements of movement reproducibility, speed, effort and aesthetic value. They found that faster movements were judged to be more enjoyable than those with slower, more uniform velocity profiles (Orlandi et al., 2020). Therefore, speed may too play a role in aesthetic evaluation of human movement stimuli, which may in turn impact the emotional expressivity assigned to those movements. Considered together, these results provide evidence for the inclusion of this kinematic property within a framework for whole-body social signalling.

2.4. The Role of Aesthetics in Emotion Recognition

As we have alluded to in the previous section, research suggests that affective and aesthetic processing may be related. This has been observed in the evaluation of paintings (Gernot et al., 2018; Marković, 2010; Jankowski et al., 2020), music (Sachs et al., 2015; Brattico et al., 2016; van den Tol et al., 2016), and more recently in the evaluation of whole-body dance movements (Christensen et al., 2021; Smith & Cross, under review). Additionally, the elicitation processes involved in the production of both affective and aesthetic experiences bear a number of strikingly common features (see Chapter 1 for a more thorough discussion of the similarities). Despite the wealth of evidence for a link between these processes, there is a lack of research directly exploring the impact of aesthetic evaluation on the emotional expressivity assigned to such stimuli, as typically this relationship is explored in the inverse; determining whether emotional expressivity influences the aesthetic value assigned to such stimuli. However, in this experiment, we obtained evidence to suggest that aesthetic processing did influence the emotional expressivity assigned to human movement stimuli in this experiment.

During this experiment, participants were not asked to provide aesthetic evaluation of the McNorm Library movement clips, and none of the questions asked in the post-task interview made any reference to aesthetic parameters. Despite the absence of aesthetic considerations within the experimental proceedings, the majority of the participants provided unprompted aesthetic value judgements (e.g., beauty, liking, interest) for the whole-body movement sequences. This, therefore, provides further evidence that affective and aesthetic processing of whole-body dance movements may be intrinsically linked.

One participant discussed a more specific example of aesthetic evaluation directly influencing their emotion judgements for the stimuli. This dancer participant noted that they struggled to assign the label of ‘anger’ to a movement sequence, despite believing it to be the ‘correct’ answer, because they could not reconcile the expression of that negative emotion with something that was beautiful. While this was the only excerpt where a participant explicitly mentioned the direct influence of aesthetic parameters on the emotional expressivity they assigned to the movement sequences, a wealth of additional evidence for this idea can be found when examining the aesthetic value judgements other participants assigned to the movement sequences, and the emotion judgments that followed. It should be noted that the vast majority of the aesthetic judgements were positive (there was only one mention of 'boring' and 'dislike'), and the subsequent decision in the emotion recognition task was typically that of 'happiness'. This finding is unsurprising when we consider previous research in support of the idea that 'what is beautiful is good'. This perspective can be traced back to the great philosophers of ancient Greece (e.g., Sappho, Plato; Dayan, 2011), and more recent evidence from facial attractiveness research also support this notion (Little et al., 2006; Lorenzo et al., 2010). It is implied that perhaps the spontaneous, positive aesthetic evaluation (i.e., what is beautiful) of the McNorm Library clips in this experiment may have guided participants to assign positive affective labels (i.e., the positive valence, or good, label of 'happiness') to those same clips.

As such specific comments about aesthetic evaluation interfering with emotion judgements were made by only one participant, and the more frequent behaviour of assigning happiness to clips perceived to be beautiful, likeable or interesting provides only indirect evidence of this relationship, these examples should be considered carefully. However, they do add to growing bank of evidence suggesting that affective and aesthetic evaluation of human movement may be intertwined, and reciprocal.

2.5. Kinaesthetic Empathy

The role of kinaesthetic empathy has already been discussed in relation to the quantitative emotion recognition rates obtained in this experiment, but kinaesthetic empathy was also a key component of the qualitative interview discussions.

Previous research has observed that both long-term and short-term dance training can result in neurophysiological changes within brain structures involved in the performance, observation and simulation of movement (the AON; Cross et al., 2009), and those involved in empathic processes (Neal & Kilner, 2010; Avenanti et al., 2013; Nummenmaa et al., 2014; Keysers & Gazzola, 2006; Caspers et al., 2010). From such results, it has been suggested that individuals with previous dance training experience greater motor simulation when observing movement (evident from greater activity in the AON), and that this greater degree of kinaesthetic empathy may be responsible for the improved performance in emotion recognition tasks observed among dance experienced samples (Christensen et al., 2016; 2019; 2021). However, as noted earlier in this manuscript, such conclusions rely on extrapolating the rationale from findings of studies testing similar, but empirically distinct, research questions. To our knowledge, this experiment is the first to explicitly ask observers whether they made use of kinaesthetic empathy during an emotion recognition task, specifically.

Through the interviews, we observed several findings which are relevant to the evaluation of the use of kinaesthetic empathy in observing and judging the emotional expressivity of whole-body movement sequences. First, we obtained evidence to suggest that while kinaesthetic empathy was used by the majority of the dancer participants during the emotion recognition task, this was not universal across the sample. Three out of the ten dancer participants noted that they did not imagine themselves executing the movements while they were watching them. There are several factors to consider when dissecting this finding. First, it should be noted that one of these dancers who did not report using kinaesthetic empathy was male. Classical ballet is deeply rooted in strict choreographic rules and tradition, which is reflected in gender roles typically employed in the choreography of performative ballets, and in turn is emphasised in the diverging movement repertoires taught to male and female dancers. For example, it is typically only female ballet dancers who perform movements *en pointe* (dancing on the tip of the toes, aided by a specific kind of shoe reinforced with wooden blocks), the act of which facilitates a particular set of dance movements, and it is typically only male ballet dancers who perform lifts during choreographic sequences (Oliver & Risner, 2017; Angyal, 2021). While gender-based choreographic repertoires may not be so evident in different dance styles, the McNorm Library clips for the most part contained Western Classical Ballet movements from a female dancer's repertoire. It is possible that this male dancer participant did not make use of kinaesthetic empathy when observing the clips in this experiment because they did not have physical experience of performing those movements themselves, but may make use of kinaesthetic empathy when observing dance movements from a male ballet dancer's repertoire.

Further, many of the dancer participants noted being familiar with more than one dance style, and were not required to be most familiar with Western classical ballet to be included within this study. Therefore, the observers may not have had experience performing the specific movements included within the McNorm Library, potentially limiting their ability to simulate the movements within their own bodies. It is possible that kinaesthetic empathy may have been used more ubiquitously if the stimuli included movements from a variety of different dance styles. Reviewing these findings, perhaps it would be useful to account for such dance style-specific considerations when creating future movement stimuli, by including sequences from the repertoires of both male and female dancers (when using western classical ballet), and expanding the breadth of dance styles captured within the resulting movement repository.

Secondly, it was observed that the use of kinaesthetic empathy was not exclusive to dancers, as six out of the ten non-dancers who participated in this experiment also reported that they imagined the movements within their own bodies to assist them in assigning emotion labels to the stimuli. Finding that the majority of the non-dancer participants reported using kinaesthetic empathy is surprising, as previous research has suggested that kinaesthetic empathy has an expertise-specific influence on the processing of human movement stimuli (Reason & Reynolds, 2010; Behrends et al., 2016; Seifert, 2018). The evidence provided in this experiment, however, indicates that kinaesthetic empathy may be used more extensively, and by a less specific subset of observers, in the evaluation of emotionally expressive whole-body human movement. Given these results, it is likely that additional factors (e.g., previous experience or personality traits), other than previous motor or observational dance expertise, contribute to whether an individual imagines simulating movement within their own bodies when observing another in motion. However, as a result of the lack of research exploring this possibility, these factors are currently unclear and merit due consideration in future work.

2.6. Previous Experience

The role of previous experience in emotion recognition tasks using human movement tends to be discussed specifically through the lens of motor experience (e.g., dance training), and with reference to the concept of kinaesthetic empathy or motor simulation. However, the findings highlighted in the previous section suggest that there are likely to be additional aspects of an individual's previous experiences that can influence the evaluation of human movement. Several of these factors were raised by participants during the semi-structured interviews, and each is discussed in more detail in the following sections.

2.6.1. Learned Associations from Dance

A number of dancer participants noted that they held associations between specific movements and the expression of particular emotions based on what they had learned during their dance training. The most common example of this noted by the dancer participants was an association between certain movements, like plies for example, and the absence of expressivity due to prior experience of performing those movements primarily in the context of warm-up exercises, rather than performances. This specific association was noted by seven out of the ten dancer participants, and this provides evidence to suggest that certain movements may elicit a particular emotional context among dance-experienced samples.

When creating the McNorm Library, we aimed to develop a series of movement sequences which were devoid of emotionally-salient gestures (e.g., shaking fist in anger, throwing arms into the air in joy), and did not include movements which were associated with a narrative from any known performative ballet (e.g., the "dying swan" arm movements which characterise the grief felt by Odette in Swan Lake); to allow the dancer to express emotion through performative qualities of the movement, rather than through emotion associations with specific movements (Smith & Cross, 2022). However, we did not consider the possibility that dancers may have associated movements with the absence of expression (e.g., neutrality), and that this may provide additional assistance in an emotion recognition task using the McNorm Library Stimuli. It is unclear whether there are other movements that elicit specific kinds of emotional expressivity in performative dance, and this possibility should be addressed in future research. For example, it may be useful to conduct follow up interviews with a larger group of dancers to identify any similarly common movements that are linked to the expression of a particular emotion, or to the lack of expressivity. Further, it may also be useful to invite choreographers to participate in these discussions, to address this same question from the alternative perspective of the choreographic process.

Several participants also likened certain movements contained within the McNorm Library clips to movements they had observed in performative ballets, including *Swan Lake*, *Giselle* and the *Nutcracker*. While most of the participants who made such comments were dancers, three had no previous dance training. This provides further evidence to support the idea that observers may draw upon observational experience of dance to assist them with assigning emotional expressivity to whole-body human movement. Behavioural and neurophysiological changes associated with watching dance have been noted in a number of studies (Cross et al., 2006; Kirsch et al., 2015; Cross et al., 2011; Bläsing et al., 2012; Orgs et al., 2018), but here we obtained direct evidence to indicate that awareness of movements included in performative dance stories can influence the lens through which observers perceive social aspects of others in motion.

2.6.2. Learned Associations from Real-World Social Experiences

A number of participants in this experiment mentioned that the movements contained within the McNorm clips conjured memories of specific social situations they had experienced in their personal lives, which they used to assign emotional expressivity to the movements. The majority of the examples were associated with negative emotions, particularly of times when participants had been engaged in some type of confrontation (e.g., argument, workplace squabble), but on other occasions participants noted that the movements reminded them of celebrations or positive interactions (e.g., parties, seeing friends). It is interesting that the vast majority of these comments arose when participants were assigning the labels of anger or happiness, respectively, as these emotions are often confused with one another in emotion recognition tasks (both in this experiment, and in the previous literature; Smith & Cross, 2022; Badshah et al., 2017; Kadiri et al., 2015; Kuchibhotla et al., 2014; Emerich et al., 2009). It is possible that successful discrimination between happiness and anger necessitates stimuli that depict the emotionally expressive movements of more than one performer interacting with one another. Perhaps the expression of happiness and anger are more dependent on the response of an interaction partner to comprehend the social context, and ultimately assign the correct emotion judgement. Future research could explore this possibility by creating movement stimuli which depict the interactive motion of two, or more, performers. Submitting this kind of stimuli to assessment in emotion recognition tasks may clarify on the relative importance of social interaction in the recognition of specific kinds of expressivity.

2.6.3. Learned Associations from Popular Culture

Finally, a number of participants referenced things from popular culture when evaluating the stimuli clips, noting that elements of the movements reminded them of things they had witnessed in films, or on television; some of which centred around dance (e.g., Strictly Ballroom, Strictly Come Dancing).

These comments imply that perhaps asking participants how often they observe dance in a formal performance setting may not provide an adequate picture of their observational experience of dance. Additionally, during the Covid-19 pandemic, a number of theatre and dance companies began broadcasting their performances online, making the consumption of performative dance easier than ever before. It is possible that our current understanding of observational experience may need to be adjusted to account for the range of ways in which arts consumers can now access and interact with dance, as existing measures may not capture an accurate picture of an individual's experience with dance in various forms. This may be worth considering in future studies exploring the role of observational experience in social processing of emotionally expressive whole-body movements, particularly in post-pandemic world.

However, not all popular culture references were associated with dance, specifically. For example, one participant imagined music from the animated film Toy Story playing when they observed one of the movement sequences. They noted that the movements made them vividly imagine specific upsetting scenes from the film, and that this guided them to assign the sad label to the clip. To our knowledge, no studies exploring affective evaluation of human dance movement have included any measures of engagement with other kinds of media (e.g., film, television, theatre, musicals). It is possible that exposure to certain emotionally-salient media, and individual differences in sensitivity to such media, may influence behaviour during emotion recognition tasks. The idea that engagement with different kinds of media can influence behaviour in social psychology experiments is already given consideration in artificial intelligence research. In a number of experiments in the field of social robotics, researchers have chosen to include questionnaires which allow participants to provide an overview of the context in which they have engaged with robots and virtual and augmented reality, for example, as it has been found that this can influence the ways in which participants can engage with such technology (Riek et al., 2011; Sundar et al., 2016; Saffari et al., 2021). In responding to such measures, participants can indicate which popular science fiction films (e.g., Ex Machina, Wall-E), television programmes (e.g., Westworld, Black Mirror), and fiction novels (e.g., 2001: A Space Odyssey, Do Androids Dream of Electric Sheep) they have watched, or read, and how this has influenced their attitudes towards various kinds of AI. Extrapolating this rationale, it may be useful in affective research to explore exposure and sensitivity to emotion-inducing media in participants to understand whether engaging with such media can influence the success with which they can assign expressivity judgements to emotionally-salient stimuli.

CONCLUSIONS

In this study quantitative results showed that participants recognised emotional expressivity from the McNorm Library movement clips with a similar level of accuracy that was reported in the original validation experiments, thus providing further evidence for their viability in social perception research. However, contrary to the previous literature, in this experiment we did not find evidence to support the idea that dancers perform better in emotion recognition tasks than non-dancers. Further, while we did not obtain evidence to suggest that kinaesthetic empathy assisted with recognising the expression of specific emotional states, the results suggest that mentally simulating the movements in the McNorm clips assisted participants in discriminating between emotionally expressive, and non-expressive whole-body movement.

Through a series of semi-structured interviews that composed the qualitative component of this study, we discovered a number of factors that influenced decision-making during the emotion recognition task. First, we observed that a number of the participants experienced difficulty in deciding what constituted an emotion within the context of the task. Second, we found that many of the participants mentioned elements of their perceptual processes that were related in some way to the basic James-Lange, Cannon-Bard, and Schachter and Singer theories of emotion. More specifically, despite the reduced visual format of the McNorm Library clips, many of the participants noted that they visualised the image of some form of environmental trigger that the dancer was responding to when making their decisions, and this appeared to be particularly salient for the recognition of fear. Additionally, we have previously suggested that failing to account for the social nature of emotions in real-world settings is a criticism of these basic emotion theories. Comments made by participants in this experiment provide further evidence that the social context of emotions is important for the successful communication and interpretation of emotional expressivity. Thirdly, we noted that a number of researchers have attempted to identify unique profiles of movement kinematics that correspond to the expression of different emotional states, with mixed results. We obtained a substantial amount of evidence from the participants in this experiment to suggest that movement speed plays a vital role in discriminating between different kinds of expressivity from the motion of others.

Another factor found to be important for decision making in this emotion recognition task was the role of aesthetic evaluation. Previous research has suggested that affective and aesthetic processing of whole-body human movement may be intrinsically linked, and we obtained further support for this notion as almost all of the participants in this experiment provided spontaneous, unprompted aesthetic value judgements for the movements in the McNorm Library clips during the experiment.

Finally, it has been suggested that previous dance experience facilitates emotion recognition through kinaesthetic empathy, or motor simulation. However, during these interviews, we observed that the use of kinaesthetic empathy was neither universal, nor unique, to dancers when assigning emotional expressivity to the dance movement sequences. Several of the dancer participants explicitly stated that they did not imagine simulating the movements within their own bodies, and the majority of the non-dancer participants did report doing so to assist them with their decision making during the task. Further, we found that there were other elements of previous experience that contributed to the decision-making processes employed by participants in this emotion recognition task, some of which were related to dance experience, and others which related to real-life experiences more broadly. We found that dancers made use of learned associations from their training to assist them in assigning emotion labels to the movements. Additionally, both dancers and non-dancers relied on their observational experience of dance to guide their decisions, and this occurred through likening movements in the stimuli to choreographic sequences in performative ballets, and to dance performances observed in more informal settings; such as those depicted on television and in films. Finally, it was also found that memories of "real-world" social experiences permeated the emotion judgements made in response to the stimuli in this experiment. Participants reported comparing certain elements of the observed movements to their own movements during instances of social engagement with others, in both positive and negative contexts.

In sum, we believe this study has provided novel perspectives on the individual differences that can sculpt decision-making in emotion recognition tasks, and these results provide several avenues for future research in this domain.

CHAPTER 6

GENERAL DISCUSSION

GENERAL DISCUSSION

1. Summary of Thesis Findings

The overall aim of this present thesis was to explore the factors influencing emotion recognition from expressive whole-body movement and to suggest a new framework for how emotional expressivity is interpreted from the human body in motion. In addressing these goals, a new repository of expressive whole-body dance movements (The McNorm Library; Smith & Cross, 2022) was created, and a series of quantitative and qualitative experiments were conducted to examine observer evaluations of these movements (in terms of emotional and aesthetic content), and the factors which may have sculpted these social judgements (i.e., individual differences in empathy, and previous dance experience). In summary of the main experimental findings, there are several important results to mention here. First, through several replication experiments, movement clips from the McNorm Library were found to be recognised with a consistent accuracy of around 50% by 147 observers. Second, aesthetic value judgements assigned to these same movements showed that the presence of emotional expressivity had a significant impact on how beautiful observers found the movements to be, and how much they liked them; suggesting a link between affective and aesthetic processing of human dance movement. Finally, through the use of semi-structured interviews, we found that both physical and observational dance experience sculpts the lens through which observers evaluate the choreographic content of dance sequences, and we obtained evidence to suggest that kinaesthetic empathy is not uniquely nor universally employed by those with previous movement training during emotion recognition tasks.

In this general discussion, I will begin by summarising the aims, predictions and results from each of the empirical chapters included within this thesis. I will then situate results from these empirical chapters within the context of theoretical contributions from the previous literature, before suggesting a potential framework for emotion communication through the movement of the human body.

1.1. The McNorm Library: Creating and Validating a New Library of Emotionally Expressive Whole-Body Dance Movements

In Chapter 2, I describe the creation of a new library of emotionally expressive whole-body dance movement sequences, named the McNorm Movement Library. One of the primary aims behind developing this movement library was to address several of the key limitations associated with existing movement repositories. These movement stimuli were validated in two experiments, and we found that participants could reliably discriminate the intended emotion for a subset of these clips with accuracy rates of more than 50% (where chance is 20%). Across both experiments, observers recognised non-expressive and sad movement sequences with the greatest accuracy, and movements conveying fear were consistently recognised least accurately. We provided several potential explanations for the relatively poor recognition of fear. First, we suggest that fear may be the most idiosyncratic of the emotions explored in this experiment; due to the various ways in which it can manifest physically within the human body (i.e., fight, flight, or freeze). Alternatively, the basic emotion theories (i.e., James-Lange, Cannon-Bard, Schachter & Singer) imply the need for some external trigger to induce the feeling of fear. For obvious reasons, it was not possible to introduce a genuine threat into the environment when generating the McNorm Library clips, and this lack of a stimulus to react to may have produced an inauthentic expression of fear; which evidence from facial expression research suggests observers may be particularly sensitivity to.

Through these validation experiments, we also explored the potential influence of previous dance experience and trait empathy, but found that neither had a significant impact on accuracy in the emotion recognition task. This result was surprising, considering the weight of previous literature suggesting the mediating role empathy and experience play in sculpting perceptions of human movement. However, in discussing these results, we note that the majority of existing work makes use of highly experienced dancer participants, and only a very small percentage of our participant samples classified themselves as such. Indeed, the majority of our dancer participants classified themselves as "Beginner" or "Intermediate" dancers. Less is known about the role of amateur, or hobby, dance training in sculpting the social evaluation of emotionally expressive whole-body movement, but these results suggest that specific aspects of formal dance training may be required to significantly change the way we perceive others in motion.

1.2. Moving Me, Moving You: Emotional Expressivity, Empathy, and Prior Experience Shape Whole-Body Movement Preferences

In Chapter 3, we sought to examine how individual differences in empathy and previous experience influence aesthetic evaluations of dance movements communicating different kinds of emotional expressivity. In this experiment, observers completed a self-report measure of empathy, provided an overview of their previous dance experience, and rated 20 videos of emotionally expressive whole-body movements from the McNorm Library (Smith & Cross, 2022) in terms of beauty and liking. We found that observers preferred emotionally expressive over non-expressive movement. In addition, specific types of emotional expressivity influenced liking, but not beauty, judgements. More specifically, observers showed a preference for angry and happy movement sequences over sad sequences, in terms of their liking ratings. Finally, our finding revealed that both a greater degree of previous dance experience and higher trait empathy resulted in increased beauty and liking ratings for the McNorm Library dance clips. We conclude that emotional expressivity plays a significant role in the aesthetic appraisal of whole-body human movement and this may suggest that these evaluation processes may be intrinsically linked.

1.3. Read Between the Lines: Behavioural and Computational Exploration of Factors Involved in the Emotion and Aesthetic Evaluation of Motion-Generated Continuous Line Drawing

In this experiment, we explored whether emotionally expressivity from whole-body human motion could be recognised when that motion was represented as visual images of the movement trajectories. To this end, we created a series of continuous line drawings using the movement trajectories of the wrists and ankles of the dancer who performed the emotionally expressive whole-body movements contained within the McNorm Library. In an emotion recognition experiment, we examined whether these drawings could function as a viable proxy for human motion, but found that recognition of the target emotions (neutral, happy, sad, angry, and fearful) was significantly impoverished compared to point-light display representations of those movements, and recognition was not significantly greater than chance.

In a second experiment, we explored the possibility that these drawings could function as visual artworks in their own right, in isolation from their movement origins. Previous image analysis research has suggested that observers prefer images that are larger in size (i.e., the "bigger is better" hypothesis), which have contours that are more rounded than jagged, and that are average in terms of complexity. This research also suggests that the factors of size, curvature, and complexity play specific roles in the affective ratings assigned to various kinds of visual art (e.g., paintings, simple polygons). In this experiment, when considering these trajectory drawings as visual artworks instead of alternative ways to visualise human movement, we obtained evidence to support a general preference for curved contours, and an inverse U-shaped preference for complexity. However, we did not observe a general preference for larger images.

From these results, we conclude that while this stimulus creation method did not produce a viable alternative to existing human movement stimuli, it did provide a novel method for generating novel visual artworks that might be of use or interest for future empirical aesthetics research. And, perhaps the most important contribution from this chapter, we also propose a new method for calculating image complexity. Currently, a universal operational definition of "image complexity" does not exist, and existing methods rely on specific kinds of stimuli (e.g., images containing greyscale, images of simple polygons such as circles and triangles), or require some degree of manual labour which inflates the potential for human-error (e.g., manually counting the number of turns in an image). Our measure of "image complexity" is an aggregated factor which comprises the output of several fully-automated methods for calculating complexity and, to our knowledge, is the first of its kind to be suggested for computational image analysis research.

1.4. "A Little More Conversation": A Qualitative Assessment of Decision-Making by Dancers and Non-Dancers During a Whole-Body Movement Emotion Recognition Task

Forced-choice tasks are used extensively in social cognition research, but such "button-pressing" tasks have often (and rightly) been accused of being too reductive to accurately reflect the complexities of human affective communication. Further, the individual differences that can sculpt the success of emotion recognition (e.g., trait empathy, kinaesthetic empathy, dance experience) are typically measured using similarly reductive quantitative methods (e.g., a single empathy quotient score, single value for "years of experience"). An over-reliance on these kinds of measurements can limit the scope of information we can obtain about whole-body social signalling. In the experiment described in Chapter 4, we explored how observers assign judgements of happiness, sadness, anger, fear and neutrality to whole-body movement sequences, through a mixed methods experiment. In this task, 10 dancers and 10 non-dancers observed emotionally expressive whole-body dance movement sequences from the McNorm Library, assigned emotion labels to each of the videos, and were asked a series of interview questions to probe the reasons for their decisions. Through these interviews, we observed that the imagination of contextual cues (i.e., environmental triggers, and imagined interaction partners), speed of motion, aesthetic evaluation (e.g., beauty, liking, interest), and individual differences in kinaesthetic empathy and previous experiences had contributed substantially to the emotion judgements made by observers. Through the work presented in this chapter, we highlight a hybrid approach to the study of decision-making in emotion recognition tasks that combines simple quantitative with richly qualitative design elements, with results suggesting a new perspective on the factors that may sculpt social perception in this domain.

2. *Contribution of the Work*

In the following section, I discuss the main contributions of this thesis to the fields of social perception, emotion science, and empirical aesthetics.

2.1. The McNorm Library

The work documented in this thesis provides several contributions to the field of social perception. First, the McNorm Library provides an open-access repository of emotionally expressive whole-body dance movement sequences for use in future work which was designed to resolve several issues present within existing libraries. For example, such previous libraries generally do not account for the dyadic nature of social communication, as emotion labels are sometimes assigned to items in these sets without accounting for the expressive intentions of the movement performer (de Meijer, 1989; Michalak et al., 2009; Melzer et al., 2019; Christensen et al., 2014; Christensen et al., 2016; Smith & Pollick, forthcoming). Additionally, existing libraries have tended to depict certain emotions through the use of specific, distinct sets of movements. Others have included culturally or socially-salient gestural cues (e.g., shaking the fist in anger, throwing arms in the air in joy – Atkinson et al., 2004; Wallbott, 1989; Montepare et al., 1999; Pasch & Poppe, 2007; Dael et al., 2012; Shafir et al., 2016), or movements from performative ballets which bring to mind emotionally-charged narratives, particularly in dance-experienced samples (e.g., "dying swan" arm movements associated with the grief of *Swan Lake's* protagonist Odette – Christensen et al., 2014; Christensen et al., 2016; Smith & Pollick, forthcoming). Finally, previous movement libraries which depict movement in the form of full-light displays (e.g., Atkinson et al., 2004; Montepare et al., 1999; Shafir et al., 2016; Melzer et al., 2019; Christensen et al., 2014; Christensen et al., 2016) contain a wealth of extraneous visual information which may render them unsuitable for testing certain questions in the field of social perception (e.g., visual attention research).

In generating the McNorm Library, I addressed these issues in several ways. First, I ensured that the movement sequences created for this library were devoid of both gestures and choreographies from performative ballet. Second, I generated five versions of each movement sequence, wherein the specific choreography was maintained but the dancer performed the movements with different kinds of expressivity (neutral/non-expressive, happiness, sadness, anger, and fear). Doing so allowed us to account for the perspective of the movement performer in the creation of emotionally-relevant stimuli (to better account for the dyadic nature of affective communication), and provided the opportunity to explore how movement changed when imbued with different kinds of expressivity, without necessitating the performance of specific, emotionally-relevant movements. Finally, I chose to depict the dancer's movement in the form of point-light, rather than full-light, displays in order to create a movement library that would be suitable for use in specific types of psychophysiological research (i.e., eye-tracking).

A further benefit of the McNorm Library lies in its test-retest reliability. Rarely are movement libraries subjected to such rigorous validation, but for this library recognition of the intended target emotion was explored in 127 participants across the two original validation experiments, and by a further 20 in the final mixed-methods experiment. Average recognition rates obtained across these three experiments were 52.2%, 49%, and 52%, respectively. These results show the extremely high degree of consistency (variance of 3.2%) in recognition of the target emotions from the emotionally expressive whole-body movement stimuli contained within the McNorm Library, and provide additional evidence for its value in social perception research.

2.2. The Benefits of Re-Using Research Data

Due to disruptions to research caused by the global Covid-19 pandemic, restrictions in access to laboratory equipment and space led to a fairly drastic restructuring of the shape and scope of this thesis in 2021. In order to adapt to new ways of conducting experiments remotely, and a resultant inability to generate any new stimuli, I decided to reuse the McNorm Library data in a number of additional experiments.

From the early noughties, there has been a growing interest in finding ways to responsibly reuse research data, particularly in STEM fields (Piwowar, 2011). Reuse of existing data offers a number of notable benefits. First, reusing data increases the efficiency of research, in terms of both time and resources (Boté & Térmens, 2019). Creating new stimuli or data repositories can require extensive access to specific equipment, facilities, and specialists; all of which can add significant costs and delays to the timescales of research projects, and consume large portions of allotted research funding. In capturing the movements included in the McNorm Library I required several days of uninterrupted access to the Motion Capture Laboratory within the School of Psychology at the University of Glasgow, and several days of input from a professional dancer and choreographer who was financially reimbursed at an appropriate rate to reflect their expertise. After the recording sessions, a significant portion of time was required to manually label joint-markers within the Vicon Nexus software, and to fill gaps within the recorded motion capture data (that resulted from technical issues). Therefore, the process involved in the creation of this movement repository involved extensive access requirements, funding, and time, and similar demands are to be expected for the creation of any such stimulus library. These practicalities alone demonstrate the value of reusing existing stimuli sets and data, but there are additional benefits to doing so.

Reusing data also allows researchers to test the consistency of findings obtained from studies using the same data. It has already been noted in the previous section that emotion recognition rates obtained from three validation experiments using the McNorm Library movement clips were remarkably consistent, and rarely are such movement repositories subjected to such rigorous retesting. By using the same stimuli in multiple experiments in this way, we can gain greater clarity about the robustness of research findings, which will ultimately strengthen the conclusions we can draw from empirical studies. This may be of particular relevance to studies exploring social perception of whole-body human movement, as the findings in this area vary widely in terms of reported recognition rates (varying from 23% to 100% depending on the specific emotion category; Atkinson et al., 2004; Crane & Gross, 2013; Gross et al., 2010; Roether et al., 2009) and kinematic profiles associated with different emotional states (Sawada et al., 2003; Wallbott, 1998; Gross et al., 2010; Brownlow et al., 1997; Camurri et al., 2002), and the specific methodologies employed (i.e., in terms of stimulus type – point-light displays, full-light displays, stylised avatars; movement styles – dance sequences, walking, knocking, free movement; and recognition techniques – automated/machine learning, spectator judgements).

A further benefit of reusing the McNorm Library clips was that it encouraged me to be creative in my approach to studying the evaluation of emotional expressivity, and aesthetic value, of whole-body human movement. Disruptions caused by the Covid-19 pandemic forced me to explore novel methods (e.g., qualitative research, computational image analysis) that I had not considered using at the beginning of my PhD. While there would have been nothing inherently wrong with performing only a series of online questionnaire experiments, I felt that doing so would have limited the breadth of my research experience and impacted my future employment opportunities.

Finally, reusing data in this way allows me to cite the McNorm Library in a number of additional publications, hopefully bringing more attention to my research output which is beneficial to me as an early career researcher, and to my funders who wish to see "bang for their buck". By using the McNorm

Library in several experiments, we have brought additional awareness of this repository to researchers in the fields of empirical aesthetics, computing science, and qualitative research methods – in addition to those who work in emotion perception, as was the original intention of this work.

2.3. Development of Novel Research Methods

At various points throughout the thesis, new methods and approaches were proposed which could be employed in future social perception research. First, in validating the McNorm Library, recognition rates obtained in the first experiment were lower than expected based on the previous literature. The average overall recognition, across all 73 clips in the full McNorm Library, was around 30%. While this did meet the threshold for significance, indicating that recognition was better than would be obtained by chance (20%), this indicated that perhaps some movement clips were not as representative of their target emotion as others. Indeed, examining recognition rates for each item in the stimuli set there were significant inconsistencies across the full range of stimuli (some were never assigned the target emotion label, while others were recognised with accuracy of greater than 65%). Based on these inconsistencies, we decided to conduct a follow-up experiment, wherein we presented the 20 clips with the highest recognition accuracy (4 each; neutral, happy, sad, angry, fearful) to a new group of participants, to examine whether this subset of clips was more representative. In this second validation experiment, the overall average recognition rate (49%) was extremely similar to the average recognition rate obtained for that same subset in the first experiment (52.2%). To our knowledge, no other researchers have employed this method of condensing stimuli sets to include only those that are most "successful", and therefore most useful for future research. We suggest selecting a subset of the original stimuli set and submitting it to further testing represents a new solution for other researchers who also obtain relatively low recognition rates for their full movement repositories.

Secondly, in Chapter 4 of this thesis, we propose a novel method for calculating image complexity. It has been argued that there is a problematic fluidity in what constitutes "complexity" in visual aesthetics research. Researchers have proposed a number of computational methods for calculating complexity, but these measures vary widely in scope, and often require extensive manual effort (which increases the likelihood of human-error) or specific kinds of stimuli for those methods to be applicable. To our knowledge, no researchers have suggested a robust, fully-automated method for calculating complexity that can be applied to non-traditional images (i.e., not paintings, greyscale images, or simple shapes like circles or triangles). In the fourth thesis chapter, we describe the development of a fully-automated method for calculating complexity present within abstract black and white images (without any shade gradation) that produces aesthetic evaluations in the expected direction (i.e., an inverse U-shaped preference profile). As the factor we describe involves averaging across several possible elements of complexity, this provides some allowances for the problematic fluidity in defining computational complexity, and permits researchers to include their own measures of complexity into the aggregated factor. It is hoped that despite the lack of an operational definition of complexity in the context of computational image analysis, using such an aggregated factor allows us to generate a more robust measure that captures the idiosyncrasies of what makes an image "complex".

Finally, the experiment presented in Chapter 5 presents a novel method for exploring factors influencing decision-making during an emotion recognition task. Results from a number of quantitative experiments have explored emotion recognition from movement through the use of force-choice tasks, and have suggested factors that may contribute to this kind of affective decision-making based on the statistical analysis of neuroscientific and behavioural data. However, to my knowledge, no previous research has explored what features of movement, or aspects of previous experiences, are perceptually important to observers through combined quantitative and qualitative approaches. In employing mixed methods when addressing this question, we have gained a more detailed understanding of what is perceptually important for the recognition of emotion from the movement of others, and a number of new avenues for future research have emerged.

To our knowledge, this is also the first time observers from various backgrounds have been surveyed to understand the extent to which they believe kinaesthetic empathy (or motor simulation) is important for understanding emotional expressivity displayed by others. Previous neuroscientific evidence has suggested that individuals with previous motor experience (i.e., dancers) make greater use of kinaesthetic empathy than those without such training, implied by greater activation observed in brain areas of the Action Observation Network in these individuals (Burzynska et al., 2017; Cross et al., 2009; Cross, 2010; Gardner et al., 2015; Kirsch et al., 2018; Liew et al., 2013). However, in this experiment, we found that this use of kinaesthetic empathy was not universal, nor exclusive, to expert dancer participants. Specifically, we found that, at least in our sample, several expert dancers did not report imagining themselves performing the movements they were observing, while the majority of non-dancers did. This result is extremely surprising when considering the previous research, and suggests that perhaps the assumption that kinaesthetic empathy results for the most part from extensive movement training should be revisited in future work. It is possible that both dancers and non-dancers do use kinaesthetic empathy, but that this occurs through different cognitive mechanisms, or with differing degrees of success. Until future research can carefully and systematically explore this possibility, however, this idea remains speculative.

Finally, this experiment suggests that certain kinematic parameters are perceived to influence emotion judgements more than others. Previous research in this area has explored movement kinematics through computational methods (Crane & Gross, 2007; Castellano et al., 2007; Berhardt & Robinson, 2007; Roether et al., 2009), or through evaluation of movement along subjective scales (e.g., Laban Movement qualities; Crane & Gross, 2013; Gross et al., 2012; Gross et al., 2010; Shafir et al., 2016), but no previous research has directly addressed whether these elements of movement are of conscious importance in the decisions of observers from a variety of backgrounds. During the semi-structured interviews, every single one of our participants referred to movement speed when explaining their rationale for assigning a specific emotional expressivity to the movement sequences. This suggests that speed may be particularly important for social evaluation of the human body in motion. Further support for this notion can be found in a study conducted by Orlandi and colleagues (2020). In a series of experiments participants assigned aesthetic value judgements for dance movement sequences varying in velocity profiles, but not choreographic content, and it was found that observers demonstrated a significant preference for faster movement sequences – over sequences with slower, more uniform velocity profiles (Orlandi et al., 2020). Considered together, these results suggest that it may be distinctly important to include speed in any framework for social-signalling through whole-body movement.

However, it should be noted here that there were a diverse range of additional kinematic parameters mentioned, much more sporadically, by different participants throughout the interviews (e.g., *"fluid"*, *"sharp"*, *"heavy"*, *"floaty"* etc…). Researchers have previously attempted to distil a more diverse range of movement parameters into specific kinematic profiles for different kinds of emotional expressivity, as other researchers have done with "action units" in the context of facial expressions. However, these studies have produced extremely mixed, and sometimes contradictory, results. For example, researchers have identified both indirect (Sawada et al., 2003) and direct (Wallbott, 1998) movement trajectories to be associated with the communication of sadness. The expression of happiness has also been associated with contradictory kinematic parameters; with relaxed motion in some studies (Gross et al., 2010), and with high energy motion in others (Brownlow et al., 1997). Further clarity will be required if we hope to develop a complementary set of unique movement components that are associated with the expression of different emotional states; something which exists for the classification of emotional facial expressions (i.e., the Universality Hypothesis). Therefore, we suggest that future researchers should look to capitalise on the open structure of qualitative research, to explore the complexities associated with this question.

2.4. Toward a Theoretical Framework

Throughout the chapters of this thesis, I have referenced the James-Lange, Cannon-Bard, and Schachter and Singer basic theories of emotion (James, 1884; Lange, 1885; Cannon, 1927; Bard, 1928; Schachter & Singer, 1962). Revisiting them now, it is important to reiterate that while these theories differ in both the specific order and relative importance of the mechanisms involved in producing an emotional experience, they broadly agree that emotions occur in the following way. First, we encounter an emotionally-salient input stimulus (e.g., a bear in the typically described vignette) which, in some order, prompts a physiological response within our body (e.g., increased heart rate, sweating) and cognitive processing of that stimulus (e.g., growling indicates an aggressive animal which implies danger), which then leads to the production of an emotion (e.g., “I am afraid”). In this thesis, I extend this logic for the elicitation of emotional experience, by proposing that similar processes may be involved in the production of emotion comprehension from human movement. A visualisation of the proposed mechanisms can be found in Figure 6.1.

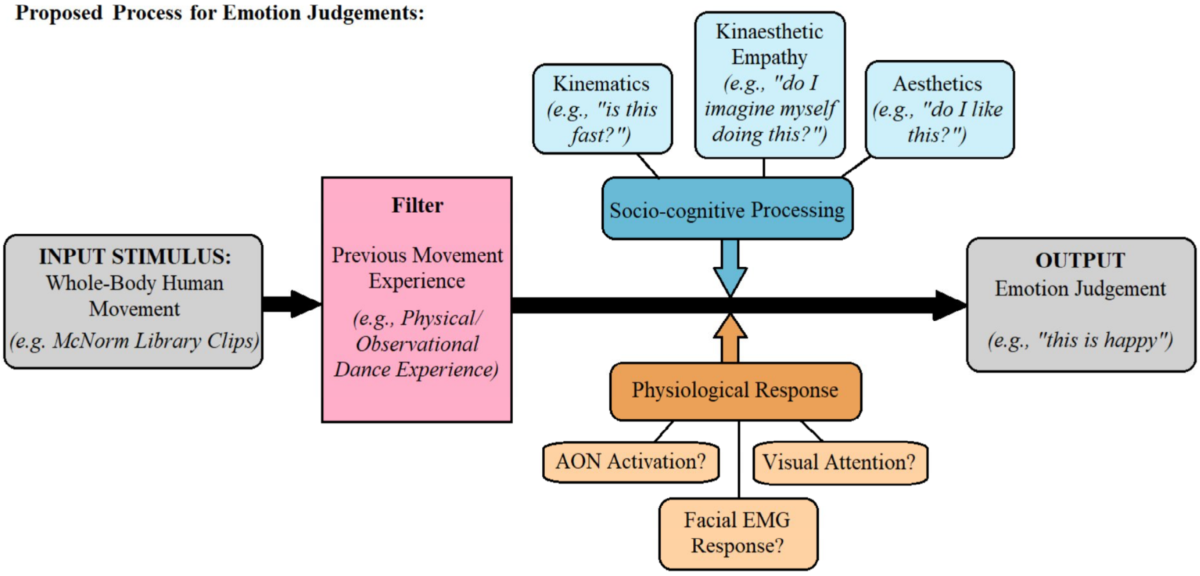


Figure 6.1: A visualisation of the proposed process for the production of an emotion judgement for whole-body human movement, based on the processes proposed for emotion elicitation in the James-Lange, Cannon-Bard and Schachter & Singer theories. This proposed process involves encountering an input stimulus (whole-body human movement), which passes through a filter of previous movement experience (including both physical and observational aspects of experience) before sociocognitively processing the stimulus and reacting to it physiologically. In this thesis we have suggested a number of aspects of cognition which may be involved (kinematic parameters like speed, the use of kinaesthetic empathy, and aesthetic evaluation; all within a real-world, social context). In section 4 of this chapter, we outline several future research projects to explore physiological reactivity including; neural activity within the Action Observation Network, micro-activation of the facial muscles, and specific visual attention strategies. The combination of these sociocognitive evaluations and bodily responses produces an emotion judgement for the whole-body movement input stimulus.

In this thesis, I have argued for the role specific aspects of cognitive decision-making play in an individual being able to assign emotional expressivity to a whole-body human movement sequence. These factors relate to the concepts of kinaesthetic empathy (e.g., “can I do this movement?”, “do I imagine myself doing this movement while I watch it?”), spontaneous aesthetic appraisal (e.g., “do I like this movement?”, “is this movement beautiful?”, “is this movement boring?”), and human movement kinematics (e.g., “is this movement fast?”, “is this movement slow?”). Over the course of four empirical chapters employing quantitative methods, I have discussed the potential role of each of these cognitive evaluations, the factors which may prompt the contemplation of these questions, and the limitations of our experimental paradigms which may have influenced the salience of these factors. In the final empirical chapter, I moved beyond the quantifiable role of these factors to explore the subjective importance of these aspects of cognitive evaluation in influencing the judgements made by observers. Therefore, based on the evaluation of previous literature in this domain, and in reviewing the contents of this thesis, I believe it is reasonable to suggest that this cognitive processing may be relevant to the **comprehension** of emotion, not only to the elicitation of these states.

I acknowledge here that this thesis did not explore bodily responses to human movement, and therefore in arguing for the role of physiological processes this requires a degree of speculation and inference from results obtained by other researchers. However, a wealth of evidence suggests that bodily responses may mediate emotional expressivity judgements assigned to socially-salient stimuli. For example, masses of neuroscientific evidence show the wide network of brain areas recruited during the detection of emotion and affect from various social cues, including faces (Jehna et al., 2011; Phan et al., 2002; Menting-Henry et al., 2022; Zhang et al., 2018; Adolphs, 2002) and voices (Drolet et al., 2012; Drolet et al., 2013; Zhang et al., 2018). However, certain neural activation patterns specifically associated with the recognition of emotional expressivity present in human body motion also appear to exist (Heberlein & Atkinson, 2009; Kret et al., 2011; Basil et al., 2017). Beyond neuroimaging research, a number of psychophysiological substrates of emotion recognition have been identified through behavioural tasks. For example, through several such studies, an intimate relationship between physiological measures like heart rate variability and galvanic skin responses and the success of emotion detection have been suggested (Bal et al., 2010; Quintana et al., 2012; Ménard et al., 2015; Gouizi et al., 2011; Du et al., 2020). Further, studies using electromyography and measures of corticospinal excitability suggest that micro-activation of bodily and facial muscles reflects the affective evaluation of human movement in observers (Jola et al., 2012; Cannon et al., 2010; Coombes et al., 2009; Engelen et al., 2018). Therefore, although I have not directly explored any measure of psychophysiology in the empirical chapters of this thesis, there is extensive evidence to suggest that, as they do in the elicitation of emotion, physiological responses to socially-salient human movement play a role in the comprehension of emotional expressivity.

Given the influence of cognitive processing, and the likely role of psychological responsiveness, in the affective evaluation of whole-body human movement, I believe that this provides evidence that the rationale present in the traditional theories of emotion elicitation (James-Lange, Cannon-Bard, and Schachter and Singer) can be extrapolated to describe mechanisms involved in other aspects of the experience of human emotions; more specifically in the comprehension of these states from the movement of others. However, in applying this theoretical framework to the recognition of emotional expressivity, I propose two slight adjustments to this account.

First, it is important to reiterate that I believe an important limitation of these theories to be the lack of acknowledgement given to the role of social factors in the experience of emotions. In real-world settings, emotions do not occur in a vacuum, they are social in nature and play a vital role in forging and maintaining relationships, communicating successfully with others, and ultimately facilitating survival in a social society. Additionally, through the empirical chapters of this thesis (and Chapter 5 in particular), we have observed the perceived importance of interpersonal factors in the evaluation of emotional expressivity from human movement. Therefore, I propose that in applying the principles of the emotion elicitation model in the context of emotion comprehension instead, *cognitive processing* should be renamed ***socio***-*cognitive processing*. This slight adjustment in terminology reflects the involvement of interpersonal evaluation which mediates interpretation of the affective content of whole-body human motion.

This inclusion of social parameters may be of particular salience in the evaluation of human dance movement. In a seminal review paper produced by Orgs and colleagues, they suggest, "a new theory of aesthetics in the performing arts that is based on communication via movement" (Orgs et al., 2016). In presenting this theory, the authors argue that the aesthetic impact of dance relies on dyadic social interaction between performers and audiences, and more specifically that interactive and bidirectional communication is an essential component of dance. Indeed, this notion was the main justification for our decision to categorise the expressive movement clips from the McNorm Library into emotion categories based on not only the subjective judgements of observers, but also through the communicative intentions of the movement performer.

Finally, I propose that before an individual can submit an input stimulus to sociocognitive processing, and respond to that stimulus physiologically, this stimulus will pass through a "filter" of previous movement experience. There are several reasons for the inclusion of this experience filter at this stage in a model of emotion comprehension. Numerous studies have shown that previous motor experience sculpts the perception and evaluation of human movement, and results in different bodily responses to observing another in motion. Previous dance experience can alter social judgements observers assign to dance movement, including cognitive assessments of beauty, liking, and interest (Kirsch et al., 2015; Kirsch et al., 2013; Orgs et al., 2018), but this previous experience also permeates the psychophysiological responses observers have when observing dance movement (Christensen et al., 2016; Jola et al., 2012; Sze et al., 2010; Vicary et al., 2017). Although the work contained in this thesis shows mixed evidence for the quantifiable role of previous experience in the recognition of emotional expressivity in the earlier empirical chapters of this thesis, comments made by participants during the interviews in the final mixed-methods experiment show that participants believed their previous experiences to play a substantial role in their decision-making processes. Therefore, these results, and those provided by the previous literature suggest that previous motor experience influences both the sociocognitive processing and bodily responsiveness processes when observing the human body in motion, and I believe this justifies the need to award motor experience due consideration in a theoretical framework for emotion comprehension from whole-body movement.

3. *Limitations of This Thesis*

A more detailed discussion of the limitations associated with each specific empirical study can be found in the discussion sections of their respective chapters, but in the following, I provide an overview of the main limitations this thesis.

3.1. The Physiological Component of Emotion Recognition

In this thesis, I propose that emotion recognition may occur through similar mechanisms to those proposed for the elicitation of emotion in the James-Lange, Cannon-Bard, and Schachter and Singer theories. Namely, that we encounter an input stimulus (e.g., emotionally expressive whole-body movement clips from the McNorm Library), which passes through a "filter" of previous movement experience, before we respond to it physiologically (e.g., through galvanic skin responses, specific visual-fixation patterns, and micro-facial expressions), and submit it to cognitive processing (e.g., "could I do this?", "do I visualise myself doing this while I'm watching it?", "is it fast, or slow?") through a social lens. Then bodily responses, and sociocognitive processing leads to an emotion judgement (e.g., "This movement is happy"). However, in this thesis, while I have conducted a thorough exploration of various aspects of cognitive processing, I have not explored any aspect of psychophysiological responses.

In order to gain a comprehensive understanding of whole-body social signalling, and to propose a robust framework for how expressivity is communicated through whole-body human movement, a thorough understanding is required of how socio-cognitive (i.e., those aspects discussed in this thesis) and physiological aspects of the emotional experience interact. While a substantial volume of previous research indicates that the kinds of individual differences explored in this thesis (e.g., previous dance experience, trait empathy) can alter physiological responses to human movement, I have only explored the role of cognitive processing in the recognition of emotional expressivity through the chapters contained in this thesis. Looking toward the future, I can already articulate concrete steps for future studies that can help fill this gap. More detailed discussion of these proposed studies can be found in Sections 4.1 – 4.4 of this chapter.

3.2. Role of Aesthetics Primed by the Concept of Dance as an Artform

Through results from the experiments described in this thesis, and considering evidence from previous empirical aesthetics research, I have proposed that affective and aesthetic evaluation of human movement stimuli may be intrinsically linked.

A potential criticism of this notion may be that observers may have provided spontaneous aesthetic evaluations of the McNorm Library clips because they were primed with the idea of dance, which exists as an artform in its own right. The links between aesthetic appraisal and the arts is obvious, and is the central tenet of empirical aesthetics and neuroaesthetics research, therefore it is clear why presenting observers with artistically-relevant stimuli may have invoked the consideration of aesthetic factors. However, there exists several results from previous experiments, and arguments made by prominent researchers in this field that may call this assumption into question. First, it has been argued that the process of submitting dance to rigorous scientific experimentation in empirical aesthetics research results in the use of stimuli which no longer resemble the artform of performative dance. In a review of the relevant literature, Christensen and Calvo-Merino provide a thorough discussion of issues related to ecological validity when we reduce the content present within dance movement stimuli (Christensen & Calvo-Merino, 2013). Here, I will briefly discuss each of these points, and how they relate to items from the McNorm Library used in the empirical chapters of this thesis, and to movement stimuli used in studies of this nature more broadly.

First, Christensen and Calvo-Merino (2013) acknowledge that when spectators observe performative dance, they obtain information not only from the body, but also from the face. It is true that we can obtain a wealth of social information from facial expressions, and visual attention research has shown that the presence of a face drives our gaze when observing the entire body of another (Gullberg & Holmqvist, 2006; Leehr et al., 2018; Wilkinson & Light, 2011). Further, research conducted by Aviezer and colleagues shows that observers believe the face to be the most salient cue to emotional expressivity, regardless of the validity of this assumption (Aviezer et al., 2012a). Therefore, when aiming to explore the recognition of emotional expressivity from movement of the human body in isolation, the previous research suggests that faces may present an undesirable confound for these judgements, and merits their exclusion from movement stimuli for social perception research. The vast majority of researchers have chosen to address this issue through manual blurring of faces in stimuli (e.g., Christensen et al., 2014; Christensen et al., 2016; Melzer et al., 2019; Shafir et al., 2016), and others have reduced the potential influence of this cue even further by presenting movements in the form of point-light displays – as I have chosen to do in creating the McNorm Library. Further, through obscuring morphological information about the dancer's body through representing movement in the form of point-light displays removes additional visual information which spectators are privy to when watching dance in the context of performance. While employing this approach in social perception research presents a unique kind of control which is only possible in laboratory conditions, in doing so, the resulting movement stimuli are several steps removed from the artform of performative dance.

Second, they note that the movement stimuli used in social perception research rarely include music or other sounds (Christensen & Calvo-Merino, 2013). Generally, the only exception to this is within experiments using such stimuli to specifically explore how the auditory cues interact with motion to influence perceptions (e.g., the role of breathing and effort sounds on aesthetic evaluation – Jola et al., 2014; exploring the influence of emotionally-salient music on valence recognition – Christensen et al., 2014). However, historically, musical scores were developed in tandem with choreography to create a number of world-renowned classical ballets (e.g., *Swan Lake*, *The Nutcracker*), and orchestras continue to play a substantial role in performative dance productions. Again, given the potentially confounding role of auditory cues, it is clear to see why such cues are removed from stimuli used in social perception research for movement in isolation. Indeed, I did not include musical accompaniments to the McNorm Library clips for this exact reason. However, removing the presence of music brings us yet another step away from dance as an artform.

Finally, Christensen and Calvo-Merino argue that condensing the experience of a full-scale dance performance into a series of short clips (some as short at 3-seconds; Calvo-Merino, 2010; Sun et al., 2020) brings us an additional step further removed from the artform of dance. They suggest several potential consequences of reducing dance sequences in this way. For example, by selecting short excerpts of movement to meet the aim of temporal consistency across stimuli, choreographic phrases may be trimmed at inappropriate points, and this may result in the complete loss of context which is developed throughout the often hours-long duration of dance performances (Christensen & Calvo-Merino, 2013). While in the creation of the McNorm Library, we did not impose restrictions on the duration of movement sequences that could be performed by the dancer, and we did produce stimuli with significantly longer durations than those in some previous movement repositories (e.g., 3-seconds in Calvo-Merino, 2010; 6-seconds in Christensen et al., 2014; 20-seconds in Ma et al., 2006), it is still likely that the longest McNorm Library dance sequence is not equivalent in content, or context, to the almost three-hour run time of, for example, the Royal Opera House's production of *Swan Lake*. Therefore, in this way too, the McNorm Library clips may not reflect the nature of performative dance as an artform. These factors combined denote that from the perspective of a dance practitioner, or dance researcher in the arts domain, there are several degrees of separation between the content of the McNorm Library movement clips and the artform of performative dance, and this may limit the extent to which participants are primed to provide spontaneous aesthetic evaluation. However, additional support for our proposal that aesthetic appraisal is a key aspect of processing in the evaluation of human movement more generally, not only in relation to aesthetically-salient movements like dance, can be found in a series of research works exploring the evaluation of everyday actions through the use of different psychophysiological measures.

Across several experiments, researchers have found that observers provide implicit aesthetic value judgements for actions as simple as grasping movements. Hayes and colleagues asked a series of participants to reach toward different household objects (e.g., canned food, washing powder) using direct versus convoluted movement trajectories. These object interactions were recorded and later shown to those participants, in addition to a new group of observers who had not taken part in the grasping task, and they were asked to provide aesthetic value judgements (i.e., liking ratings) for the objects in the videos. They found that individual household objects were assigned higher liking ratings when the movement trajectories used to approach an object in question was more fluid (compared to convoluted). This was found to be the case for both the participants who performed those grasping actions themselves, and for the participants who had merely observed those actions (Hayes et al., 2008). In a later experiment, Cannon and colleagues measured micro-facial expressions (as a proxy for affective evaluation) in response to those same grasping actions using facial electromyography. In this experiment, participants observed a series of videos depicting grasping actions wherein orientation of the objects were, or were not, compatible with positioning of the arms, and as a result would produce fluid, or stilted motion. In the first four experimental blocks, participants were asked to perform an unrelated categorisation task while EMG data were collected (to obtain an implicit measure of affective evaluation of the movements), and only in the final block were participants asked to provide liking ratings for the objects. They observed that although participants did not provide significantly more positive aesthetic appraisals for objects grasped with movements varying in fluidity, they exhibited greater activity in the zygomaticus major (i.e., the "smiling muscle") in response to objects which were grasped through fluid arm movements (Cannon et al., 2010). Due to the higher liking ratings assigned to fluid movements in the Hayes (2008) experiment, and more positive affective evaluation of the fluid movements observed in the Cannon (2010) study, it appears there is clear and consistent evidence of an implicit relationship between affective and aesthetic evaluation of human movement even when considering mundane, everyday grasping actions. As such, this work provides further support for our proposal that aesthetic evaluation is an integral component of human movement perception, more generally.

3.3. Revisiting Limitations of the McNorm Library

The McNorm Library movement clips were at the centre of all four of our empirical research works. Therefore, evaluation of these stimuli warrant reconsideration when discussing the contents of this thesis as a whole. Before doing so, it should be acknowledged here that I believe the McNorm Library holds significant value for researchers in the fields of social perception and empirical aesthetics, for a number of reasons which have already been discussed extensively throughout this thesis. However, there are a number of potential limitations associated with this movement repository that merit deliberation here.

3.3.1. Idiosyncrasies in Emotional Communication

First, I aimed to ensure that emotion classification of our movement stimuli was robust, by accounting for the intentions of the movement performer and perceptions of the movement observers. However, it should be acknowledged that while we explored the perspectives of 127 observers, we only accounted for the expressive intentions of one movement performer. Through the various validation experiments presented in this thesis, we observed that fear was consistently recognised with the lowest level of accuracy. I have proposed that fear may be more idiosyncratic, given the various ways in which fear can manifest within the human body (i.e., the fight, flight or freeze responses; Barlow, 2004; Schmidt et al., 2008). However, in these experiments we also observed that anger was recognised with relatively low levels of accuracy. It is possible that anger, and indeed the other emotions explored in this thesis, also have idiosyncratic presentations which should be accounted for through the perspective of multiple movement performers. During the interviews presented in Chapter 5, one of the participants provided the following comment which feels relevant to this discussion;

KR281508-Dancer (P62, L24-30) *"...I think almost like with an anger, there's- I know* ***everybody when they're angry is very, very different*** *and like, to me, sometimes I'm like if I say the word angry I don't mean someone that's being really, really super violent and like vicious or whatever. Just like when I get angry, sometimes I can then get upset because I'm angry. Or it's like the major like frustration and stuff, and that to me in my mind* ***that is my anger and it's not potentially somebody else's****."*

This excerpt highlights the idea that the way in which one individual expresses anger, for example, may diverge significantly from how another person communicates that same emotion. By including only the perspective of one movement performer in the creation of a stimulus library, we may be generating a kind of emotional expressivity which is very personal, and perhaps unique to that movement performer. It may have been more beneficial, perhaps, to capture expressivity of the different emotions, anger and fear more specifically, through the movements of multiple performers. In doing so, we may have been able to identify certain commonalities to isolate more universal components of emotionally-salient movement, and account for any individual differences in the transmission of affective expression. I believe this to be a limitation of the current work, which may explain why certain kinds of expressivity were recognised with lower accuracy than others in our validation experiments.

3.3.2. A WEIRD Focus

A widely acknowledged criticism of social perception research, and psychological research more broadly, is the almost exclusive focus on the perspectives of Western, Educated, Industrialized, Rich and Democratic (WEIRD) participants. The problems associated with exploring judgements of exclusively WEIRD samples are evident from the controversy surrounding the validity of the Universality Hypothesis for the recognition of emotional facial expressions. As noted in various chapters of this thesis, it was widely believed that there are six facial expressions (happy, sad, angry, fear, disgust, and surprise) which produce unique patterns of muscular activity in the face, and that these six facial expressions are universally recognised by observers from a large number of countries (Ekman et al., 1969; Tomkins, 2008; Ekman & Friesen, 1978). However, more recent research conducted by Jack and colleagues provided a series of challenges to this long-standing assumption. In the seminal 2016 paper, Jack and colleagues observed that some of these emotional states were recognised with significantly lower accuracy among Asian participants, and suggested that there may only be four, not six, culturally common facial expressions (Jack et al., 2016). Although this work was conducted to explore emotional expressivity through cues from the human face, not the body, these results suggest that cultural differences may be a particularly relevant aspect of social communication to address.

In this thesis, I acknowledge that a focus on recognition of expressivity from Western dance styles (classical ballet and western contemporary dance), in participants from the UK, contributes to the WEIRD problem, and to a western-centric rhetoric in social science. However, some researchers in the fields of social perception and empirical aesthetics have taken steps to broaden our understanding of the universality of engaging with and perceiving dance by exploring styles practiced by performers from different cultural backgrounds within more diverse participant samples. For example, Darda and colleagues have conducted a number of studies exploring evaluation of Indian Bharatanatyam dance among culturally-diverse participant samples (Darda & Cross, 2021; Darda & Cross, 2022). Christensen and colleagues highlight the importance of doing so, in a recent paper discussing Persian "classical" dance. These authors argue several reasons for the importance of exploring such underrepresented movement styles in experimental work, but mainly emphasize that due to the social and political climate in Iran, there exists a widespread prohibition of Persian "classical" dance. Further, they argue that without raising awareness of, and interest in, relatively unfamiliar dance styles among the academic community, they may be in danger of underdevelopment, or indeed in the case of Persian dance specifically, may risk eradication (Christensen, Khorsandi, Smith et al., submitted manuscript).

Moreover, including more diverse participant samples may be of particular importance for my research questions, as it has already been noted in this chapter that Jack and colleagues found that culture impacts emotion recognition from facial expressions (Jack et al., 2016; Jack, Caldara & Schyns, 2012; Jack et al., 2012). Across several studies Jack and colleagues have observed that participants from Asia recognise facial expressions of disgust and surprise with significantly lower accuracy than participants from the UK (Jack et al., 2016). It is likely that cultural differences may too be salient in the context of bodily cues, as Monroy and colleagues (2021) have observed differences in aesthetic value judgements for movements differing in synchrony between Japanese and British observers. More specifically, Japanese participants liked synchronous and asynchronous movements equally, while British participants showed a distinct preference for asynchronous dance movements (Monroy et al., 2021). Based on these results, and my proposal that affective and aesthetic evaluation of human movement may be intrinsically linked, this further highlights the importance of exploring judgements of less WEIRD samples for the McNorm Movement Library clips.

Considering these factors together, it is likely that culture plays a significant role in social perception from emotionally expressive stimuli, including videos depicting whole-body dance movements, and including underrepresented movement styles in empirical work may confer additional benefits in relation to conservation, and improved social awareness. I have not addressed culturally-salient aspects of social perception in this thesis, and I believe this should be rectified in future work. For example, it may be useful to expand the McNorm Library to include movement sequences from a more diverse range of dance styles (e.g., Bharatanatyam, Persian "classical" dance).

4. *Specific Future Research Directions*

To conclude this discussion, in the following section I will outline the aims and methods of four follow-up experiments that I believe would be worthwhile supplements for the contents of this thesis.

4.1. *A Comparison of Computational and Subjective Measures of Image Complexity*

In Chapter 4, we describe a study examining the emotional and aesthetic evaluations of black and white continuous line drawings. In this experiment, we also describe the creation of a novel method for quantifying the complexity present in such images. From a visual examination of the data, and given the behavioural results showing an expected inverse U-shaped relationship between complexity and aesthetic value judgements using this metric, it would appear that we have developed a robust, and fully-automated method for capturing the complexity of an image within a single value. However, it is unclear whether this metric is reflective of subjective assessments of complexity, and if it is, how this aggregated complexity factor compares with existing methods for quantifying complexity present in images.

To this end, I would like to conduct a follow-up experiment wherein a large participant sample (N = ~200) is presented with the same 40 trajectory images used in our original line drawing experiment, and are asked to provide subjective complexity ratings for each of the images along a 100-point slider scale. After obtaining an average subjective complexity score for each image, I can explore how the observer judgements align with the individual component measures, and our aggregated complexity factor. By examining which factor produces an order that is most similar to the order provided by a large group of observers, we could potentially validate our metric, or identify which of the existing measures is most representative of image complexity from a perceptual standpoint.

4.2. *The Role of Visual Attention in the Evaluation of Social Signalling from Whole-Body Movements*

An observer's visual fixation patterns provide insight into the structure of attention and awareness (Kietzmann, Geuter & König, 2011). As noted in a recent review by Lockhofen & Mulert, visual attention facilitates processing of information or stimuli based on what may be most salient for achieving a particular goal (Lockhofen & Mulert, 2021). Eye-tracking studies have been employed in a number of fields to evaluate what features of a visual stimulus are most important for the acquisition of knowledge, and comprehension of information (Hasanzadeh et al., 2018; Bard et al., 1980; Wallenborn, 2011; Hancock & Ste-Marie, 2013). In applying this logic within the context of my research questions, it may be possible to uncover whether there are particular body regions which are perceived to communicate the most salient cues to emotional expressivity through the use of an eye-tracking and emotion recognition experiment.

In such an experiment, research suggests that it may be useful to collect data from participants with various levels of expertise, in relation to both movement experience and emotion recognition capabilities. Expertise-driven visual fixation patterns have emerged among experts in a variety of professions, including construction workers (Hasanzadeh et al., 2018), gymnastics and dressage judges (Bard et al., 1980; Wallenborn, 2011), ice hockey referees (Hancock & Ste-Marie, 2013), and medical professionals (Bertram et al., 2013). But expertise driven visual attention has also been found to emerge among professional dancers and choreographers (Petrakis, 1987; Stevens et al., 2010). Therefore, when using human movement as stimuli in such an emotion recognition experiment, it may be important to collect data from participants with varying degrees of previous dance experience.

To this end, I propose the following experimental paradigm. A number of non-dancers, amateur dancers and professional dancers will be asked to complete the Goldsmith's Dance Sophistication Index (Rose et al., 2020) before assigning emotion labels to 20 clips from the McNorm Library (4 of each; neutral, happy, sad, angry, and fearful). During the emotion recognition task, eye-movement data will be collected (using EyeLink, Tobii or similar tools), to explore visual attention directed toward five target body regions (head, arms, torso, legs, and feet).

An iteration of this experiment was the focus of my Undergraduate and Masters degree theses (Smith & Pollick, forthcoming). In these earlier experiments, I used FLD movement videos of performative ballet sequences (taken from the library created by Christensen et al., 2014). These FLDs included a number of superfluous visual confounds (including stage furniture, costumes, and occasionally the presence of other dancers in the background), which results suggested may have detracted attention from the movements themselves. Further, in my earlier experiments I only explored visual attention patterns during the recognition of happiness and sadness. Therefore, these results did not provide any indication of whether visual attention or emotion recognition may differ among dancers (both amateur and professionals), and non-dancers for other basic emotions (e.g., anger and fear), or for the absence of expressivity (i.e., neutral).

Therefore, to address the limitations of my previous work, I plan to carry out an experiment which explores visual attention directed toward movement specifically, in isolation from other contextual cues, using the point-light display representations of movement from the McNorm Library. The McNorm Library allows us to examine recognition of emotion from movement sequences which contain an emotional expressivity which is agreed upon by both observers and the movement performer, contain no superficial visual information, and contains motion depicting four basic emotional states (happiness, sadness, anger, and fear), in addition to the absence of expressivity (neutral motion), without containing culturally and socially-relevant gestures or choreography from existing performative ballets.

Results obtained from this experiment may indicate whether, during the observation of movement for the purposes of emotion recognition, there are particular body regions which observers believe to convey the most salient social cues. Further, by exploring visual fixation patterns of observers with the greatest recognition accuracy may provide some insight into which body areas, or visual-search strategies may be most useful for the perception of social signals from the dynamic human body. Therefore, I believe this experiment would be a valuable next step in developing the model I propose in this thesis, and would hold benefits for the social psychology researcher community more broadly.

4.3. Using Electromyography to Explore the Conscious and Unconscious Evaluation of Emotionally Expressive Whole-Body Motion

A wealth of previous literature suggests that dancers exhibit stronger bodily responses when observing others in motion than individuals without any prior movement training. Evidence of this can be observed through the results of numerous neuroimaging experiments which show that extensive dance experience prompts greater activation of the Action Observation Network (Burzynska et al., 2017; Cross, 2010; Gardner et al., 2015; Kirsch et al., 2018; Liew et al., 2013; Calvo-Merino et al., 2005; Cross et al., 2009; Gerson et al., 2015), and through expertise-related differences observed across a number of studies using psychophysiological measurements (e.g., corticospinal excitability – Jola et al., 2012; galvanic skin response – Christensen et al., 2016; Christensen et al., 2021). Additionally, results from a study conducted by Kirsch and colleagues suggest that dance experience may provoke heightened and valence-specific physiological responses in observer's faces when evaluating human movement stimuli. Based on these results, it is likely that dancers and non-dancers would exhibit diverging EMG responses when observing dance, and that emotion-specific microfacial expressions may be more distinct among those with previous motor training.

However, to our knowledge, no research has explored expertise-related differences in the EMG responses of dancers and non-dancers while assigning a wider array of basic emotional labels to human dance motion. It is important to understand whether such differences would emerge because this may provide valuable insights into the success of both conscious and unconscious processing of socially-salient stimuli, as EMG responses may reveal microfacial expressions which align more closely with either the intended (i.e., target) or perceived (i.e., explicit judgement) expressivity of different emotional states within human movement stimuli. Therefore, to address this gap in the literature, I plan to conduct an experiment which addresses this question directly.

In this experiment, participants with varying degrees of previous experience will be asked to complete the Gold-DSI (Rose et al., 2020), and the Toronto Empathy Scale (Spreng et al., 2009) before taking part in an emotion recognition task. During this task, they will assign emotion labels to 20 whole-body movement clips from the McNorm Library (4 each; neutral, happy, sad, angry, and fearful) while EMG responses are collected from the corrugator supercilia and zygomaticus major. I will then evaluate whether activity in facial muscles reflects the intended emotional expressivity of the movement clip (i.e., the target emotion), and the perceived emotional expressivity (i.e., label chosen by participants). I will also explore whether EMG responses are stronger when the target and perceived emotions are congruous (i.e., when assigning the correct emotion label) compared to when they are incongruous (i.e., when assigning an emotion label which is different to the target emotion label). Finally, I will explore whether participants with a more extensive physical or observational dance history will produce quantifiably different EMG responses to those without such experience, in addition to exploring whether individual differences in trait empathy influence these responses.

By identifying the relationship between EMG responses to human movement, and the judgements of emotion provided explicitly by observers, we may be able to identify the extent to which physiology contributes to the identification of emotion from whole-body human movement. Regardless of the results obtained, this experiment could contribute to our appreciation of whole-body social signalling more generally, and may provide more specific benefits for the content of this thesis; by providing some insight into the role of the physiological processes in the model for the production of emotion judgements I have proposed in this chapter.

4.4. A Qualitative Exploration of Choreographic and Performative Decision-Making in the Communication of Emotional Expressivity Through Dance Movement

During the interviews, from the experiment presented in Chapter 5, a number of participants noted that they assigned emotions to the McNorm Library clips based on learned associations. Additionally, participants noted that during the task they had based their emotion judgements on the specific choreography, rather than the expressive *quality* present within the movement clips. Although in creating the McNorm Library, we aimed to create a series of movement sequences which were devoid of culturally or socially relevant gestures, and any specific choreographic sequences from performative ballet, discussions during these interviews suggest that observers may hold certain emotional associations with specific movements based on their previous experiences (with dance, from real-world social situations, or with things they had seen in the media).

It is unclear from existing research whether there are specific movements that are included within performative dance for the communication of specific emotional states through whole-body movement. Further, it is unclear how, when provided with specific movement phrases, dancers alter the way in which they perform those movements to communicate expressivity that is relevant to the emotionally-charged narratives in performative dance. To address this gap in our understanding, I plan to conduct a qualitative investigation of this question.

More specifically, in a future experiment, I will recruit a series of professional dancers and choreographers to take part in a series of semi-structured interviews or focus groups where I will explore, for example, the following kinds of questions;

For choreographers;

1. When designing a piece of choreography which aims to communicate X emotion (i.e., happiness, sadness, anger, or fear), are there any particular movements you would include within that choreography?
2. Are there any movements that have long-standing associations with, or are characteristic of X (i.e., happiness, sadness, anger, or fear) emotion in performative dance?

For dancers;

1. When you are given a movement, or movement phrase, how do you alter the way you perform that movement/those movements to express X emotion (i.e., happiness, sadness, anger, or fear)?

During the interviews (Chapter 5), it was also noted that many dancers struggled to verbalise their thought processes, often resorting to using their physicality or illustrative examples of motion to aid in this communication. In light of this, it will be extremely important to record these interviews in full using a video camera. In doing so, I can evaluate not only the verbal responses provided by the participants, but I can also annotate the gestures, movements, and physicality they employ during the discussions.

After initially identifying any specific movements which are associated with a particular kind of emotional expressivity within our sample of participants, I (or other researchers) could conduct further replications of such an experiment to explore the universality of these findings across different dance practitioners. After cross-validation, it is hoped that the resultant findings may be useful in serving as a guide for annotation of past and future emotionally expressive dance movement libraries, and may highlight movements to avoid including during the creation of such repositories.

CONCLUSIONS

Through the use of both quantitative and qualitative research methods, this thesis has explored a number of factors that can influence the success of emotion recognition from expressive whole-body movement sequences. Overall, results from a validation experiment, and several replications of that experiment, show that the movement library proposed in this thesis (The McNorm Movement Library) is a valuable resource for researchers in the fields of social psychology, empirical aesthetics, and other related disciplines due to its high level of test-retest reliability, and its more balanced reflection of the dyadic nature of affective social communication. Throughout the four main empirical chapters of this thesis, results show that the relationship between empathy, previous dance experience, and emotion recognition skills may be more complex than previous studies (primarily those conducted with neuroimaging) might suggest. Here we found that general trait empathy did not always influence recognition ability in the predicted direction, kinaesthetic empathy was neither universal nor unique to individuals with previous movement training, and both observational and physical dance experience did not always result in global improvements to performance in emotion recognition tasks.

Finally, by drawing on new insights gained through the empirical chapters, I have proposed an updated model of emotion judgement, which is grounded in the historical theories of emotion elicitation. More specifically, I propose that when an individual encounters whole-body human movement as an input stimulus, this passes through a filter of previous dance experience (including both physical and observational aspects of experience), before that stimulus is appraised in a socio-cognitive context (including aesthetic evaluation – is this beautiful?; perception of movement kinematics – is this fast?; and kinaesthetic empathy processes – do I imagine myself doing this while I watch it?), and reacted to by aspects of human physiology (e.g., microfacial expressions, specific visual fixation patterns, activity of the Action Observation Network). These processes then combine to produce an emotion judgement. While the empirical work presented in this thesis has focused on elements of socio-cognitive processing, I have proposed a number of future follow-up research directions in this final general discussion chapter to address gaps in the proposed role of physiological processes based on existing, related literature.

FINAL THOUGHTS

This framework I propose for how the dynamic human body can transmit socially salient information requires further research, which I aim to pursue in the future, and would encourage other researchers to consider too. Facial expression research has dominated the bulk of existing affective communication research, but the results presented in this thesis suggest that the human body in motion provides a wealth of social cues that are highly valuable to observers. Therefore, until testable theoretical accounts (like the one proposed in this thesis) for social cues beyond facial expressions are offered, and submitted to the same kind of rigorous empirical evaluation, our current appreciation of human social signalling, as a whole, cannot be taken at face-value.

REFERENCES

ABOUT NEUROLIVE. (n.d.). NEUROLIVE. Retrieved April 22, 2022, from https://neurolive.info/About

Adolphs, R. (2002). Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behavioral and cognitive neuroscience reviews*, *1*(1), 21-62.

Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nature Reviews Neuroscience*, *4*(3), 165-178.

Adolphs, R., & Anderson, D. J. (2018). *The neuroscience of emotion: A new synthesis*. Princeton University Press.

Alaerts, K., Nackaerts, E., Meyns, P., Swinnen, S. P., & Wenderoth, N. (2011). Action and emotion recognition from point light displays: an investigation of gender differences. *PloS one, 6*(6), e20989.

Alaerts, K., Woolley, D. G., Steyaert, J., Di Martino, A., Swinnen, S. P., & Wenderoth, N. (2014). Underconnectivity of the superior temporal sulcus predicts emotion recognition deficits in autism. *Social cognitive and affective neuroscience*, *9*(10), 1589-1600.

Albers, C., & Lakens, D. (2018). When power analyses based on pilot data are biased: Inaccurate effect size estimators and follow-up bias. *Journal of experimental social psychology*, *74*, 187-195.

Algoe, S. B., Dwyer, P. C., Younge, A., & Oveis, C. (2020). A new perspective on the social functions of emotions: Gratitude and the witnessing effect. *Journal of Personality and Social Psychology*, *119*(1), 40.

Andrew, R. J. (1963). The origin and evolution of the calls and facial expressions of the primates. *Behaviour*, 1-109.

Angyal, C. (2021). *Turning Pointe: How a New Generation of Dancers Is Saving Ballet from Itself.* Hachette UK.

Archer, D. (1997). Unspoken diversity: Cultural differences in gestures. *Qualitative sociology*, *20*(1), 79-105.

Aristidou, A., Zeng, Q., Stavrakis, E., Yin, K., Cohen-Or, D., Chrysanthou, Y., & Chen, B. (2017, July). Emotion control of unstructured dance movements. In *Proceedings of the ACM SIGGRAPH/Eurographics symposium on computer animation* (pp. 1-10).

Arnold, A., Meggs, S. M., & Greer, A. G. (2014). Empathy and aesthetic experience in the art museum. *International Journal of Education Through Art*, *10*(3), 331-347.

Arslan, R. C., Walther, M. P., & Tata, C. S. (2020). formr: A study framework allowing for automated feedback generation and complex longitudinal experience-sampling studies using R. *Behavior Research Methods*, *52*(1), 376-387.

Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception, 33*(6), 717-746.

Atkinson, A. P., Tunstall, M. L., & Dittrich, W. H. (2007). Evidence for distinct contributions of form and motion information to the recognition of emotions from body gestures. *Cognition*, *104*(1), 59-72.

Attneave, F. (1957). Physical determinants of the judged complexity of shapes. *Journal of experimental Psychology*, *53*(4), 221.

Avenanti, A., Annella, L., Candidi, M., Urgesi, C., & Aglioti, S. M. (2013). Compensatory plasticity in the action observation network: virtual lesions of STS enhance anticipatory simulation of seen actions. *Cerebral cortex*, *23*(3), 570-580.

Aviezer, H., Trope, Y., & Todorov, A. (2012a). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science*, *338*(6111), 1225-1229.

Aviezer, H., Trope, Y., & Todorov, A. (2012b). Holistic person processing: faces with bodies tell the whole story. *Journal of personality and social psychology*, *103*(1), 20.

Bachmann, J., Zabicki, A., Munzert, J., & Krüger, B. (2020). Emotional expressivity of the observer mediates recognition of affective states from human body movements. *Cognition and Emotion, 34*(7), 1370-1381.

Badshah, A. M., Ahmad, J., Rahim, N., & Baik, S. W. (2017, February). Speech emotion recognition from spectrograms with deep convolutional neural network. In *2017 international conference on platform technology and service (PlatCon)* (pp. 1-5). IEEE.

Bal, E., Harden, E., Lamb, D., Van Hecke, A. V., Denver, J. W., & Porges, S. W. (2010). Emotion recognition in children with autism spectrum disorders: Relations to eye gaze and autonomic state. *Journal of autism and developmental disorders*, *40*(3), 358-370.

Balconi, M., & Bortolotti, A. (2012). Detection of the facial expression of emotion and self-report measures in empathic situations are influenced by sensorimotor circuit inhibition by low-frequency rTMS. *Brain Stimulation*, *5*(3), 330-336.

Balconi, M., & Bortolotti, A. (2012). Detection of the facial expression of emotion and self-report measures in empathic situations are influenced by sensorimotor circuit inhibition by low-frequency rTMS. *Brain Stimulation*, *5*(3), 330-336.

Bard, C., Fleury, M., Carrière, L., & Hallé, M. (1980). Analysis of gymnastics judges' visual search. *Research Quarterly for Exercise and Sport*, *51*(2), 267-273.

Bard, P. (1928). A diencephalic mechanism for the expression of rage with special reference to the sympathetic nervous system. *American Journal of Physiology-Legacy Content, 84*(3), 490–515.

Barliya, A., Omlor, L., Giese, M. A., Berthoz, A., & Flash, T. (2013). Expression of emotion in the kinematics of locomotion. *Experimental brain research*, *225*(2), 159-176.

Barlow, D. H. (2004). *Anxiety and its disorders: The nature and treatment of anxiety and panic*. Guilford press.

Barrett, L. F. (2006). Valence is a basic building block of emotional life. *Journal of Research in Personality, 40*(1), 35-55.

Barrett, L. F., & Bliss-Moreau, E. (2009). Affect as a psychological primitive. *Advances in experimental social psychology, 41,* 167-218.

Basil, R. A., Westwater, M. L., Wiener, M., & Thompson, J. C. (2017). A causal role of the right superior temporal sulcus in emotion recognition from biological motion. *Open Mind*, *2*(1), 26-36.

Behrends, A., Müller, S., & Dziobek, I. (2016). Dancing supports empathy: The potential of interactional movement and dance for psychotherapy. *European psychotherapy, 13*, 99-131.

Bernhardt, D., & Robinson, P. (2007, September). Detecting affect from non-stylised body motions. In *International conference on affective computing and intelligent interaction* (pp. 59-70). Springer, Berlin, Heidelberg.

Bertamini, M., & Makin, A. D. (2014). Brain activity in response to visual symmetry. *Symmetry*, *6*(4), 975-996.

Bertram, R., Helle, L., Kaakinen, J. K., & Svedström, E. (2013). The effect of expertise on eye movement behaviour in medical image perception. *PloS one*, *8*(6), e66169.

Besel, L. D., & Yuille, J. C. (2010). Individual differences in empathy: The role of facial expression recognition. *Personality and individual differences*, *49*(2), 107-112.

Bishop, S. L., & Seltzer, M. M. (2012). Self-reported autism symptoms in adults with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *42*(11), 2354-2363.

Blair, R. J. R. (2003). Facial expressions, their communicatory functions and neuro–cognitive substrates. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *358*(1431), 561-572.

Bläsing, B., Calvo-Merino, B., Cross, E. S., Jola, C., Honisch, J., & Stevens, C. J. (2012). Neurocognitive control in dance perception and performance. *Acta psychologica*, *139*(2), 300-308.

Boté, J. J., & Térmens, M. (2019). Reusing data: Technical and ethical challenges. *DESIDOC Journal of Library & Information Technology*, *39*(6).

Brattico, E., & Jacobsen, T. (2009). Subjective appraisal of music: Neuroimaging evidence. *Annals of the New York Academy of Sciences*, *1169*(1), 308-317.

Brattico, E., Bogert, B., & Jacobsen, T. (2013). Toward a neural chronometry for the aesthetic experience of music. *Frontiers in psychology*, *4*, 206.

Brattico, E., Bogert, B., Alluri, V., Tervaniemi, M., Eerola, T., & Jacobsen, T. (2016). It's sad but I like it: The neural dissociation between musical emotions and liking in experts and laypersons. *Frontiers in Human Neuroscience*, *9*, 676.

Braun, V., & Clarke, V. (2012). Thematic analysis. In H. Cooper, P. M. Camic, D. L. Long, A. T. Panter, D. Rindskopf, & K. J. Sher (Eds.), *APA handbook of research methods in psychology, Vol. 2. Research designs: Quantitative, qualitative, neuropsychological, and biological* (pp. 57–71). American Psychological Association.

Breukel, A., van Dijk, C., & Spee, K. (2016). The relative importance of aesthetics in the adoption process of solar panels in the Netherlands. *EESD*, *8*, 73-80.

Bromgard, G. D., Trafimow, D., & Silvera, D. H. (2013). The influence of presentation format on the “Bigger is Better”(BIB) effect. *Psychological Reports*, *112*(2), 458-468.

Brownlow, S., Dixon, A. R., Egbert, C. A., & Radcliffe, R. D. (1997). Perception of movement and dancer characteristics from point-light displays of dance. *The Psychological Record*, *47*(3), 411-422.

Bu, F., Mak, H. W., Bone, J. K., & Fancourt, D. (2022). Longitudinal changes in home-based arts engagement during and following the first national lockdown due to the COVID-19 pandemic in the UK. *Perspectives in Public Health*, *142*(2), 117-126.

Burger, B., Thompson, M. R., Luck, G., Saarikallio, S., & Toiviainen, P. (2013). Influences of rhythm- and timbre-related musical features on characteristics of music-induced movement. *Frontiers in psychology*, *4*, 183.

Burton, E., Wattam-Bell, J., Rubin, G. S., Atkinson, J., Braddick, O., & Nardini, M. (2016). Cortical processing of global form, motion and biological motion under low light levels. *Vision research*, *121*, 39-49.

Burzynska, A. Z., Finc, K., Taylor, B. K., Knecht, A. M., & Kramer, A. F. (2017). The dancing brain: Structural and functional signatures of expert dance training. *Frontiers in human neuroscience*, 566.

Button, K. S., Ioannidis, J., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S., & Munafò, M. R. (2013). Power failure: why small sample size undermines the reliability of neuroscience. *Nature reviews neuroscience*, *14*(5), 365-376.

Calder, A. J., Keane, J., Manes, F., Antoun, N., & Young, A. W. (2000). Impaired recognition and experience of disgust following brain injury. *Nature neuroscience*, *3*(11), 1077-1078.

Calvo-Merino, B. (2010). Neural mechanisms for seeing dance. In *The neurocognition of dance* (pp. 163-186). Psychology Press.

Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E., & Haggard, P. (2005). Action observation and acquired motor skills: an FMRI study with expert dancers. *Cerebral cortex*, *15*(8), 1243-1249.

Calvo-Merino, B., Jola, C., Glaser, D. E., & Haggard, P. (2008). Towards a sensorimotor aesthetics of performing art. *Consciousness and cognition*, *17*(3), 911-922.

Camurri, A., Lagerlöf, I., & Volpe, G. (2003). Recognizing emotion from dance movement: comparison of spectator recognition and automated techniques. *International journal of human-computer studies*, *59*(1-2), 213-225.

Cannon, P. R., Hayes, A. E., & Tipper, S. P. (2010). Sensorimotor fluency influences affect: Evidence from electromyography. *Cognition & Emotion*, *24*(4), 681-691.

Cannon, W. B. (1927). The James-Lange theory of emotions: A critical examination and an alternative theory. *The American Journal of Psychology, 39*(1/4), 106–124.

Caplan, S. E. (2003). Preference for online social interaction: A theory of problematic Internet use and psychosocial well-being. *Communication research*, *30*(6), 625-648.

Caspers, S., Zilles, K., Laird, A. R., & Eickhoff, S. B. (2010). ALE meta-analysis of action observation and imitation in the human brain. *Neuroimage*, *50*(3), 1148-1167.

Castellano, G., Villalba, S. D., & Camurri, A. (2007, September). Recognising human emotions from body movement and gesture dynamics. In *International Conference on Affective Computing and Intelligent Interaction* (pp. 71-82). Springer, Berlin, Heidelberg.

Castro Jaramillo, C. M., & Panhofer, H. (2021). Dance movement therapy techniques to promote kinaesthetic empathy for couples:'I have to dance seriously with you'. *Body, Movement and Dance in Psychotherapy*, 1-17.

Celeghin, A., Diano, M., Bagnis, A., Viola, M., & Tamietto, M. (2017). Basic emotions in human neuroscience: neuroimaging and beyond. *Frontiers in Psychology*, *8*, 1432.

Chamberlain, R., Berio, D., Mayer, V., Chana, K., Leymarie, F. F., & Orgs, G. (2022). A dot that went for a walk: People prefer lines drawn with human-like kinematics. *British Journal of Psychology*, *113*(1), 105-130.

Chan, D. W. (1985). Perception and judgment of facial expressions among the Chinese. *International Journal of Psychology*, *20*(3-4), 681-692.

Chang, D. H., Ban, H., Ikegaya, Y., Fujita, I., & Troje, N. F. (2018). Cortical and subcortical responses to biological motion. *Neuroimage*, *174*, 87-96.

Chazelle, B., & Incerpi, J. (1984). Triangulation and shape-complexity. *ACM Transactions on Graphics (TOG)*, *3*(2), 135-152.

Chen, C. C., Wu, J. H., & Wu, C. C. (2011). Reduction of image complexity explains aesthetic preference for symmetry. *Symmetry*, *3*(3), 443-456.

Chen, C., & Jack, R. E. (2017). Discovering cultural differences (and similarities) in facial expressions of emotion. *Current opinion in psychology*, *17*, 61-66.

Chen, N., Tanaka, K., Matsuyoshi, D., & Watanabe, K. (2016). Cross preferences for colors and shapes. *Color Research & Application*, *41*(2), 188-195.

Chen, Y., & Sundaram, H. (2005, October). Estimating complexity of 2D shapes. In *2005 IEEE 7th Workshop on Multimedia Signal Processing* (pp. 1-4). IEEE.

Chikhman, V., Bondarko, V., Danilova, M., Goluzina, A., & Shelepin, Y. (2012). Complexity of images: Experimental and computational estimates compared. *Perception*, *41*(6), 631-647.

Christensen, J. F., & Calvo-Merino, B. (2013). Dance as a subject for empirical aesthetics. *Psychology of Aesthetics, Creativity and the Arts, 7*(1), 76.

Christensen, J. F., Azevedo, R. T., & Tsakiris, M. (2021). Emotion matters: Different psychophysiological responses to expressive and non-expressive full-body movements. *Acta Psychologica,* 212, 103215.

Christensen, J. F., Gaigg, S. B., & Calvo-Merino, B. (2018). I can feel my heartbeat: Dancers have increased interoceptive accuracy. *Psychophysiology, 55*(4), e13008.

Christensen, J. F., Gaigg, S. B., Gomila, A., Oke, P., & Calvo-Merino, B. (2014a). Enhancing emotional experiences to dance through music: the role of valence and arousal in the cross-modal bias. *Frontiers in human neuroscience*, *8*, 757.

Christensen, J. F., Gomila, A., Gaigg, S. B., Sivarajah, N., & Calvo-Merino, B. (2016). Dance expertise modulates behavioral and psychophysiological responses to affective body movement. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(8), 1139.

Christensen, J. F., Lambrechts, A., & Tsakiris, M. (2019). The Warburg dance movement library—The WADAMO library: A validation study. *Perception*, *48*(1), 26-57.

Christensen, J. F., Nadal, M., Cela-Conde, C. J., & Gomila, A. (2014b). A norming study and library of 203 dance movements. *Perception*, *43*(2-3), 178-206.

Clark, D. M. (1983). On the induction of depressed mood in the laboratory: Evaluation and comparison of the Velten and musical procedures. *Advances in Behaviour Research and Therapy*, *5*(1), 27-49.

Clark, H. H., & Gerrig, R. J. (1990). Quotations as demonstrations. *Language*, 764-805.

Clarke, A. M., & Jack, B. (1998). The benefits of using qualitative research. *Professional Nurse (London, England)*, *13*(12), 845-847.

Clarke, T. J., Bradshaw, M. F., Field, D. T., Hampson, S. E., & Rose, D. (2005). The perception of emotion from body movement in point-light displays of interpersonal dialogue. *Perception*, *34*(10), 1171-1180.

Clavel, C., Plessier, J., Martin, J. C., Ach, L., & Morel, B. (2009, September). Combining facial and postural expressions of emotions in a virtual character. In *International Workshop on Intelligent Virtual Agents* (pp. 287-300). Springer, Berlin, Heidelberg.

Clore, G. L., & Schnall, S. (2005). The influence of affect on attitude. In D. Albarracín, B. T. Johnson, & M. P., Zanna (Eds.), *The handbook of attitudes* (pp. 437-492). Mahwah: Erlbaum.

Colquhoun, D. (2014). An investigation of the false discovery rate and the misinterpretation of p-values. *Royal Society open science*, *1*(3), 140216.

Colver, M. C., & El-Alayli, A. (2016). Getting aesthetic chills from music: The connection between openness to experience and frisson. *Psychology of Music*, *44*(3), 413-427.

Coombes, S. A., Tandonnet, C., Fujiyama, H., Janelle, C. M., Cauraugh, J. H., & Summers, J. J. (2009). Emotion and motor preparation: a transcranial magnetic stimulation study of corticospinal motor tract excitability. *Cognitive, Affective, & Behavioral Neuroscience*, *9*(4), 380-388.

Corradini, A., & Antonietti, A. (2013). Mirror neurons and their function in cognitively understood empathy. *Consciousness and cognition*, *22*(3), 1152-1161.

Cowie, R., & Cornelius, R. R. (2003). Describing the emotional states that are expressed in speech. *Speech communication*, *40*(1-2), 5-32.

Crane, E. A., & Gross, M. M. (2013). Effort-shape characteristics of emotion-related body movement. *Journal of Nonverbal Behavior*, *37*(2), 91-105.

Crane, E., & Gross, M. (2007, September). Motion capture and emotion: Affect detection in whole body movement. In *International Conference on Affective Computing and Intelligent Interaction* (pp. 95-101). Springer, Berlin, Heidelberg.

Cross, E. S. (2010). Building a dance in the human brain: Insights from expert and novice dancers. In *The Neurocognition of Dance* (pp. 187-212). Psychology Press.

Cross, E. S., & Smith, R. A. (forthcoming). The embodied neuroaesthetics of watching dance. In C. Fernandes, V. Evola & C. Ribero (Eds.), *Dance Data, Cognition, and Multimodal Communication* (Chapter 11). Routledge.

Cross, E. S., Hamilton, A. F. D. C., & Grafton, S. T. (2006). Building a motor simulation de novo: observation of dance by dancers. *Neuroimage*, *31*(3), 1257-1267.

Cross, E. S., Kirsch, L., Ticini, L. F., & Schütz-Bosbach, S. (2011). The impact of aesthetic evaluation and physical ability on dance perception. *Frontiers in human neuroscience*, 102.

Cross, E. S., Kraemer, D. J., Hamilton, A. F. D. C., Kelley, W. M., & Grafton, S. T. (2009). Sensitivity of the action observation network to physical and observational learning. *Cerebral cortex*, *19*(2), 315-326.

Crozier, W., & Greenhalgh, P. (1992). The empathy principle: Towards a model for the psychology of art. *Journal for the Theory of Social Behaviour*, *22*(1), 63-79.

Dael, N., Mortillaro, M., & Scherer, K. R. (2012). Emotion expression in body action and posture. *Emotion*, *12*(5), 1085.

Dahl, S., & Friberg, A. (2007). Visual perception of expressiveness in musicians' body movements. *Music Perception, 24*(5), 433-454.

Damasio, A. (2003). Feelings of emotion and the self. *Annals of the New York Academy of Sciences*, *1001*(1), 253-261.

Darda, K. M., & Cross, E. (2022a). Toward a general model of visual art perception-the role of expertise and culture. *Journal of Vision*, *21*(9), 1864-1864.

Darda, K. M., PhD, & Cross, E. S. (2021, September 2). A Unifying Model of Visual Art Appreciation: The Role of Expertise and Culture. https://doi.org/10.31234/osf.io/q5yzb

Darwin, C. (1965). *The expression of the emotions in man and animals.* Chicago: University of Chicago Press. (Original work published 1872).

Dayan, S. H. (2011). What is beauty, and why do we care so much about it?. *Archives of Facial Plastic Surgery*, *13*(1), 66-67.

De Bruyn, L., Moelants, D., & Leman, M. (2011). An embodied approach to testing musical empathy in subjects with an autism spectrum disorder. *Music and Medicine*, *4*(1), 28-36.

De Clercq, R. (2019). Aesthetic Pleasure Explained: De Clercq. *The Journal of Aesthetics and Art Criticism*, *77*(2), 121-132.

De Meijer, M. (1989). The contribution of general features of body movement to the attribution of emotions. *Journal of Nonverbal behavior*, *13*(4), 247-268.

Dekeyser, M., Verfaillie, K., & Vanrie, J. (2002). Creating stimuli for the study of biological-motion perception. *Behavior Research Methods, Instruments, & Computers*, *34*(3), 375-382.

Demenescu, L. R., Kortekaas, R., den Boer, J. A., & Aleman, A. (2010). Impaired attribution of emotion to facial expressions in anxiety and major depression. *PloS one*, *5*(12), e15058.

Demenescu, L. R., Mathiak, K. A., & Mathiak, K. (2014). Age-and gender-related variations of emotion recognition in pseudowords and faces. *Experimental Aging Research*, *40*(2), 187-207.

Demurie, E., De Corel, M., & Roeyers, H. (2011). Empathic accuracy in adolescents with autism spectrum disorders and adolescents with attention-deficit/hyperactivity disorder. *Research in Autism Spectrum Disorders*, *5*(1), 126-134.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Experimental brain research*, *91*(1), 176-180.

Dittrich, W. H., Troscianko, T., Lea, S. E., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception*, *25*(6), 727-738.

Dixson, B. J., Grimshaw, G. M., Ormsby, D. K., & Dixson, A. F. (2014). Eye-tracking women's preferences for men's somatotypes. *Evolution and Human Behavior*, *35*(2), 73-79.

Djikic, M., Oatley, K., & Moldoveanu, M. C. (2013). Reading other minds: Effects of literature on empathy. *Scientific study of literature*, *3*(1), 28-47.

Drolet, M., Schubotz, R. I., & Fischer, J. (2012). Authenticity affects the recognition of emotions in speech: behavioral and fMRI evidence. *Cognitive, Affective, & Behavioral Neuroscience*, *12*(1), 140-150.

Drouilly, M., Nattrass, N., & O'Riain, M. J. (2021). Beauty or beast? Farmers' dualistic views and the influence of aesthetic appreciation on tolerance towards black-backed jackal and caracal. *Plos one*, *16*(3), e0248977.

Du, G., Long, S., & Yuan, H. (2020). Non-contact emotion recognition combining heart rate and facial expression for interactive gaming environments. *IEEE Access*, *8*, 11896-11906.

Du, S., Tao, Y., & Martinez, A. M. (2014). Compound facial expressions of emotion. *Proceedings of the National Academy of Sciences*, *111*(15), E1454-E1462.

Ekman, P. (1992). Facial expressions of emotion: New findings, new questions. *Psychological Science, 3*(1), 34-38.

Ekman, P. (2009). Darwin's contributions to our understanding of emotional expressions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3449-3451.

Ekman, P., & Friesen, W. V. (1978). Facial action coding system. *Environmental Psychology & Nonverbal Behavior*.

Ekman, P., & Oster, H. (1979). Facial expressions of emotion. *Annual Review of Psychology, 30,* 527–554. https://doi.org/10.1146/annurev.ps.30.020179.002523

Ekman, P., Sorenson, E. R., & Friesen, W. V. (1969). Pan-cultural elements in facial displays of emotion. *Science*, *164*(3875), 86-88.

Elfenbein, H. A., & Ambady, N. (2002). On the universality and cultural specificity of emotion recognition: a meta-analysis. *Psychological bulletin*, *128*(2), 203.

Emerich, S., Lupu, E., & Apatean, A. (2009, August). Emotions recognition by speechand facial expressions analysis. In *2009 17th European signal processing conference* (pp. 1617-1621). IEEE.

Engelen, T., de Graaf, T. A., Sack, A. T., & de Gelder, B. (2015). A causal role for inferior parietal lobule in emotion body perception. *Cortex*, *73*, 195-202.

Engelen, T., Zhan, M., Sack, A. T., & Gelder, B. D. (2018). The influence of conscious and unconscious body threat expressions on motor evoked potentials studied with continuous flash suppression. *Frontiers in neuroscience*, *12*, 480.

Etcoff, N., Stock, S., Krumhuber, E. G., & Reed, L. I. (2021). A novel test of the Duchenne marker: Smiles after botulinum toxin treatment for crow's feet wrinkles. *Frontiers in psychology*, *11*, 3799.

Farley, F. H., & Dowling, P. M. (1972). Aesthetic preference in adolescents as a function of race and visual complexity. *Studies in Art Education*, *13*(2), 23-26.

Faughnder, R. (2021, March 18). Streaming milestone: Global subscriptions passed 1 billion last year. *Los Angeles Times.* https://www.latimes.com/entertainment-arts/business/story/2021-03-18/streaming-milestone-global-subscriptions-passed-1-billion-last-year-mpa-theme-report

Federman, D. J. (2011). Kinesthetic ability and the development of empathy in dance movement therapy. *Journal of Applied Arts & Health, 2*(2), 137-154.

Feist, G. J., & Brady, T. R. (2004). Openness to experience, non-conformity, and the preference for abstract art. *Empirical Studies of the Arts*, *22*(1), 77-89.

Fiori, F., Plow, E., Rusconi, M. L., & Cattaneo, Z. (2020). Modulation of corticospinal excitability during paintings viewing: a TMS study. *Neuropsychologia*, *149*, 107664.

Flood, A. (2021, January 25). Book sales defy pandemic to hit eight-year high. *The Guardian.* https://www.theguardian.com/books/2021/jan/25/bookshops-defy-pandemic-to-record-highest-sales-for-eight-years

Fox, R., & McDaniel, C. (1982). The perception of biological motion by human infants. *Science*, *218*(4571), 486-487.

Fridenson-Hayo, S., Berggren, S., Lassalle, A., Tal, S., Pigat, D., Bölte, S., Baron-Cohen, S., & Golan, O. (2016). Basic and complex emotion recognition in children with autism: cross-cultural findings. *Molecular autism*, *7*(1), 1-11.

Friedenberg, J., & Liby, B. (2016). Perceived beauty of random texture patterns: A preference for complexity. *Acta psychologica*, *168*, 41-49.

Frith, C. (2009). Role of facial expressions in social interactions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *364*(1535), 3453-3458.

Furnham, A., & Walker, J. (2001). Personality and judgements of abstract, pop art, and representational paintings. *European Journal of Personality*, *15*(1), 57-72.

Gardner, T., Goulden, N., & Cross, E. S. (2015). Dynamic modulation of the action observation network by movement familiarity. *Journal of Neuroscience*, *35*(4), 1561-1572.

Garrido, S. (2017). Towards A Model for Understanding Sad Music Listening. In *Why Are We Attracted to Sad Music?* (pp. 253-266). Palgrave Macmillan, Cham.

Garrido, S., & Schubert, E. (2011). Individual differences in the enjoyment of negative emotion in music: A literature review and experiment. *Music Perception*, *28*(3), 279-296.

Gauvrit, N., Soler-Toscano, F., & Guida, A. (2017). A preference for some types of complexity comment on "perceived beauty of random texture patterns: A preference for complexity". *Acta psychologica*, *174*, 48-53.

Gernot, G., Pelowski, M., & Leder, H. (2018). Empathy, Einfühlung, and aesthetic experience: The effect of emotion contagion on appreciation of representational and abstract art using fEMG and SCR. *Cognitive Processing, 19*(2), 147-165.

Gerson, S. A., Bekkering, H., & Hunnius, S. (2015). Short-term motor training, but not observational training, alters neurocognitive mechanisms of action processing in infancy. *Journal of Cognitive Neuroscience*, *27*(6), 1207-1214.

Giese, M. A., Thornton, I., & Edelman, S. (2008). Metrics of the perception of body movement. *Journal of Vision*, *8*(9), 13-13.

Golan, O., Baron-Cohen, S., Hill, J. J., & Rutherford, M. (2007). The 'Reading the Mind in the Voice'test-revised: a study of complex emotion recognition in adults with and without autism spectrum conditions. *Journal of autism and developmental disorders*, *37*(6), 1096-1106.

Gómez-Puerto, G., Rosselló, J., Corradi, G., Acedo-Carmona, C., Munar, E., & Nadal, M. (2018). Preference for curved contours across cultures. *Psychology of Aesthetics, Creativity, and the Arts*, *12*(4), 432.

Gouizi, K., Bereksi Reguig, F., & Maaoui, C. (2011). Emotion recognition from physiological signals. *Journal of medical engineering & technology*, *35*(6-7), 300-307.

Graham, D. J., Friedenberg, J. D., McCandless, C. H., & Rockmore, D. N. (2010, February). Preference for art: similarity, statistics, and selling price. In *Human Vision and Electronic Imaging XV* (Vol. 7527, pp. 377-386). SPIE.

Grezes, J., & Decety, J. (2001). Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis. *Human brain mapping*, *12*(1), 1-19.

Grezes, J., Pichon, S., & De Gelder, B. (2007). Perceiving fear in dynamic body expressions. *Neuroimage, 35*(2), 959-967.

Grosbras, M. H., Beaton, S., & Eickhoff, S. B. (2012). Brain regions involved in human movement perception: A quantitative voxel-based meta-analysis. *Human brain mapping*, *33*(2), 431-454.

Gross, J. J., & Levenson, R. W. (1995). Emotion elicitation using films. *Cognition & emotion*, *9*(1), 87-108.

Gross, M. M., Crane, E. A., & Fredrickson, B. L. (2010). Methodology for assessing bodily expression of emotion. *Journal of Nonverbal Behavior*, *34*(4), 223-248.

Gross, M. M., Crane, E. A., & Fredrickson, B. L. (2012). Effort-shape and kinematic assessment of bodily expression of emotion during gait. *Human movement science*, *31*(1), 202-221.

Gullberg, M., & Holmqvist, K. (2006). What speakers do and what addressees look at: Visual attention to gestures in human interaction live and on video. *Pragmatics & Cognition*, *14*(1), 53-82.

Gunns, R. E., Johnston, L., & Hudson, S. M. (2002). Victim selection and kinematics: A point-light investigation of vulnerability to attack. *Journal of Nonverbal Behavior, 26*(3), 129–158.

Hall, C., Hogue, T., & Guo, K. (2011). Differential gaze behavior towards sexually preferred and non-preferred human figures. *Journal of Sex Research*, *48*(5), 461-469.

Halovic, S., & Kroos, C. (2018). Not all is noticed: Kinematic cues of emotion-specific gait. *Human movement science*, *57*, 478-488.

Hancock, D. J., & Ste-Marie, D. M. (2013). Gaze behaviors and decision making accuracy of higher- and lower-level ice hockey referees. *Psychology of Sport and Exercise*, *14*(1), 66-71.

Hänggi, J., Koeneke, S., Bezzola, L., & Jäncke, L. (2010). Structural neuroplasticity in the sensorimotor network of professional female ballet dancers. *Human brain mapping*, *31*(8), 1196-1206.

Hasanzadeh, S., Esmaeili, B., & Dodd, M. D. (2018). Examining the relationship between construction workers' visual attention and situation awareness under fall and tripping hazard conditions: Using mobile eye tracking. *Journal of construction engineering and management*, *144*(7).

Haslam, S. A., Haslam, C., Cruwys, T., Jetten, J., Bentley, S. V., Fong, P., & Steffens, N. K. (2022). Social identity makes group-based social connection possible: Implications for loneliness and mental health. *Current opinion in psychology*, *43*, 161-165.

Hayes, A. E., Paul, M. A., Beuger, B., & Tipper, S. P. (2008). Self produced and observed actions influence emotion: The roles of action fluency and eye gaze. *Psychological research*, *72*(4), 461-472.

Heberlein, A. S., & Atkinson, A. P. (2009). Neuroscientific evidence for simulation and shared substrates in emotion recognition: beyond faces. *Emotion Review*, *1*(2), 162-177.

Heberlein, A. S., Adolphs, R., Tranel, D., & Damasio, H. (2004). Cortical regions for judgments of emotions and personality traits from point-light walkers. *Journal of Cognitive Neuroscience, 16*(7), 1143–1158.

Heron-Delaney, M., Quinn, P. C., Lee, K., Slater, A. M., & Pascalis, O. (2013). Nine-month-old infants prefer unattractive bodies over attractive bodies. *Journal of experimental child psychology*, *115*(1), 30-41.

Hess, U., & Thibault, P. (2009). Darwin and emotion expression. *American Psychologist*, *64*(2), 120.

Hevner, K. (1935). The affective character of the major and minor modes in music. *The American Journal of Psychology*, *47*(1), 103-118.

Hofmann, S. G., & Doan, S. N. (2018). Defining emotions. In S. G. Hofmann & S. N. Doan, *The social foundations of emotion: Developmental, cultural, and clinical dimensions* (pp. 11–21). American Psychological Association.

Holland, A. C., O'Connell, G., & Dziobek, I. (2021). Facial mimicry, empathy, and emotion recognition: a meta-analysis of correlations. *Cognition and Emotion*, *35*(1), 150-168.

Holmes, M. (2013). The force of face-to-face diplomacy: mirror neurons and the problem of intentions. *International organization*, *67*(4), 829-861.

Hoyle, H., Norton, B., Dunnett, N., Richards, J. P., Russell, J. M., & Warren, P. (2018). Plant species or flower colour diversity? Identifying the drivers of public and invertebrate response to designed annual meadows. *Landscape and Urban Planning*, *180*, 103-113.

Huis in 't Veld, E. M., Van Boxtel, G. J., & de Gelder, B. (2014a). The Body Action Coding System I: muscle activations during the perception and expression of emotion. *Social neuroscience*, *9*(3), 249-264.

Huis In't Veld, E. M., van Boxtel, G. J., & de Gelder, B. (2014b). The Body Action Coding System II: muscle activations during the perception and expression of emotion. *Frontiers in behavioral neuroscience*, *8*, 330.

Hung, S. M., Nieh, C. H., & Hsieh, P. J. (2016). Unconscious processing of facial attractiveness: Invisible attractive faces orient visual attention. *Scientific Reports*, *6*(1), 1-8.

Ishizu, T., & Zeki, S. (2017). The experience of beauty derived from sorrow. *Human Brain Mapping*, *38*(8), 4185-4200.

Israelashvili, J., Sauter, D., & Fischer, A. (2020). Two facets of affective empathy: concern and distress have opposite relationships to emotion recognition. *Cognition and Emotion, 34*(6), 1112-1122.

Izard, C. E. (1994). Innate and universal facial expressions: Evidence from developmental and cross-cultural research. *Psychological Bulletin, 115*, 288-299. http://dx.doi.org/10.1037/0033-2909.115.2.288

Jack, R. E. (2013). Culture and facial expressions of emotion. *Visual Cognition*, *21*(9-10), 1248-1286.

Jack, R. E., Blais, C., Scheepers, C., Schyns, P. G., & Caldara, R. (2009). Cultural confusions show that facial expressions are not universal. *Current biology*, *19*(18), 1543-1548.

Jack, R. E., Caldara, R., & Schyns, P. G. (2012). Internal representations reveal cultural diversity in expectations of facial expressions of emotion. *Journal of Experimental Psychology: General*, *141*(1), 19.

Jack, R. E., Garrod, O. G., & Schyns, P. G. (2014). Dynamic facial expressions of emotion transmit an evolving hierarchy of signals over time. *Current biology*, *24*(2), 187-192.

Jack, R. E., Garrod, O. G., Yu, H., Caldara, R., & Schyns, P. G. (2012). Facial expressions of emotion are not culturally universal. *Proceedings of the National Academy of Sciences*, *109*(19), 7241-7244.

Jack, R. E., Sun, W., Delis, I., Garrod, O. G., & Schyns, P. G. (2016). Four not six: Revealing culturally common facial expressions of emotion. *Journal of Experimental Psychology: General*, *145*(6), 708.

Jack, R., Garrod, O., Yu, H., Caldara, R., & Schyns, P. (2012). Beyond Darwin: Revealing Culture--specificities in the Temporal Dynamics of 4D Facial Expressions. *Journal of Vision*, *12*(9), 971-971.

Jacob, P. (2008). What do mirror neurons contribute to human social cognition?. *Mind & Language*, *23*(2), 190-223.

James, W. (1884). What is an emotion? *Mind, 9*, 188–205.

Jankowski, T., Francuz, P., Oleś, P., & Chmielnicka-Kuter, E. (2020). The effect of temperament, expertise in art, and formal elements of paintings on their aesthetic appraisal. *Psychology of Aesthetics, Creativity, and the Arts*, *14*(2), 209. Sachs, Damasio & Habbibi, 2015

Järveläinen, J., Schürmann, M., Avikainen, S., & Hari, R. (2001). Stronger reactivity of the human primary motor cortex during observation of live rather than video motor acts. *Neuroreport, 12*(16), 3493-3495.

Jehna, M., Neuper, C., Ischebeck, A., Loitfelder, M., Ropele, S., Langkammer, C., Ebner, F., Fuchs, S., Schmidt, R., Fazekas, F., & Enzinger, C. (2011). The functional correlates of face perception and recognition of emotional facial expressions as evidenced by fMRI. *Brain Research*, *1393*, 73-83.

Joffe, H. (2012). Thematic analysis. *Qualitative research methods in mental health and psychotherapy: A guide for students and practitioners*, *1*, 210-223.

Johansson, G. (1975). Visual motion perception. *Scientific American, 232*(6), 76-89.

Johansson, M., & Laike, T. (2007). Intention to respond to local wind turbines: the role of attitudes and visual perception. *Wind energy: an international journal for progress and applications in wind power conversion technology*, *10*(5), 435-451.

Johnson, K. L., Gill, S., Reichman, V., & Tassinary, L. G. (2007). Swagger, sway, and sexuality: Judging sexual orientation from body motion and morphology. *Journal of Personality and Social Psychology, 93*(3), 321.

Johnson, W. F., Emde, R. N., Scherer, K. R., & Klinnert, M. D. (1986). Recognition of emotion from vocal cues. *Archives of General Psychiatry*, *43*(3), 280-283.

Jola, C., & Grosbras, M. H. (2013). In the here and now: Enhanced motor corticospinal excitability in novices when watching live compared to video recorded dance. *Cognitive Neuroscience*, *4*(2), 90-98.

Jola, C., Abedian-Amiri, A., Kuppuswamy, A., Pollick, F. E., & Grosbras, M. H. (2012). Motor simulation without motor expertise: enhanced corticospinal excitability in visually experienced dance spectators. *PloS one*, *7*(3), e33343.

Jola, C., Davis, A., & Haggard, P. (2011). Proprioceptive integration and body representation: insights into dancers' expertise. *Experimental Brain Research*, *213*(2), 257-265.

Jola, C., Pollick, F. E., & Calvo-Merino, B. (2014). "Some like it hot": spectators who score high on the personality trait openness enjoy the excitement of hearing dancers breathing without music. *Frontiers in Human Neuroscience*, *8*, 718.

Jospe, K., Flöel, A., & Lavidor, M. (2018). The interaction between embodiment and empathy in facial expression recognition. *Social cognitive and affective neuroscience*, *13*(2), 203-215.

Kadiri, S. R., Gangamohan, P., Gangashetty, S. V., & Yegnanarayana, B. (2015). Analysis of excitation source features of speech for emotion recognition. In *Sixteenth annual conference of the international speech communication association*.

Kainz, H., Graham, D., Edwards, J., Walsh, H. P., Maine, S., Boyd, R. N., Lloyd, D.G., Modenese, L., & Carty, C. P. (2017). Reliability of four models for clinical gait analysis. *Gait & posture*, *54*, 325-331.

Kaniušonytė, G., Truskauskaitė-Kunevičienė, I., Žukauskienė, R., & Crocetti, E. (2019). Knowing who you are for not feeling lonely? A longitudinal study on identity and loneliness. *Child Development*, *90*(5), 1579-1588.

Kettlewell, N., Lipscomb, S., Evans, L., & Rosston, K. (1990). The effect of subject matter and degree of realism on aesthetic preferences for paintings. *Empirical Studies of the Arts*, *8*(1), 85-93.

Keysers, C., & Gazzola, V. (2006). Towards a unifying neural theory of social cognition. *Progress in brain research*, *156*, 379-401.

Khachatryan, H., Rihn, A., Hansen, G., & Clem, T. (2020). Landscape aesthetics and maintenance perceptions: Assessing the relationship between homeowners' visual attention and landscape care knowledge. *Land Use Policy*, *95*, 104645.

Kietzmann, T. C., Geuter, S., & König, P. (2011). Overt visual attention as a causal factor of perceptual awareness. *PloS one*, *6*(7), e22614.

Killgore, W. D., Cloonan, S. A., Taylor, E. C., & Dailey, N. S. (2020). Loneliness: A signature mental health concern in the era of COVID-19. *Psychiatry research, 290*, 113117.

Kim, J. H. (2015). Kinaesthetic Empathy as Aesthetic Experience of Music. A Phenomenological Approach. *Les Cahiers Philosophiques de Strasbourg*, (38), 119-138.

Kirsch, L. P., & Cross, E. S. (2015). Additive routes to action learning: layering experience shapes engagement of the action observation network. *Cerebral Cortex*, *25*(12), 4799-4811.

Kirsch, L. P., & Cross, E. S. (2018). The influence of sensorimotor experience on the aesthetic evaluation of dance across the life span. *Progress in Brain Research*, *237*, 291-316.

Kirsch, L. P., Dawson, K., & Cross, E. S. (2015). Dance experience sculpts aesthetic perception and related brain circuits. *Annals of the New York Academy of Sciences*, *1337*(1), 130-139.

Kirsch, L. P., Diersch, N., Sumanapala, D. K., & Cross, E. S. (2018). Dance training shapes action perception and its neural implementation within the young and older adult brain. *Neural Plasticity*, *2018*.

Kirsch, L. P., Drommelschmidt, K. A., & Cross, E. S. (2013). The impact of sensorimotor experience on affective evaluation of dance. *Frontiers in human neuroscience*, *7*, 521.

Kirsch, L. P., Snagg, A., Heerey, E., & Cross, E. S. (2016). The impact of experience on affective responses during action observation. *Plos one*, *11*(5), e0154681.

Kleinginna, P. R., & Kleinginna, A. M. (1981). A categorized list of emotion definitions, with suggestions for a consensual definition. *Motivation and emotion*, *5*(4), 345-379.

Kleinman, J., Marciano, P. L., & Ault, R. L. (2001). Advanced theory of mind in high-functioning adults with autism. *Journal of autism and developmental disorders*, *31*(1), 29-36.

Knuuttila, S. (2014). Emotions from Plato to the Renaissance. In *Sourcebook for the History of the Philosophy of Mind* (pp. 463-497). Springer, Dordrecht.

Koch, S. C., Mehl, L., Sobanski, E., Sieber, M., & Fuchs, T. (2015). Fixing the mirrors: A feasibility study of the effects of dance movement therapy on young adults with autism spectrum disorder. *Autism, 19*(3), 338-350.

Kok, F. M., Groen, Y., Becke, M., Fuermaier, A. B., & Tucha, O. (2016). Self-reported empathy in adult women with autism spectrum disorders–a systematic mini review. *PloS one*, *11*(3), e0151568.

Konrath, S., Corneille, O., Bushman, B. J., & Luminet, O. (2014). The relationship between narcissistic exploitativeness, dispositional empathy, and emotion recognition abilities. *Journal of Nonverbal Behavior*, *38*(1), 129-143.

Konstan, D. (2006). The Concept of" Emotion" from Plato to Cicero. *Méthexis*, *19*, 139-151.

Kosonogov, V. (2012). Why the mirror neurons cannot support action understanding. *Neurophysiology*, *44*(6), 499-502.

Kothari, R., Skuse, D., Wakefield, J., & Micali, N. (2013). Gender differences in the relationship between social communication and emotion recognition. *Journal of the American Academy of Child & Adolescent Psychiatry*, *52*(11), 1148-1157.

Kourmousi, N., Amanaki, E., Tzavara, C., Merakou, K., Barbouni, A., & Koutras, V. (2017). The toronto empathy questionnaire: reliability and validity in a nationwide sample of greek teachers. *Social Sciences*, *6*(2), 62.

Krch, D. (2011). Cognitive Processing. In: Kreutzer, J.S., DeLuca, J., Caplan, B. (Eds.), *Encyclopedia of Clinical Neuropsychology*. Springer, New York, NY.

Kret, M. E., Pichon, S., Grèzes, J., & de Gelder, B. (2011). Similarities and differences in perceiving threat from dynamic faces and bodies. An fMRI study. *Neuroimage*, *54*(2), 1755-1762.

Kreutz, G., Ott, U., Teichmann, D., Osawa, P., & Vaitl, D. (2008). Using music to induce emotions: Influences of musical preference and absorption. *Psychology of music*, *36*(1), 101-126.

Kreutz, G., Schubert, E., & Mitchell, L. A. (2008). Cognitive styles of music listening. *Music Perception*, *26*(1), 57-73.

Krüger, B., Kaletsch, M., Pilgramm, S., Schwippert, S. S., Hennig, J., Stark, R., Lis, S., Gallhofer, B., Sammer, G., Zentgraf, K., & Munzert, J. (2018). Perceived intensity of emotional point–light displays is reduced in subjects with ASD. *Journal of autism and developmental disorders*, *48*(1), 1-11.

Krumhuber, E. G., & Manstead, A. S. (2009). Can Duchenne smiles be feigned? New evidence on felt and false smiles. *Emotion*, *9*(6), 807.

Kuchibhotla, S., Vankayalapati, H. D., Vaddi, R. S., & Anne, K. R. (2014). A comparative analysis of classifiers in emotion recognition through acoustic features. *International Journal of Speech Technology*, *17*(4), 401-408.

Kuppens, P., Tuerlinckx, F., Russell, J. A., & Barrett, L. F. (2013). The relation between valence and arousal in subjective experience. *Psychological bulletin, 139*(4), 917.

Lange, C. G. (1885). The mechanism of the emotions. In *The classical psychologists* (pp. 672–684).

Langner, R., & Eickhoff, S. B. (2013). Sustaining attention to simple tasks: a meta-analytic review of the neural mechanisms of vigilant attention. *Psychological bulletin*, *139*(4), 870.

Leder, H., Bär, S., & Topolinski, S. (2012). Covert painting simulations influence aesthetic appreciation of artworks. *Psychological Science*, *23*(12), 1479-1481.

Leder, H., Ring, A., & Dressler, S. G. (2013). See me, feel me! Aesthetic evaluations of art portraits. *Psychology of Aesthetics, Creativity, and the Arts*, *7*(4), 358.

Leder, H., Tinio, P. P., & Bar, M. (2011). Emotional valence modulates the preference for curved objects. *Perception*, *40*(6), 649-655.

Lee, W. (2020, April 21). Record high Netflix subscriptions in coronavirus crisis. *Los Angeles Times.* https://www.latimes.com/entertainment-arts/business/story/2020-04-21/netflix-usage-profits-surge-during-coronavirus-crisis

Leehr, E., Giel, K. E., Schaeffeler, N., Mack, I., Thiel, A., Zurstiege, G., & Zipfel, S. (2018). Where do you look visual attention to human bodies across the weight spectrum in individuals with normal weight or with obesity. *Obesity facts*, *11*(4), 277-286.

Le-Hoang, P. V. (2020). Direct and indirect impact of aesthetics on intention to buy smartphones. *Independent Journal of Management & Production*, *11*(7), 2685-2708.

Leighton, S. R. (1982). Aristotle and the Emotions. *Phronesis*, 144-174.

Levy, J. A., & Duke, M. P. (2003). The Use of Laban Movement Analysis in the Study of Personality, Emotional State and Movement Style: An Exploratory Investigation of the Veridicality of" Body Language". *Individual Differences Research*, *1*(1).

Li, M., Lafarge, F., & Marlet, R. (2020). Approximating shapes in images with low-complexity polygons. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 8633-8641).

Liew, S. L., Sheng, T., Margetis, J. L., & Aziz-Zadeh, L. (2013). Both novelty and expertise increase action observation network activity. *Frontiers in human neuroscience*, *7*, 541.

Lim, M. H., Allen, K. A., Furlong, M. J., Craig, H., & Smith, D. C. (2021). Introducing a dual continuum model of belonging and loneliness. *Australian Journal of Psychology*, *73*(1), 81-86.

Limbrecht-Ecklundt, K., Scheck, A., Jerg-Bretzke, L., Walter, S., Hoffmann, H., & Traue, H. C. (2013). The effect of forced choice on facial emotion recognition: a comparison to open verbal classification of emotion labels. *GMS Psycho-Social-Medicine*, *10*.

Little, A. C., Burt, D. M., & Perrett, D. I. (2006). What is good is beautiful: Face preference reflects desired personality. *Personality and Individual Differences*, *41*(6), 1107-1118.

Locher, P. (2006). The usefulness of eye movement recordings to subject an aesthetic episode with visual art to empirical scrutiny. *Psychology Science*, *48*(2), 106.

Lockhofen, D. E. L., & Mulert, C. (2021). Neurochemistry of visual attention. *Frontiers in Neuroscience*, *15*, 643597.

Lorenzo, G. L., Biesanz, J. C., & Human, L. J. (2010). What is beautiful is good and more accurately understood: Physical attractiveness and accuracy in first impressions of personality. *Psychological science*, *21*(12), 1777-1782.

Lu, X., Suryanarayan, P., Adams Jr, R. B., Li, J., Newman, M. G., & Wang, J. Z. (2012, October). On shape and the computability of emotions. In *Proceedings of the 20th ACM international conference on Multimedia* (pp. 229-238).

Ma, Y., Paterson, H. M., & Pollick, F. E. (2006). A motion capture library for the study of identity, gender, and emotion perception from biological motion. *Behavior research methods*, *38*(1), 134-141.

Mandy, W. P., Charman, T., & Skuse, D. H. (2012). Testing the construct validity of proposed criteria for DSM-5 autism spectrum disorder. *Journal of the American Academy of Child & Adolescent Psychiatry*, *51*(1), 41-50.

Manera, V., Becchio, C., Cavallo, A., Sartori, L., & Castiello, U. (2011). Cooperation or competition? Discriminating between social intentions by observing prehensile movements. *Experimental brain research*, *211*(3), 547-556.

Mar, R. A., Oatley, K., Hirsh, J., Dela Paz, J., & Peterson, J. B. (2006). Bookworms versus nerds: Exposure to fiction versus non-fiction, divergent associations with social ability, and the simulation of fictional social worlds. *Journal of research in personality*, *40*(5), 694-712.

Marin, M. M., Lampatz, A., Wandl, M., & Leder, H. (2016). Berlyne revisited: Evidence for the multifaceted nature of hedonic tone in the appreciation of paintings and music. *Frontiers in human neuroscience*, *10*, 536.

Marković, S. (2010). Aesthetic experience and the emotional content of paintings. *Psihologija*, *43*(1), 47-64.

Massin, O. (2018). Bitter Joys and Sweet Sorrow. In C. Tappolet, F. Teroni & A. Konzelmann (Eds.), *Shadows of the Soul: Philosophical Perspectives on Negative Emotions* (pp. 50–59). Routledge.

Mastandrea, S., Bartoli, G., & Carrus, G. (2011). The automatic aesthetic evaluation of different art and architectural styles. *Psychology of Aesthetics, Creativity, and the Arts*, *5*(2), 126.

Mastrominico, A., Fuchs, T., Manders, E., Steffinger, L., Hirjak, D., Sieber, M., Thomas, E., Holzinger, A., Konrad, A., Bopp, N., & Koch, S. C. (2018). Effects of dance movement therapy on adult patients with autism spectrum disorder: A randomized controlled trial. *Behavioral Sciences, 8*(7), 61.

Masuda, M., & Kato, S. (2010, September). Motion rendering system for emotion expression of human form robots based on laban movement analysis. In *19th international symposium in robot and human interactive communication* (pp. 324-329). IEEE.

Matsumoto, D., & Hwang, H. S. C. (2019). Culture, emotion, and expression. *Cross-Cultural Psychology: Contemporary Themes and Perspectives*, 501-515.

Matsumoto, D., Keltner, D., Shiota, M. N., O'Sullivan, M., & Frank, M. (2008). Facial expressions of emotion. In M. Lewis, J. M. Haviland-Jones, & L. F. Barrett (Eds.), *Handbook of emotions* (pp. 211–234). The Guilford Press.

Mazurek, K. A., Rouse, A. G., & Schieber, M. H. (2018). Mirror neuron populations represent sequences of behavioral epochs during both execution and observation. *Journal of Neuroscience*, *38*(18), 4441-4455.

Mazzoni, N., Landi, I., Ricciardelli, P., Actis-Grosso, R., & Venuti, P. (2020). Motion or emotion? Recognition of emotional bodily expressions in children with autism spectrum disorder with and without intellectual disability. *Frontiers in Psychology*, *11*, 478.

McAleer, P., Todorov, A., & Belin, P. (2014). How do you say 'Hello'? Personality impressions from brief novel voices. *PloS one*, *9*(3), e90779.

McCrae, R. R. (2007). Aesthetic chills as a universal marker of openness to experience. *Motivation and Emotion*, *31*(1), 5-11.

McGarry, L. M., & Russo, F. A. (2011). Mirroring in dance/movement therapy: Potential mechanisms behind empathy enhancement. *The Arts in Psychotherapy, 38*(3), 178-184.

McLellan, T., Johnston, L., Dalrymple-Alford, J., & Porter, R. (2010). Sensitivity to genuine versus posed emotion specified in facial displays. *Cognition and Emotion*, *24*(8), 1277-1292.

Melzer, A., Shafir, T., & Tsachor, R. P. (2019). How do we recognize emotion from movement? Specific motor components contribute to the recognition of each emotion. *Frontiers in psychology*, 1389.

Ménard, M., Richard, P., Hamdi, H., Daucé, B., & Yamaguchi, T. (2015, February). Emotion Recognition based on Heart Rate and Skin Conductance. In *Proceedings of the 2nd International Conference on Physiological Computing Systems* (pp. 26-32).

Menting-Henry, S., Hidalgo-Lopez, E., Aichhorn, M., Kronbichler, M., Kerschbaum, H., & Pletzer, B. (2022). Oral contraceptives modulate the relationship between resting brain activity, amygdala connectivity and emotion recognition–a resting state fMRI study. *Frontiers in Behavioral Neuroscience*, *16*.

Merriam-Webster(a). (n.d.). Aesthetic. In *Merriam-Webster*. Retrieved January 3, 2023, from https://www.merriam-webster.com/dictionary/aesthetic

Merriam-Webster(b). (n.d.). Emotion. In *Merriam-Webster*. Retrieved April 1, 2022, from https://www.merriam-webster.com/dictionary/emotion

Metcalfe, D., McKenzie, K., McCarty, K., & Pollet, T. V. (2019). Emotion recognition from body movement and gesture in children with Autism Spectrum Disorder is improved by situational cues. *Research in developmental disabilities*, *86*, 1-10.

Michaelis, K., Wiener, M., & Thompson, J. C. (2014). Passive listening to preferred motor tempo modulates corticospinal excitability. *Frontiers in Human Neuroscience*, *8*, 252.

Michalak, J., Troje, N. F., Fischer, J., Vollmar, P., Heidenreich, T., & Schulte, D. (2009). Embodiment of sadness and depression—gait patterns associated with dysphoric mood. *Psychosomatic medicine*, *71*(5), 580-587.

Mignault, A., & Chaudhuri, A. (2003). The many faces of a neutral face: Head tilt and perception of dominance and emotion. *Journal of nonverbal behavior*, *27*(2), 111-132.

Mileva, M., Kramer, R. S., & Burton, A. M. (2019). Social evaluation of faces across gender and familiarity. *Perception*, *48*(6), 471-486.

Mitchell, R. W., & Curry, C. (2016). Self-recognition and other-recognition in point-light displays. *Open Journal of Philosophy, 6*(01), 42.

Monroy, E., Imada, T., Sagiv, N., & Orgs, G. (2022). Dance across cultures: Joint action aesthetics in Japan and the UK. *Empirical Studies of the Arts*, *40*(2), 209-227.

Montepare, J., Koff, E., Zaitchik, D., & Albert, M. (1999). The use of body movements and gestures as cues to emotions in younger and older adults. *Journal of Nonverbal Behavior*, *23*(2), 133-152.

Moran, A., Byrne, A., & McGlade, N. (2002). The effects of anxiety and strategic planning on visual search behaviour. *Journal of sports sciences*, *20*(3), 225-236.

Morris, M. W., & Keltner, D. (2000). How emotions work: The social functions of emotional expression in negotiations. *Research in organizational behavior*, *22*, 1-50.

Müller, M., Höfel, L., Brattico, E., & Jacobsen, T. (2010). Aesthetic judgments of music in experts and laypersons—An ERP study. *International Journal of Psychophysiology*, *76*(1), 40-51.

Munafò, M. (2016). Open science and research reproducibility. *ecancermedicalscience, 10*.

Myers, G. (1999). Functions of reported speech in group discussions. *Applied Linguistics*, *20*(3), 376-401.

Näsänen, R., Kukkonen, H., & Rovamo, J. (1993). Spatial integration of band-pass filtered patterns in noise. *Vision Research*, *33*(7), 903-911.

Neal, A., & Kilner, J. M. (2010). What is simulated in the action observation network when we observe actions?. *European Journal of Neuroscience*, *32*(10), 1765-1770.

Neale, C., Griffiths, A., Chalmin-Pui, L. S., Mendu, S., Boukhechba, M., & Roe, J. (2021). Color aesthetics: A transatlantic comparison of psychological and physiological impacts of warm and cool colors in garden landscapes. *Wellbeing, Space and Society*, *2*, 100038.

Neumann, D., Zupan, B., Malec, J. F., & Hammond, F. (2014). Relationships between alexithymia, affect recognition, and empathy after traumatic brain injury. *The Journal of head trauma rehabilitation, 29*(1), E18-E27.

Nummenmaa, L., Smirnov, D., Lahnakoski, J. M., Glerean, E., Jääskeläinen, I. P., Sams, M., & Hari, R. (2014). Mental action simulation synchronizes action–observation circuits across individuals. *Journal of Neuroscience*, *34*(3), 748-757.

Olderbak, S., Sassenrath, C., Keller, J., & Wilhelm, O. (2014). An emotion-differentiated perspective on empathy with the emotion specific empathy questionnaire. *Frontiers in Psychology*, *5*, 653.

Oliver, W., & Risner, D. (2017). An introduction to dance and gender. *Dance and Gender: An Evidence-Based Approached*, 1-19.

Open Science Collaboration. (2012). An open, large-scale, collaborative effort to estimate the reproducibility of psychological science. *Perspectives on Psychological Science, 7*(6), 657-660.

Orgs, G. Calvo-Merino, B & Cross, E. S. (2018). Knowing dance or knowing how to dance? Sources of expertise in aesthetic appreciation of human movement. In: B. Bläsing; M. Puttke and T. Schack, (Eds). *The Neurocognition of Dance: Mind, Movement and Motor Skills* (pp. 238-257). Routledge.

Orgs, G., Caspersen, D., & Haggard, P. (2016). You move, I watch, it matters: Aesthetic communication in dance. In S. S. Obhi & E. S. Cross (Eds.), *Shared representations: Sensorimotor foundations of social life* (pp. 627–653). Cambridge University Press. https://doi.org/10.1017/CBO9781107279353.031

Orgs, G., Hagura, N., & Haggard, P. (2013). Learning to like it: aesthetic perception of bodies, movements and choreographic structure. *Consciousness and Cognition*, *22*(2), 603-612.

Orlandi, A., & Proverbio, A. M. (2019). Bilateral engagement of the occipito-temporal cortex in response to dance kinematics in experts. *Scientific reports*, *9*(1), 1-14.

Orlandi, A., Cross, E. S., & Orgs, G. (2020). Timing is everything: Dance aesthetics depend on the complexity of movement kinematics. *Cognition*, *205*, 104446.

Oyibo, K., & Vassileva, J. (2017, July). What drives perceived usability in mobile web design: classical or expressive aesthetics?. In *International Conference of Design, User Experience, and Usability* (pp. 445-462). Springer, Cham.

Pasch, M., & Poppe, R. (2007). Person or puppet? the role of stimulus realism in attributing emotion to static body postures. In *International Conference on Affective Computing and Intelligent Interaction* (pp. 83-94). Springer, Berlin, Heidelber.

Paterson, H. M., Pollick, F. E., & Sanford, A. J. (2001). The role of velocity in affect discrimination. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 23, No. 23).

Patuano, A. (2018). Measuring naturalness and complexity using the fractal dimensions of landscape photographs. *Journal of Digital Landscape Architecture*, *3*, 328-335.

Pepper, K. L., Demetriou, E. A., Park, S. H., Boulton, K. A., Hickie, I. B., Thomas, E. E., & Guastella, A. J. (2019). Self-reported empathy in adults with autism, early psychosis, and social anxiety disorder. *Psychiatry Research*, *281*, 112604.

Peters, D. (2015). Musical Empathy, Emotional Co-Constitution, and the" Musical Other". *Empirical Musicology Review*, *10*(1-2), 2-15.

Petrakis, E. (1987). Analysis of visual search patterns of dance teachers. *Journal of Teaching in Physical Education*, *6*(2), 149-156.

Petrides, K. V., Niven, L., & Mouskounti, T. (2006). The trait emotional intelligence of ballet dancers and musicians. *Psicothema, 18*, 101-107.

Petrini, K., Dahl, S., Rocchesso, D., Waadeland, C. H., Avanzini, F., Puce, A., & Pollick, F. E. (2009). Multisensory integration of drumming actions: musical expertise affects perceived audiovisual asynchrony. *Experimental brain research*, *198*(2), 339-352.

Petrini, K., Pollick, F. E., Dahl, S., McAleer, P., McKay, L., Rocchesso, D., Waadeland, C. H., Love, S., Avanzini, F., & Puce, A. (2011). Action expertise reduces brain activity for audiovisual matching actions: an fMRI study with expert drummers. *Neuroimage*, *56*(3), 1480-1492.

Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage*, *16*(2), 331-348.

Philip, R. C., Whalley, H. C., Stanfield, A. C., Sprengelmeyer, R., Santos, I. M., Young, A. W., Atkinson, A. P., Calder, A. J., Johnstone, E. C., Lawrie, S. M., & Hall, J. (2010). Deficits in facial, body movement and vocal emotional processing in autism spectrum disorders. *Psychological medicine*, *40*(11), 1919-1929.

Phillips, M. L., Young, A. W., Senior, C., Brammer, M., Andrew, C., Calder, A. J., Bullmore, E. T., Perrett, D. I., Rowland, D., Williams, S. C. R., Gray, J. A., & David, A. S. (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature*, *389*(6650), 495-498.

Pilarska, A. (2018). Big-Five personality and aspects of the self-concept: Variable-and person-centered approaches. *Personality and Individual Differences*, *127*, 107-113.

Piwek, L., Petrini, K., & Pollick, F. (2016). A dyadic stimulus set of audiovisual affective displays for the study of multisensory, emotional, social interactions. *Behavior research methods*, *48*(4), 1285-1295.

Piwowar, H. A. (2011). Who shares? Who doesn't? Factors associated with openly archiving raw research data. *PloS one*, *6*(7), e18657.

Pollick, F. E., Kay, J. W., Heim, K., & Stringer, R. (2005). Gender recognition from point-light walkers. *Journal of Experimental Psychology: Human Perception and Performance, 31*(6), 1247.

Pollick, F. E., Paterson, H. M., Bruderlin, A., & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition*, *82*(2), B51-B61.

Polo, C., Lund, K., Plantin, C., & Niccolai, G. P. (2016). Group emotions: The social and cognitive functions of emotions in argumentation. *International Journal of Computer-Supported Collaborative Learning*, *11*(2), 123-156.

Posner, J., Russell, J. A., & Peterson, B. S. (2005). The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and psychopathology, 17*(3), 715-734.

Press, C., Weiskopf, N., & Kilner, J. M. (2012). Dissociable roles of human inferior frontal gyrus during action execution and observation. *Neuroimage*, *60*(3), 1671-1677.

Quintana, D. S., Guastella, A. J., Outhred, T., Hickie, I. B., & Kemp, A. H. (2012). Heart rate variability is associated with emotion recognition: Direct evidence for a relationship between the autonomic nervous system and social cognition. *International journal of psychophysiology*, *86*(2), 168-172.

Rawlings, D., Barrantes i Vidal, N., & Furnham, A. (2000). Personality and aesthetic preference in Spain and England: Two studies relating sensation seeking and openness to experience to liking for paintings and music. *European journal of personality*, *14*(6), 553-576.

Reason, M., & Reynolds, D. (2010). Kinesthesia, empathy, and related pleasures: An inquiry into audience experiences of watching dance. *Dance research journal*, *42*(2), 49-75.

Reber, R., Winkielman, P., & Schwarz, N. (1998). Effects of perceptual fluency on affective judgments. *Psychological science*, *9*(1), 45-48.

Remington, N. A., Fabrigar, L. R., & Visser, P. S. (2000). Reexamining the circumplex model of affect. *Journal of personality and social psychology, 79*(2), 286.

Richardson, T., Elliott, P., & Roberts, R. (2017). Relationship between loneliness and mental health in students. *Journal of Public Mental Health*, *16*(2), 48-54.

Rieffe, C., Ketelaar, L., & Wiefferink, C. H. (2010). Assessing empathy in young children: Construction and validation of an Empathy Questionnaire (EmQue). *Personality and individual differences*, *49*(5), 362-367.

Riek, L. D., Adams, A., & Robinson, P. (2011, March). Exposure to cinematic depictions of robots and attitudes towards them. In *Proceedings of international conference on human-robot interaction, workshop on expectations and intuitive human-robot interaction* (Vol. 6).

Rimé, B., Corsini, S., & Herbette, G. (2002). Emotion, verbal expression, and the social sharing of emotion. In *The verbal communication of emotions* (pp. 193-216). Psychology Press.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.*, *27*, 169-192.

Rizzolatti, G., & Craighero, L. (2005). Mirror neuron: a neurological approach to empathy. In *Neurobiology of human values* (pp. 107-123). Springer, Berlin, Heidelberg.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive brain research*, *3*(2), 131-141.

Robins, V., Wood, P. J., & Sheppard, A. P. (2011). Theory and algorithms for constructing discrete Morse complexes from grayscale digital images. *IEEE Transactions on pattern analysis and machine intelligence*, *33*(8), 1646-1658.

Roether, C. L., Omlor, L., Christensen, A., & Giese, M. A. (2009). Critical features for the perception of emotion from gait. *Journal of Vision, 9*(6), 15–15.

Rose, D., Müllensiefen, D., Lovatt, P., & Orgs, G. (2020). The Goldsmiths Dance Sophistication Index (Gold-DSI): A psychometric tool to assess individual differences in dance experience. *Psychology of Aesthetics, Creativity, and the Arts.* Advance online publication. https://doi.org/10.1037/aca0000340

Rosenthal, G. G., Flores Martinez, T. Y., García de León, F. J., & Ryan, M. J. (2001). Shared preferences by predators and females for male ornaments in swordtails. *The American Naturalist*, *158*(2), 146-154.

Ross, P., Polson, L., & Grosbras, M. H. (2012). Developmental changes in emotion recognition from full-light and point-light displays of body movement. *PLoS ONE.*, *7*(9), e44815.

Rova, M. (2017). Embodying kinaesthetic empathy through interdisciplinary practice-based research. *The Arts in Psychotherapy*, *55*, 164-173.

Russell, J. A. (1980). A circumplex model of affect. *Journal of personality and social psychology, 39*(6), 1161.

Russell, J. A. (1993). Forced-choice response format in the study of facial expression. *Motivation and Emotion*, *17*(1), 41-51.

Rutherford, M. D., Baron-Cohen, S., & Wheelwright, S. (2002). Reading the mind in the voice: A study with normal adults and adults with Asperger syndrome and high functioning autism. *Journal of autism and developmental disorders*, *32*(3), 189-194.

Saffari, E., Hosseini, S. R., Taheri, A., & Meghdari, A. (2021). "Does cinema form the future of robotics?": a survey on fictional robots in sci-fi movies. *SN Applied Sciences*, *3*(6), 1-13.

Sams, J. (2010). Quoting the unspoken: An analysis of quotations in spoken discourse. *Journal of Pragmatics*, *42*(11), 3147-3160.

Sandelowski, M. (1996). One is the liveliest number: The case orientation of qualitative research. *Research in nursing & health*, *19*(6), 525-529.

Santos, I., & Young, A. (2005). Exploring the perception of social characteristics in faces using the isolation effect. *Visual Cognition*, *12*(1), 213-247.

Sato, W., & Aoki, S. (2006). Right hemispheric dominance in processing of unconscious negative emotion. *Brain and cognition*, *62*(3), 261-266.

Sawada, M., Suda, K., & Ishii, M. (2003). Expression of emotions in dance: Relation between arm movement characteristics and emotion. *Perceptual and motor skills*, *97*(3), 697-708.

Schachter, S., & Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review, 69*(5), 379.

Scheff, T. (2015). Toward defining basic emotions. *Qualitative Inquiry*, *21*(2), 111-121.

Scherer, K. R. (2005). What are emotions? And how can they be measured?. *Social science information*, *44*(4), 695-729.

Scherer, K. R. (2013). Vocal markers of emotion: Comparing induction and acting elicitation. *Computer Speech & Language*, *27*(1), 40-58.

Schlotz, W., Wallot, S., Omigie, D., Masucci, M. D., Hoelzmann, S. C., & Vessel, E. A. (2020). The Aesthetic Responsiveness Assessment (AReA): A screening tool to assess individual differences in responsiveness to art in English and German. *Psychology of Aesthetics, Creativity, and the Arts*.

Schmidt, N. B., Richey, J. A., Zvolensky, M. J., & Maner, J. K. (2008). Exploring human freeze responses to a threat stressor. *Journal of Behavior Therapy and Experimental Psychiatry, 39*(3), 292–304.

Schneider, S., Christensen, A., Häußinger, F. B., Fallgatter, A. J., Giese, M. A., & Ehlis, A. C. (2014). Show me how you walk and I tell you how you feel—a functional near-infrared spectroscopy study on emotion perception based on human gait. *NeuroImage, 85*, 380–390.

Schreiber, M. (2017). Audiences, aesthetics and affordances analysing practices of visual communication on social media. *Digital Culture & Society*, *3*(2), 143-164.

Seidel, A., & Prinz, J. (2018). Great works: A reciprocal relationship between spatial magnitudes and aesthetic judgment. *Psychology of Aesthetics, Creativity, and the Arts*, *12*(1), 2.

Seifert, T. (2018). *The Relationship between Proprioceptive Acuity, Emotion Recognition and Emotional Empathy*. University of Haifa (Israel).

Shafir, T., Tsachor, R. P., & Welch, K. B. (2016). Emotion regulation through movement: unique sets of movement characteristics are associated with and enhance basic emotions. *Frontiers in psychology*, *6*, 2030.

Shankar, A., Rafnsson, S. B., & Steptoe, A. (2015). Longitudinal associations between social connections and subjective wellbeing in the English Longitudinal Study of Ageing. *Psychology & health*, *30*(6), 686-698.

Shikanai, N., Sawada, M., & Ishii, M. (2013). Development of the movements impressions emotions model: Evaluation of movements and impressions related to the perception of emotions in dance. *Journal of Nonverbal Behavior*, *37*(2), 107-121.

Shimada, S., & Hiraki, K. (2006). Infant's brain responses to live and televised action. *Neuroimage, 32*(2), 930-939.

Sievers, B., Polansky, L., Casey, M., & Wheatley, T. (2013). Music and movement share a dynamic structure that supports universal expressions of emotion. *Proceedings of the National Academy of Sciences*, *110*(1), 70-75.

Silvera, D. H., Josephs, R. A., & Giesler, R. B. (2002). Bigger is better: The influence of physical size on aesthetic preference judgments. *Journal of Behavioral Decision Making*, *15*(3), 189-202.

Silvia, P. J., & Barona, C. M. (2009). Do people prefer curved objects? Angularity, expertise, and aesthetic preference. *Empirical studies of the arts*, *27*(1), 25-42.

Silvia, P. J., & Berg, C. (2011). Finding movies interesting: How appraisals and expertise influence the aesthetic experience of film. *Empirical Studies of the Arts*, *29*(1), 73-88.

Silvia, P. J., Fayn, K., Nusbaum, E. C., & Beaty, R. E. (2015). Openness to experience and awe in response to nature and music: personality and profound aesthetic experiences. *Psychology of Aesthetics, Creativity, and the Arts*, *9*(4), 376.

Sinigaglia, C. (2008). Mirror neurons: This is the question. *Journal of Consciousness Studies*, *15*(10-11), 70-92.

Smith, M. L., Cottrell, G. W., Gosselin, F., & Schyns, P. G. (2005). Transmitting and decoding facial expressions. *Psychological science*, *16*(3), 184-189.

Smith, R. A., & Cross, E. S. (2022). The McNorm library: creating and validating a new library of emotionally expressive whole body dance movements. *Psychological research*, 1-25.

Smith, R. A., & Cross, E. S. (under review). Moving Me, Moving You: Emotional expressivity, empathy, and prior experience shape whole-body movement preferences. *Psychology of Aesthetics, Creativity, and the Arts*.

Smith, R. A., & Pollick, F. E. (forthcoming). The Role of Dance Experience, Visual Processing Strategies, and Quantitative Movement Features in Recognition of Emotion from Whole-Body Movements. In C. Fernandes, V. Evola & C. Ribero (Eds.), *Dance Data, Cognition, and Multimodal Communication* (Chapter 11). Routledge.

Sofer, C., Dotsch, R., Wigboldus, D. H., & Todorov, A. (2015). What is typical is good: The influence of face typicality on perceived trustworthiness. *Psychological Science*, *26*(1), 39-47.

Soto, J. A., & Levenson, R. W. (2009). Emotion recognition across cultures: the influence of ethnicity on empathic accuracy and physiological linkage. *Emotion, 9*(6), 874.

Spaulding, S. (2013). Mirror neurons and social cognition. *Mind & Language*, *28*(2), 233-257.

Spreng, R. N., McKinnon, M. C., Mar, R. A., & Levine, B. (2009). The Toronto Empathy Questionnaire: Scale development and initial validation of a factor-analytic solution to multiple empathy measures. *Journal of personality assessment*, *91*(1), 62-71.

Stevens, C., Winskel, H., Howell, C., Vidal, L. M., Latimer, C., & Milne-Home, J. (2010). Perceiving dance schematic expectations guide experts' scanning of a contemporary dance film. *Journal of Dance Medicine & Science*, *14*(1), 19-25.

Sun, G., Wong, Y., Cheng, Z., Kankanhalli, M. S., Geng, W., & Li, X. (2020). DeepDance: music-to-dance motion choreography with adversarial learning. *IEEE Transactions on Multimedia*, *23*, 497-509.

Sun, Y., & Yin, L. (2008, October). Facial expression recognition based on 3D dynamic range model sequences. In *European Conference on Computer Vision* (pp. 58-71). Springer, Berlin, Heidelberg.

Sundar, S. S., Waddell, T. F., & Jung, E. H. (2016, March). The Hollywood robot syndrome media effects on older adults' attitudes toward robots and adoption intentions. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 343-350). IEEE.

Susskind, J. M., Hinton, G. E., Movellan, J. R., & Anderson, A. K. (2008). Generating facial expressions with deep belief nets. *Affective Computing, Emotion Modelling, Synthesis and Recognition*, 421-440.

Sze, J. A., Gyurak, A., Yuan, J. W., & Levenson, R. W. (2010). Coherence between emotional experience and physiology: does body awareness training have an impact?. *Emotion*, *10*(6), 803.

Tan, K. W., Tiddeman, B., & Stephen, I. D. (2018). Skin texture and colour predict perceived health in Asian faces. *Evolution and Human Behavior*, *39*(3), 320-335.

Ticini, L. F., Rachman, L., Pelletier, J., & Dubal, S. (2014). Enhancing aesthetic appreciation by priming canvases with actions that match the artist's painting style. *Frontiers in Human Neuroscience*, *8*, 391.

Ticini, L. F., Urgesi, C., & Calvo-Merino, B. (2015). Embodied aesthetics: insight from cognitive neuroscience of performing arts. In *Aesthetics and the embodied mind: Beyond art theory and the Cartesian mind-body dichotomy* (pp. 103-115). Springer, Dordrecht.

Todorov, A., & Duchaine, B. (2008). Reading trustworthiness in faces without recognizing faces. *Cognitive Neuropsychology*, *25*(3), 395-410.

Todorova, G. K., Hatton, R. E. M., & Pollick, F. E. (2019). Biological motion perception in autism spectrum disorder: a meta-analysis. *Molecular autism*, *10*(1), 1-28.

Tomkins, S. S. (2008). *Affect imagery consciousness: the complete edition: two volumes*. Springer publishing company.

Tonks, J., Williams, W. H., Frampton, I., Yates, P., & Slater, A. (2007). Assessing emotion recognition in 9–15-years olds: Preliminary analysis of abilities in reading emotion from faces, voices and eyes. *Brain injury*, *21*(6), 623-629.

Torrents, C., Castañer, M., Jofre, T., Morey, G., & Reverter, F. (2013). Kinematic parameters that influence the aesthetic perception of beauty in contemporary dance. *Perception*, *42*(4), 447-458.

Tracy, J. L., & Robins, R. W. (2004). Show your pride: Evidence for a discrete emotion expression. *Psychological science*, *15*(3), 194-197.

Tramacere, A., Pievani, T., & Ferrari, P. F. (2017). Mirror neurons in the tree of life: mosaic evolution, plasticity and exaptation of sensorimotor matching responses. *Biological Reviews*, *92*(3), 1819-1841.

Trimmer, E., McDonald, S., & Rushby, J. A. (2017). Not knowing what I feel: Emotional empathy in autism spectrum disorders. *Autism*, *21*(4), 450-457.

Tripathi, S., Tripathi, S., & Beigi, H. (2018). Multi-modal emotion recognition on iemocap dataset using deep learning. *arXiv preprint arXiv:1804.05788*.

Trutoiu, L. C., Carter, E. J., Pollard, N., Cohn, J. F., & Hodgins, J. K. (2014). Spatial and temporal linearities in posed and spontaneous smiles. *ACM Transactions on Applied Perception (TAP)*, *11*(3), 1-15.

Tschacher, W., Greenwood, S., Kirchberg, V., Wintzerith, S., van den Berg, K., & Tröndle, M. (2012). Physiological correlates of aesthetic perception of artworks in a museum. *Psychology of Aesthetics, Creativity, and the Arts*, *6*(1), 96.

Tsiourti, C., Weiss, A., Wac, K., & Vincze, M. (2019). Multimodal integration of emotional signals from voice, body, and context: Effects of (in) congruence on emotion recognition and attitudes towards robots. *International Journal of Social Robotics*, *11*(4), 555-573.

Val-Calvo, M., Álvarez-Sánchez, J. R., Ferrández-Vicente, J. M., Díaz-Morcillo, A., & Fernández-Jover, E. (2020). Real-time multi-modal estimation of dynamically evoked emotions using EEG, heart rate and galvanic skin response. *International journal of neural systems*, *30*(04), 2050013.

Valentine, C. W. (2015). *The experimental psychology of beauty*. Routledge.

Van den Tol, A. J., Edwards, J., & Heflick, N. A. (2016). Sad music as a means for acceptance-based coping. *Musicae Scientiae*, *20*(1), 68-83.

Van der Schyff, D., & Krueger, J. (2019). Musical empathy, from simulation to 4E interaction. *Music, Sound & Mind, ed. AF Corrêa,(Rio de Janeiro: ABCM)*.

Van Dyck, E., Burger, B., & Orlandatou, K. (2017). The communication of emotions in dance. In *The Routledge companion to embodied music interaction* (pp. 122-130). Routledge.

Van Dyck, E., Maes, P. J., Hargreaves, J., Lesaffre, M., & Leman, M. (2012). The impact of induced emotions on free movement. In *12th International conference on Music Perception and Cognition (ICMPC 2012); 8th Triennial conference of the European Society for the Cognitive Sciences of Music (ESCOM 2012)* (pp. 1050-1056). Aristotle University of Thessaloniki. School of Music Studies.

Van Dyck, E., Moelants, D., Demey, M., Deweppe, A., Coussement, P., & Leman, M. (2012). The impact of the bass drum on human dance movement. *Music Perception: An Interdisciplinary Journal*, *30*(4), 349-359.

Van Dyck, E., Vansteenkiste, P., Lenoir, M., Lesaffre, M., & Leman, M. (2014). Recognizing induced emotions of happiness and sadness from dance movement. *PloS one*, *9*(2), e89773.

Van Dyke, J. (2012). Questioning trends in university dance. *Journal of Dance Education*, *12*(2), 31-36.

Van Kleef, G. A., Cheshin, A., Fischer, A. H., & Schneider, I. K. (2016). The social nature of emotions. *Frontiers in Psychology*, *7*, 896.

Van Paasschen, J., Bacci, F., & Melcher, D. P. (2015). The influence of art expertise and training on emotion and preference ratings for representational and abstract artworks. *PloS one*, *10*(8), e0134241.

Vanrie, J., & Verfaillie, K. (2004). Perception of biological motion: A stimulus set of human point-light actions. *Behavior Research Methods, Instruments, & Computers*, *36*(4), 625-629.

Vasileiou, K., Barnett, J., Thorpe, S., & Young, T. (2018). Characterising and justifying sample size sufficiency in interview-based studies: systematic analysis of qualitative health research over a 15-year period. *BMC medical research methodology*, *18*(1), 1-18.

Vernon, P. E. (1930). Synaesthesia in music. *Psyche*.

Verschuere, B., Crombez, G., & Koster, E. (2001). The international affective picture system. *Psycholo Belgica, 41,* 205-217.

Vicary, S., Sperling, M., Von Zimmermann, J., Richardson, D. C., & Orgs, G. (2017). Joint action aesthetics. *Plos one*, *12*(7), e0180101.

Vlami, V., Danek, J., Zogaris, S., Gallou, E., Kokkoris, I. P., Kehayias, G., & Dimopoulos, P. (2020). Residents' views on landscape and ecosystem services during a wind farm proposal in an island protected area. *Sustainability*, *12*(6), 2442.

Voos, A., Cordeaux, C., Tirrell, J., & Pelphrey, K. (2013). Developmental neuroscience of social perception. *Neural Circuit Development and Function in the Healthy and Diseased Brain: Comprehensive Developmental Neuroscience*, *3*, 351.

Vuoskoski, J. K., Thompson, W. F., McIlwain, D., & Eerola, T. (2011). Who enjoys listening to sad music and why?. *Music Perception*, *29*(3), 311-317.

Wallbott, H. G. (1998). Bodily expression of emotion. *European journal of social psychology*, *28*(6), 879-896.

Wallenborn, A. (2011). *Visual Attention in Dressage Judges*.

Wang, K., Zeng, X., Yang, J., Meng, D., Zhang, K., Peng, X., & Qiao, Y. (2018, October). Cascade attention networks for group emotion recognition with face, body and image cues. In *Proceedings of the 20th ACM International Conference on Multimodal Interaction* (pp. 640-645).

Wang, L., Xia, L., & Zhang, D. (2017). Face-body integration of intense emotional expressions of victory and defeat. *Plos one*, *12*(2), e0171656.

Wanzer, D. L., Finley, K. P., Zarian, S., & Cortez, N. (2020). Experiencing flow while viewing art: Development of the Aesthetic Experience Questionnaire. *Psychology of Aesthetics, Creativity, and the Arts*, *14*(1), 113.

Wegemer, C. M. (2020). Selflessness, depression, and neuroticism: An interactionist perspective on the effects of self-transcendence, perspective-taking, and materialism. *Frontiers in psychology*, *11*, 2556.

Weichselbaum, H., Leder, H., & Ansorge, U. (2018). Implicit and explicit evaluation of visual symmetry as a function of art expertise. *i-Perception*, *9*(2), 2041669518761464.

Weld, H. P. (1912). An experimental study of musical enjoyment. *The American Journal of Psychology*, *23*(2), 245-308.

White, S., & Egermann, H. (2021). Do Free Dance Movements Communicate How We Feel? Investigating Emotion Recognition in Dance. *Jahrbuch Musikpsychologie*, *30*, 1-21.

Whiting, C. M., Kotz, S. A., Gross, J., Giordano, B. L., & Belin, P. (2020). The perception of caricatured emotion in voice. *Cognition*, *200*, 104249.

Wilkinson, K. M., & Light, J. (2011). Preliminary Investigation of Visual Attention to Human Figures in Photographs: Potential Considerations for the Design of Aided AAC Visual Scene Displays. *Journal of Speech, Language & Hearing Research*, *54*(6).

Wilkinson, Z., Cunningham, R., & Elliott, M. A. (2021). The influence of empathy on the perceptual response to visual art. *Psychology of Aesthetics, Creativity, and the Arts*.

Williams, E. H., & Cross, E. S. (2018). Decreased reward value of biological motion among individuals with autistic traits. *Cognition*, *171*, 1-9.

Williams, E. H., Cristino, F., & Cross, E. S. (2019). Human body motion captures visual attention and elicits pupillary dilation. *Cognition*, *193*, 104029.

Xu, Y., Zhang, N., Wei, P., Sang, G., Li, L., & Yuan, F. (2020). Deep neural framework with visual attention and global context for predicting image aesthetics. *IEEE Access*.

Yalta, K., Ozturk, S., & Yetkin, E. (2016). Golden Ratio and the heart: A review of divine aesthetics. *International Journal of Cardiology*, *214*, 107-112.

Yeh, Y. C., Hsu, W. C., & Li, P. H. (2018). The modulation of personal traits in neural responses during the aesthetic experience of mundane art. *Trends in Neuroscience and Education*, *10*, 8-18.

Yeung, M. K. (2021). A systematic review and meta-analysis of facial emotion recognition in autism spectrum disorder: The specificity of deficits and the role of task characteristics. *Neuroscience & Biobehavioral Reviews*, 104518.

Yu, H., & Winkler, S. (2013, July). Image complexity and spatial information. In *2013 Fifth International Workshop on Quality of Multimedia Experience (QoMEX)* (pp. 12-17). IEEE.

Zhang, T., Zheng, W., Cui, Z., Zong, Y., & Li, Y. (2018). Spatial–temporal recurrent neural network for emotion recognition. *IEEE transactions on cybernetics*, *49*(3), 839-847.

Zhang, X., Yin, L., Cohn, J. F., Canavan, S., Reale, M., Horowitz, A., & Liu, P. (2013, April). A high-resolution spontaneous 3d dynamic facial expression database. In *2013 10th IEEE international conference and workshops on automatic face and gesture recognition (FG)* (pp. 1-6). IEEE.

Supplementary Table S.1: A summary of movement characteristics, identified in the previous kinematic analysis literature, which are considered to be associated with particular emotional states.

Authors	Type of Motion	Kinematics Methodology	Kinematics Findings
Crane & Gross (2007)	Whole Body, Walking Motions	Computational	Joy: fast, high amplitude elbow flexion Anger: fast, increased step frequency Sad: slow, head oriented down, decreased motion in upper body, lower range of motion
Crane & Gross (2013)	Whole Body, Walking Motions	LMA Efforts	Joy: expansion, direct, strong, fast, bound Anger: direct, strong, bound Sad: contraction, indirect, light, slow, relaxed
Wallbott (1998)	Whole Body	Subjective Coding	Elated Joy: raised shoulders, head oriented backwards, raised arms (upward and frontal), opening and closing hands, high movement activity, expansion Happy: low movement activity Cold Anger: arms stretched (frontal), lateralised hand movements

			Hot Anger: raised shoulders, arms stretched (frontal), lateralised hand movements, pointing gestures Sadness: collapsed upper body, low movement activity Despair: shoulders forward, hands opening and closing, expansive Fear: hands opening and closing Terror: arms stretched (sides), high movement dynamics
Montepare, Koff, Zaitchik & Albert (1999)	Whole Body	Subjective Coding	Happy: jerky, loose, fast, soft, expansive, high movement dynamics Angry: jerky, stiff, fast, hard, expansion, high movement dynamics Sad: smooth, loose, slow, soft, contracted, low movement dynamics Neutral: smooth, loose, slow, soft, contracted, low movement dynamics
Castellano, Villalba & Camurri (2007)	Raising and Lowering Arms (in Coronal Plane)	Computational	Joy: high quantity of motion, expansion Pleasure: low quantity of motion, expansion Anger: high quantity of motion, contraction

			Sad: low quantity of motion, contraction
Bernhardt & Robinson (2007)	Whole Body	Computational	Anger: energetic, forceful Sad: slow, slack
Sawada, Suda & Ishii (2003)	Arm Movements, Dance Motions	Computational	Joy: high travel distance Anger: high acceleration Sad: low travel distance, low acceleration
Roether, Omlor, Christensen & Giese (2009)	Whole Body, Walking Motions	Computational	Happy: fast, high movement dynamics Anger: fast, high movement dynamics, angular arrangement of body segments Sad: slow, low movement dynamics, head oriented down Fear: slow, low movement dynamics, reduced motion in lower body (knees, hips)

Gross, Crane & Fredrickson (2012)	Whole Body, Walking Motions	LMA Efforts & Computational	Joy: fast, high amplitude of joint motion (hip, shoulder, elbow, pelvis), high torso motion, torso expansion, shoulder depression Anger: fast, high amplitude of joint motion (hip, shoulder, elbow, pelvis), high torso motion Sad: slow, high flexion (neck, torso)
Gross, Crane & Fredrickson (2010)	Arm Movements, Knocking Motion	LMA Efforts & Computational	Joy: fast, expansion, relaxed Content: slow Proud: fast Angry: fast, expansion, strong Sad: slow, tense Anxious: slow, expansive, slow, tense
Camurri, Lagerlöf & Volpe (2003)	Whole Body, Dance Motion	Subjective Coding & Computational	Joy: expansion, high quantity of motion, high fluidity Angry: high quantity of motion, low fluidity

			Grief: slow, contraction, low quantity of motion, low fluidity, frequent transitions between movements Fear: contraction, low fluidity, more movement phases
Shafir, Tsachor & Welch (2016)	Whole Body	LMA Efforts	Happy: free, light, spreading, rising, upward-accented motion, jumping, rhythmic Angry: strong, sudden, direct, advancing Sad: passive weight (limp), sinking, head oriented down, arms close to upper body Fear: bound, retreating, condensing, enclosing, upper body oriented backward
Halovic & Kroos (2018)	Whole Body, Walking Motion	Subjective Coding & Computational	Happy: upward-accented motion, high quantity of motion (arms), long strides Angry: fast, heavy, long strides, high quantity of motion (arms) Sad: slow, slouching, short strides, low quantity of motion (arms) Fear: mixed speed (fast and slow), short strides, low quantity of motion (arms)
Shikanai, Sawada & Ishii (2013)	Whole Body, Dance Motion	Subjective Coding	Joy: expansion, high stability, turning, jumping, high quantity of motion, expansion (limbs), body closing

			Anger: high stability, high quantity of motion, body closing Sad: low quantity of motion, contraction, low stability, body closing
Masuda & Kato (2010)	Whole Body	LMA Efforts	Pleasure: fast, strong, expansion, reclining posture, upward-accented motion Relaxation: slow, weak, mixed posture (reclining and straight) Anger: fast, strong, expansion, movement in various directions Sad: slow, weak, contraction, forward-tilted posture, downward-accented motion
Smith & Pollick (in Preparation)	Whole Body, Dance Motion	LMA Efforts	Happy: fast, Sad: slow

Supplementary Table S.2: A brief overview of the methods and results of previous emotion recognition, and stimulus library validation experiments. This table includes the following details: the authors and date of publication, the sample size used in the experiment, the nature of the participants (e.g., university students, dancers), the type of movements explored (e.g., dance, walking - and where possible whether gestures were included or excluded), the duration of the stimuli, the way the body was represented in the stimuli (e.g., full-light displays, point-light displays, avatars), the method for collecting and measuring emotion recognition, the emotions explored in the experiment, and recognition rates reported in the papers. It should be noted that recognition rates (expressed as percentage) were not always available. Some authors used different measures (e.g., Crane & Gross (2013) used mean decoding accuracy expressed as a decimal value; Christensen et al., (2016), (2019) and (2021) used a 0-100 slider scale for the extent of expressivity, and Dahl & Friberg (2007) used a 6-point Likert scale to measure the extent of expressivity of each specific emotion), and some authors did not provide emotion-specific recognition rates (e.g., Atkinson, Tunstall & Dittrich (2007) show recognition rates for each emotion in Figure 2, but the specific values were not available in the text).

Authors	Sample Size	Nature of Participants	Movement Type	Stimuli Duration	Experimental Design	Chance Level Recognition	Stimuli Type	Emotion Category	Recognition Rate
Atkinson, Dittrich, Gemmell & Young (2004)	N = 36	University Students	Free Movement *(including gestural cues)*	6 seconds	Forced-choice emotion recognition task	20%	Point-Light Videos & Full-Light Videos	*Point-Light Videos* Anger Disgust Fear Happiness Sadness	 71.94% 63.06%■ 79.72% 84.17%▲ 82.22%

								<u>*Full-Light Videos*</u> Anger Disgust Fear Happiness Sadness	 85.55% 75.28%■ 91.11%▲ 86.67% 86.94%
Crane & Gross (2007)	N = 60	University Students *(who performed the walking motions)*	Walking motions after emotion elicitation phase	Various durations	Binary choice for agreement with the felt emotion (agreement or disagreement)	50%	Point-Light Videos	Neutral Sad Content Joy Anger	83%▲ 76% 74% 67% 62%■
Crane & Gross (2013)	N = 42	University Students *(who performed the walking motions)*	Walking motions after emotion elicitation phase	Various durations	Forced-choice emotion recognition task	.10	Full-Light Videos	Anger Content Joy Neutral Sad	.23■ .23■ .24 .25 .43▲ *(mean decoding accuracies)*
Gross, Crane &	N= 30	University Students	Walking motions	Various durations	Forced-choice emotion recognition task	10%	Full-Light Videos	Anger Joy Sad	22% 20% 43%▲

Fredrickson (2012)								Content Neutral	19%■ 25%
Montepare, Koff, Zaitchik & Albert (1999)	N = 41	University Students *(N = 20)* & Older Adults *(N = 21)*	Dramatized depictions of emotions *(including gestural cues)*	3 seconds	Forced-choice emotion recognition task	25%	Full-Light Videos	*University Students* Neutral Happy Angry Sad *Older Adults* Neutral Happy Angry Sad	 95.74%▲ 73.68%■ 90.99% 82.41% 92.34%▲ 71.92% 78.66% 68.77%■
Pasch & Poppe (2007)	N = 48	University Students & Staff	Static Images of Body Postures	N/A	Forced-choice emotion recognition task	16.67%	Full-Light 3D Renderings of Human Bodies and Human-Like Wooden Avatars	*Human Body Renderings* Anger Disgust Fear Happiness Sadness Surprise	 32% 9%■ 36% 57% 58%▲ 32%

								Wooden Avatars Anger Disgust Fear Happiness Sadness Surprise	 39% 24% 23% 62%▲ 54% 16%■
Bernhardt & Robinson (2007)	N/A	Machine Learning Algorithm	Everyday Motions *(e.g., knocking, throwing, lifting)*	Various durations	Machine Learning Algorithm based on Movement Kinematics	25%	Point-Light Displays	*Using Unbiased Features* Neutral Happy Angry Sad *Using Biased Features* Neutral Happy Angry Sad	 37.9%■ 41.1% 59.1% 62.4%▲ 74.2% 65.3%■ 92.4%▲ 92.4%▲
Dael, Mortillaro &	N/A	Algorithm based on	Emotions Portrayed in Scenario-Based	Various Durations	Algorithm based on	8.33%	Full-Light Videos	Amusement Pride Elated Joy	70% 50% 90%▲

Scherer (2012)		Movement Characteristics	Interaction Settings by Actors		Movement Characteristics			Panic Fear Anxiety Despair Sadness Hot Anger Irritation Interest Pleasure Relief	60% 50% 40% 30%■ 70% 30%■ 60% 60% 50%
Roether, Omlor, Christensen & Giese (2009)	N = 21	University Students	Walking motion	Various Durations	Forced-choice emotion recognition task	25%	3D Avatar Renderings	Anger Happiness Neutral Fear Sadness	76.0% 65.1%■ 71.5% 80.0% 92.0%▲
Gross, Crane & Fredrickson (2010)	N = 35	University Students	Knocking motion after emotion elicitation phase	Various durations	Intensity scores for each emotion including the target emotion on a 5-point Likert scale (classified as recognised if	N/A	Full-Light Videos	*Across All Clips (Emotions Felt and Unfelt)* Angry Anxious Sad Joyful Proud Content	 78%▲ 65% 56% 39% 56% 33%■

					the target emotion was rated 'moderately' (3) or above in intensity			*Across Only Clips where Emotion was Felt* Angry Anxious Content Joy Proud Sad	 8%▲ 6% 3%■ 3%■ 6% 6%
Camurri, Lagerlöf & Volpe (2003)	N = 32	University Students	Modern Dance movements *(specifically excluding gestures)*	Various durations	Forced-choice emotion recognition task & Intensity Ratings for All Emotions	25%	Full-Light Videos	*Across All Participants and Portrayals* Anger Joy Fear Grief	 60.6% 55.0% 39.8%■ 70.4%▲
Melzer, Shafir & Tsachor (2019)	N = 62	General Population	Free movement *(including gestural cues)*	3 seconds	Forced-choice emotion recognition task	20.7% (for the entire sample)	Full-Light Videos	Happy Sadness Fear Anger Neutral	81.39%▲ 78.57% 51.15% 47.29%■ 67.42%

Alaerts et al., (2011)	N = 32 (after technical issues)	University Students	Everyday Motions *(e.g., walking, jumping, kicking)*	3 seconds	Forced-choice emotion recognition task (happier', 'sadder, 'angrier' or 'not-different' to a neutral clip)	25%	Point-Light Videos	Neutral Happy Sad Angry	54.3% 44.2%■ 45.8% 58.6%▲
Atkinson, Tunstall & Dittrich (2007)	N = 32	University Students	Free movement *(including gestural cues)*	3 seconds	Forced-choice emotion recognition task	16.67%	Full-Light Videos & Point-Light Videos	<u>*Full-Light Videos*</u> Overall <u>*Point-Light Videos*</u> Overall *Note: No means for emotion-specific recognition rates were available in the paper	 84.2% 81.7% From Figure 2, the general order of recognition rates (from highest to lowest) was

									1. Sadness▲ 2. Happiness 3. Neutral 4. Anger 5. Fear 6. Disgust■
*Bachmann, Zabicki, Munzert & Kr*üger (2020)	N = 30	General Population	Emotions Portrayed in Interaction Settings by Actors	4 seconds	Forced-choice emotion recognition task	20%	Point-Light Videos	***Full-Body Stimuli*** Overall Anger Happiness Sadness Affection ***Arms-Only Stimuli*** Overall ***Trunk-Only Stimuli*** Overall	 80.7% 87.8%▲ 85.6% 82.1% 67.6%■ 68.9% 57.8%

Christensen et al., (2016)	N = 44	University Students *(N = 24)* & Ballet Dancers *(N = 20)*	Ballet Dance Movements	5-6 seconds	Slider Scale Response from 'Very Sad' (0) to 'Very Happy' (100)	Slider Scale Value for Perfect Recognition of Happy = 100 Slider Scale Value for Perfect Recognition of Sad = 0	Full-Light Videos	Happy Sad	60.54 39.68
Christensen, Azevedo & Tsakiris (2021)	N = 41	General Population *(with varying levels of prior dance experience)*	Western Ballet and Contemporary Dance Movements *(Emotionally Expressive-Positive, Emotionally Expressive-Negative &*	6-8 seconds	Slider Scale Response from 'Not at all (expressive)' (0) to 'Very (expressive)' (100)	Slider Scale Value for Perfect Recognition of Expression = 100 Slider Scale Value for Perfect	Full-Light Videos	Expressive Contemporary Not-Expressive Contemporary Expressive Ballet Not-Expressive Ballet	61.88 43.34 57.22 46.40

			Non-Expressive)			Recognition of Non-Expression = 0			
Christensen, Lambrechts & Tsakiris (2019)	N=80 (in each experiment)		Western Ballet and Contemporary Dance Movements *(Emotionally Expressive-Positive, Emotionally Expressive-Negative & Non-Expressive)*	6 seconds	Slider Scale Response from 'Not at all (expressive)' (0) to 'Very (expressive)' (100)	Slider Scale Value for Perfect Recognition of Expression = 100 Slider Scale Value for Perfect Recognition of Non-Expression = 0	Full-Light Videos	Expressive-Positive Contemporary Expressive-Negative Contemporary Not-Expressive Contemporary Expressive-Positive Ballet Expressive-Negative Ballet Not-Expressive Ballet	57.99 56.16 50.46 58.49 56.51 46.22

Dittrich, Troscianko, Lea & Morgan (1996)	N = 72	University Students	Emotional Dance Movements	5 seconds	Rating Questionnaire (numbers assigned to emotions based on likelihood of being the intended emotion)	Rescaled values range from 0 (least likely) to 5 (most likely) where 5 is Perfect Recognition of the Intended Emotion Chance = 2.5	Full-Light Videos & Point-Light Videos	*Full-Light Videos* Surprise Fear Anger Disgust Grief Joy *Point-Light Videos* Surprise Fear Anger Disgust Grief Joy	 4.80▲ 3.92 4.43 3.50■ 4.65 4.79 3.11 3.02■ 3.06 2.77 3.72▲ 3.69
Grèzes, Pichon & de Gelder (2007) [Pilot Study]	N = 12	Information Unavailable	Emotions Portrayed in Scenarios by Actors	3 seconds	Forced-choice emotion Recognition Task	33.33%	Full-Light Videos	Fear Neutral	79%■ 88%▲

Dahl & Friberg (2007)	N = 20	University Students	Musician Playing the Marimba with Different Emotion Expressions	30-50 seconds	Rating the Expression of Each Emotion on a 6-Point Scale from 'Nothing' (0) to "Very much' (6)	N/A	Full-Light Videos	Happy Sad Angry Fearful *Note: No exact values for emotion-specific recognition were available in the paper	~ 4.0▲ ~ 4.5 ~ 3.5 ~ 2.0■

Where ■ denotes the emotion with the lowest level of recognition, and where ▲ denotes the emotion with the highest level of recognition.

Supplementary Table S.3: A summary of average recognition rates (%) for different types of stimuli reported in the previous emotion recognition tasks (from the publications summarised in Supplementary Table S.1). Recognition rates expressed as percentage were not always available.

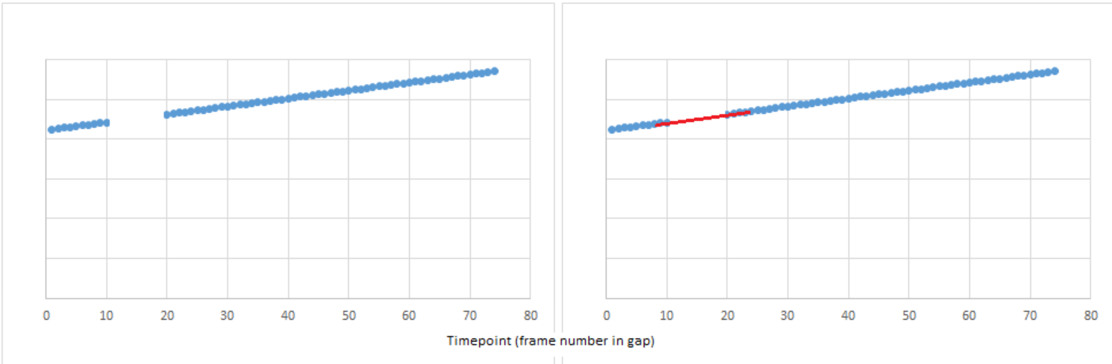
Stimuli Type	Emotion Category	Average Recognition Rate (%)	Number of Data Points Included in Calculation	Range in Reported Recognition Rates (%)
Avatar	**Overall**	**54.78**	**11**	**16 – 92**
Avatar	Neutral	71.5	1	N/A
Avatar	Happy	63.55	2	62 – 65.1
Avatar	Sad	73	2	54 – 92
Avatar	Anger	57.5	2	39 – 76
Avatar	Fear	51.5	2	23 – 80
Avatar	Other	20	2	16 – 24
FLD	**Overall**	**50.67**	**64**	**3 – 95.74**
FLD	Neutral	65.58	6	25 – 95.74
FLD	Happy	54.7	11	3 – 90
FLD	Sad	55.26	12	6 – 86.94
FLD	Anger	54.19	12	8 – 90.99
FLD	Fear	59.51	6	36 – 91.11
FLD	Other	38.74	18	3 – 75.28
PLD	**Overall**	**69**	**26**	**37.9 – 92.4**
PLD	Neutral	62.35	4	37.9 – 83
PLD	Happy	64.56	6	41.1 – 85.6

PLD	Sad	73.49	6	45.8 – 92.4
PLD	Anger	71.97	6	58.6 – 92.4
PLD	Fear	79.72	1	N/A
PLD	Other	68.22	3	63.06 – 74
Overall	**Neutral**	**64.95**	**11**	**25 – 95.74**
Overall	**Happy**	**58.74**	**19**	**3 – 90**
Overall	**Sad**	**62.5**	**20**	**6 – 92.4**
Overall	**Anger**	**60.15**	**19**	**8 – 92.4**
Overall	**Fear**	**59.98**	**9**	**23 – 91.11**
Overall	**Other**	**40.95**	**23**	**3 – 75.28**

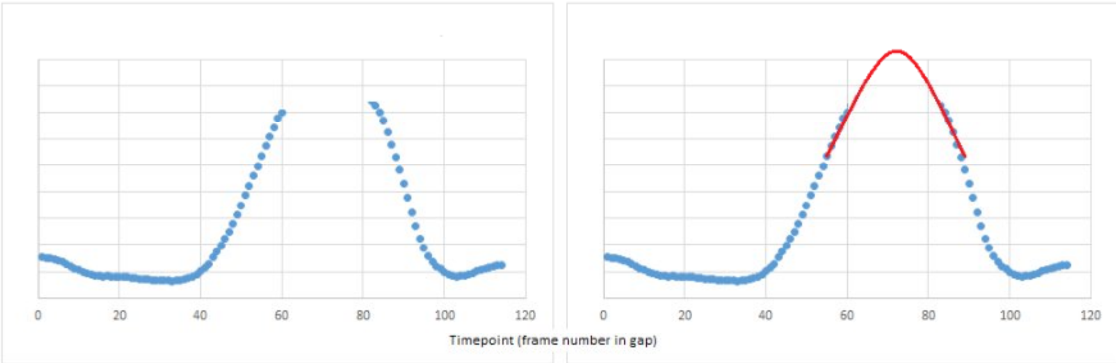
SECTION A:

MOTION CAPTURE GAP-FILLING TECHNIQUE

For each instance of missing data, the 100 preceding and succeeding points were graphed to show the intended trajectory of the coordinates across each time point of the gap. If the gap appeared to show a linear trajectory, the data was interpolated linearly (Supplementary Figure S.1). If the gap in the trajectory appeared to be parabolic then a different method was used. For seemingly parabolic gaps, the data was extrapolated linearly from the datapoints preceding the gap, and from the datapoints following the gap, the point where these lines intersected was used as the peak of the parabola and the missing data was extrapolated accordingly (Supplementary Figure S.2).



Supplementary Figure S.1: Case-by-Case Linear Extrapolation: Shows an example of missing data which appears to follow a linear trajectory. The red line depicts the points which were extrapolated to fill this type of gap.



Supplementary Figure S.2: Depicts an example of missing data which appears to follow a parabolic trajectory. On each side of the gap a linear extrapolation was carried out. The point where these two lines intersected was identified as the peak of the parabola and the surrounding points were manually smoothed to create a curve in the trajectory.

This was identified as the best method for estimating the missing data points when examining the visual output of these manipulations (compared to linear extrapolation and other such methods), but it should be acknowledged that this case-by-case method may have led to several errors that should be considered when running a computational kinematic analysis on this data. Movements contained in these sequences were often mathematically unpredictable, and the trajectories created using this method may not have accurately captured all of the complexities of this motion. However, for the creation of these point light displays, the aim was only to create stimuli with no missing movement information (i.e., to avoid markers disappearing throughout the duration of the sequence) and to minimise the amount of movement perturbation. Therefore, for this purpose, this method of gap filling was deemed sufficient.

Supplementary Table S.4: Information about each individual clip within the full, original McNorm Library.

Movement Sequence	Emotion Portrayal	Stimuli Code	Validated in Experiment 1?	Validated in Experiment 2?	Duration *(seconds)*	Video Size *(pixel width x pixel height)*	Average Recognition Accuracy (%) in Experiment 1	Average Recognition Accuracy (%) in Experiment 2	Average Recognition Accuracy (%) Across Experiments 1 & 2
rdej	Neutral	N1	Y	N	31.6	496 x 370	42	—	—
rdej	Happy	H1	Y	N	30.6	496 x 370	38	—	—
rdej	Sad	S2	Y	Y	32.6	496 x 370	52	45.28	48.64
rdej	Angry	A1	Y	N	30.2	496 x 370	12	—	—
rdej	Fearful	F1	Y	N	30.4	496 x 370	14	—	—
tendu	Neutral	N2	Y	Y	22.3	496 x 370	62	60.38	61.19
tendu	Happy	H2	Y	N	23.0	496 x 370	32	—	—
tendu	Sad	S3	Y	N	23.8	496 x 370	42	—	—
tendu	Angry	A2	Y	N	21.5	496 x 370	2	—	—
tendu	Fearful	F2	Y	N	22.6	496 x 370	20	—	—
ever_con	Neutral	N3	Y	N	12.6	496 x 370	20	—	—
ever_con	Happy	H3	Y	N	13.6	496 x 370	22	—	—
ever_con	Sad	S4	Y	N	16.4	496 x 370	26	—	—
ever_con	Angry	—	N	N	—	—	—	—	—
ever_con	Fearful	F3	Y	Y	17.2	496 x 370	42	45.28	43.64
bitter	Neutral	N4	Y	N	17.1	496 x 370	6	—	—

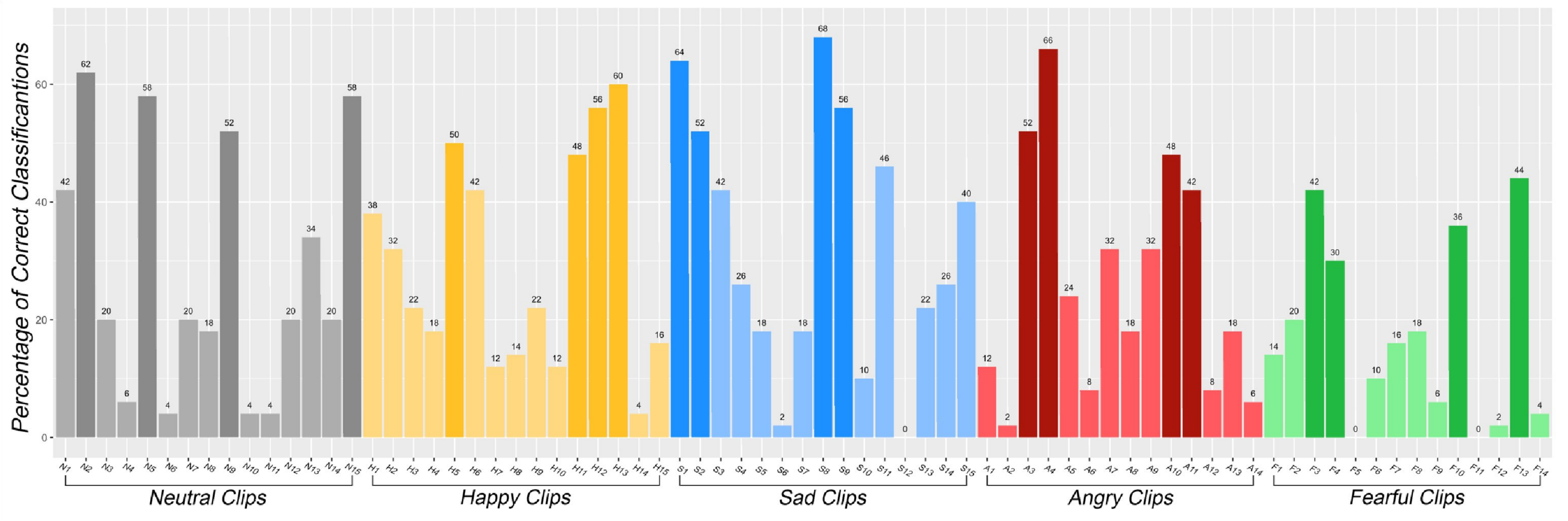
bitter	Happy	H4	Y	N	17.1	496 x 370	18	—	—
bitter	Sad	S5	Y	N	17.7	496 x 370	18	—	—
bitter	Angry	A3	Y	Y	17.4	496 x 370	52	35.85	43.93
bitter	Fearful	F4	Y	Y	18.2	496 x 370	30	30.19	30.10
char	Neutral	N6	Y	N	6.6	496 x 370	4	—	—
char	Happy	H5	Y	Y	8.1	496 x 370	50	56.60	53.3
char	Sad	S6	Y	N	9.2	496 x 370	2	—	—
char	Angry	A4	Y	Y	7.3	496 x 370	66	64.23	65.12
char	Fearful	F5	Y	N	7.8	496 x 370	0	—	—
jete	Neutral	N5	Y	Y	19.9	496 x 370	58	58.49	58.25
jete	Happy	H6	Y	N	20.9	496 x 370	42	—	
jete	Sad	S7	Y	N	20.2	496 x 370	18	—	—
jete	Angry	A5	Y	N	19.4	496 x 370	24	—	—
jete	Fearful	F6	Y	N	18.8	496 x 370	10	—	—
fondu	Neutral	N7	Y	N	37.2	496 x 370	20	—	—
fondu	Happy	H7	Y	N	42.0	496 x 370	12	—	—
fondu	Sad	S8	Y	Y	37.3	496 x 370	68	67.92	67.96
fondu	Angry	A6	Y	N	37.2	496 x 370	8	—	—
fondu	Fearful	F7	Y	N	42.8	496 x 370	16	—	—
free_mov	Neutral	N8	Y	N	26.1	496 x 370	18	—	—
free_mov	Happy	H8	Y	N	27.9	496 x 370	14	—	—
free_mov	Sad	S9	Y	Y	36.8	496 x 370	56	54.72	55.36

free_mov	Angry	A7	Y	N	25.1	496 x 370	32	—	—
free_mov	Fearful	F8	Y	N	28.1	496 x 370	18	—	—
frappe	Neutral	N9	Y	Y	16.3	496 x 370	52	52.83	52.42
frappe	Happy	H9	Y	N	15.4	496 x 370	22	—	—
frappe	Sad	S10	Y	N	16.4	496 x 370	10	—	—
frappe	Angry	A8	Y	N	15.0	496 x 370	18	—	—
frappe	Fearful	F9	Y	N	15.8	496 x 370	6	—	—
ice	Neutral	N10	Y	N	25.3	496 x 370	4	—	—
ice	Happy	H10	Y	N	22.0	496 x 370	12	—	—
ice	Sad	S11	Y	N	27.2	496 x 370	46	—	—
ice	Angry	A9	Y	N	19.8	496 x 370	32	—	—
ice	Fearful	F10	Y	Y	26.4	496 x 370	36	50.94	43.47
balanch	Neutral	N11	Y	N	8.2	496 x 370	4	—	—
balanch	Happy	H11	Y	Y	8.2	496 x 370	48	47.17	47.59
balanch	Sad	S12	Y	N	10.9	496 x 370	0	—	—
balanch	Angry	A10	Y	Y	10.1	496 x 370	48	37.74	42.87
balanch	Fearful	F11	Y	N	9.4	496 x 370	0	—	—
onegin	Neutral	N12	Y	N	24.6	496 x 370	20	—	—
onegin	Happy	H12	Y	Y	25.7	496 x 370	56	43.40	49.7
onegin	Sad	S13	Y	N	31.1	496 x 370	22	—	—
onegin	Angry	A11	Y	Y	22.6	496 x 370	42	28.30	35.15
onegin	Fearful	F12	Y	N	24.4	496 x 370	2	—	—

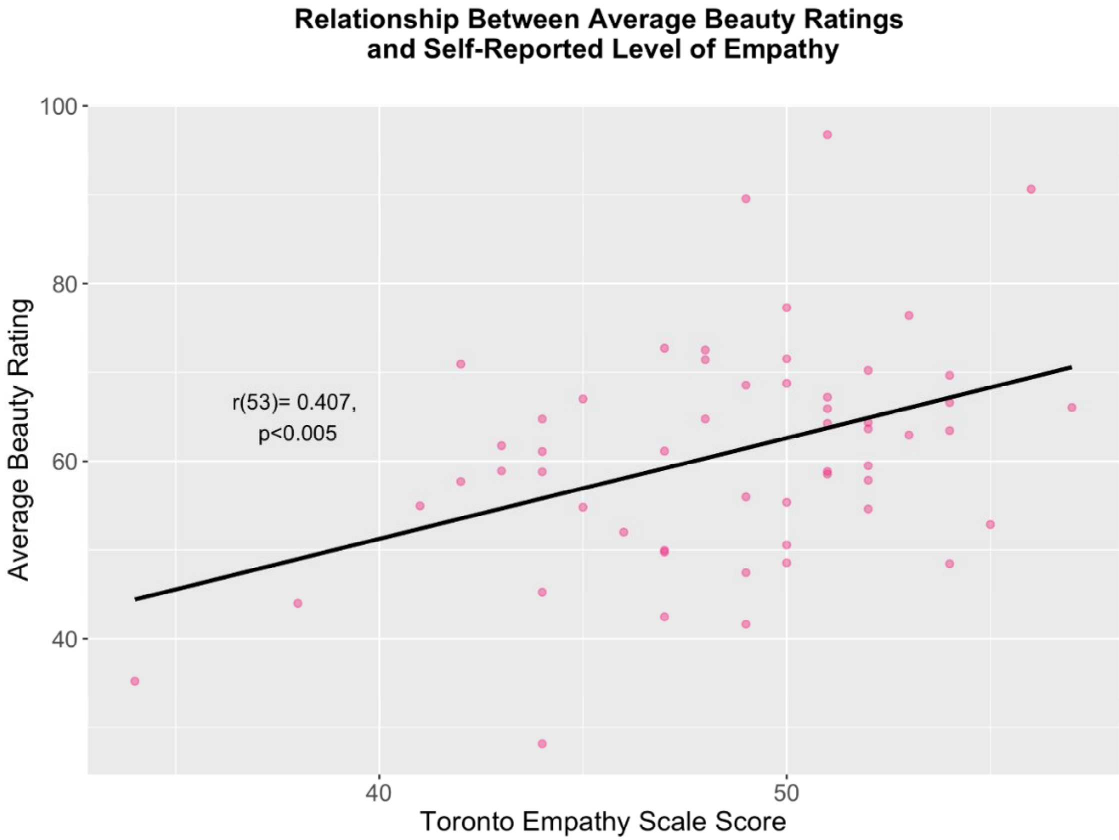
pirouette	Neutral	N13	Y	N	15.4	496 x 370	34	—	—
pirouette	Happy	H13	Y	Y	20.2	496 x 370	60	41.51	50.76
pirouette	Sad	S14	Y	N	19.9	496 x 370	26	—	—
pirouette	Angry	A12	Y	N	18.4	496 x 370	8	—	—
pirouette	Fearful	-	N	N	—	—	—	—	—
partner	Neutral	N14	Y	N	23.4	496 x 370	20	—	—
partner	Happy	H14	Y	N	23.0	496 x 370	4	—	—
partner	Sad	S15	Y	N	28.9	496 x 370	40	—	—
partner	Angry	A13	Y	N	20.4	496 x 370	18	—	—
partner	Fearful	F13	Y	Y	24.1	496 x 370	44	30.19	37.095
plie	Neutral	N15	Y	Y	29.6	496 x 370	58	62.26	60.13
plie	Happy	H15	Y	N	30.6	496 x 370	16	—	—
plie	Sad	S1	Y	Y	35.7	496 x 370	64	64.15	64.08
plie	Angry	A14	Y	N	24.4	496 x 370	6	—	—
plie	Fearful	F14	Y	N	28.4	496 x 370	4	—	—

Supplementary Table S.5: A summary of the average durations and recognition rates for clips from each emotion category in the McNorm Library.

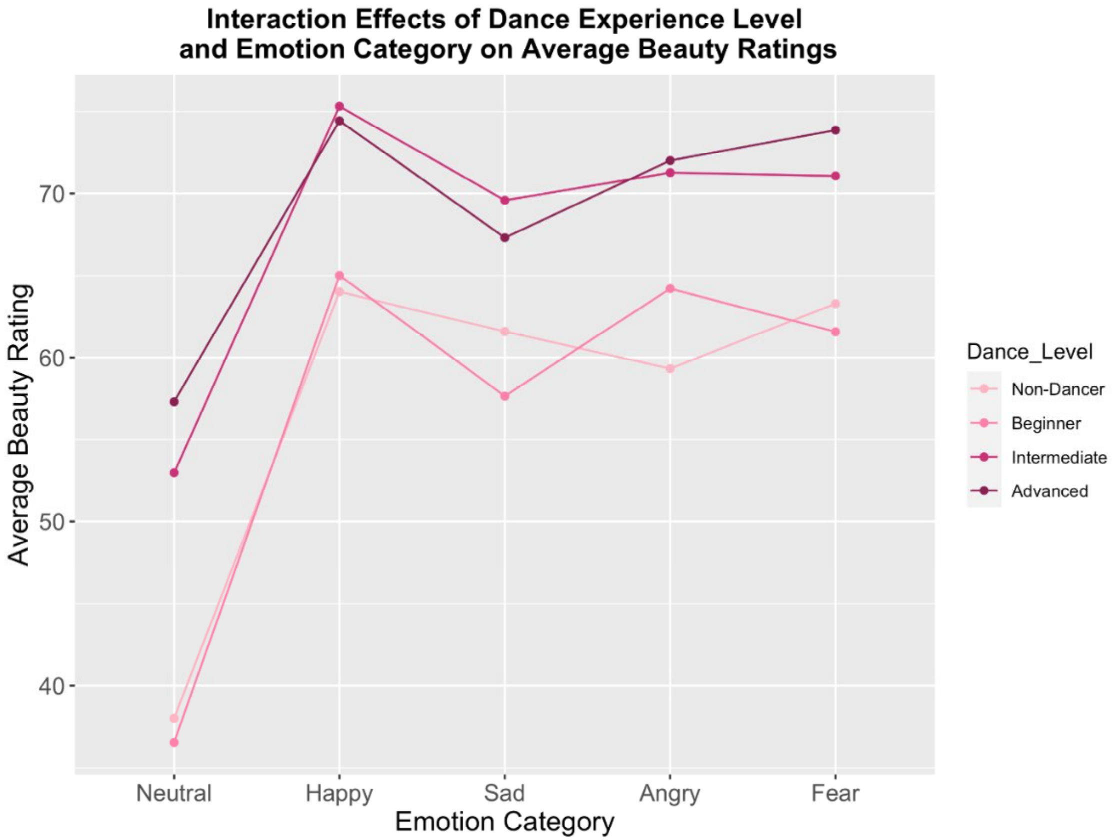
Clip Emotion Category	Mean Duration *(seconds)*	Minimum Duration *(seconds)*	Maximum Duration *(seconds)*	Mean Recognition Accuracy (%) Across All Clips in the Full Stimuli Set from Experiment 1	Mean Recognition Accuracy (%) for Subset of Clips Across Experiment 1 & 2
Neutral	21.08	6.6	37.2	28.13 *(N = 15)*	58.00 *(N = 4)*
Happy	21.89	8.1	42	29.73 *(N = 15)*	50.34 *(N = 4)*
Sad	24.27	9.2	37.3	32.67 *(N = 15)*	59.01 *(N = 4)*
Angry	20.63	7.3	37.2	26.00 *(N = 14)*	46.77 *(N = 4)*
Fearful	22.46	7.8	42.8	17.29 *(N = 14)*	38.58 *(N = 4)*



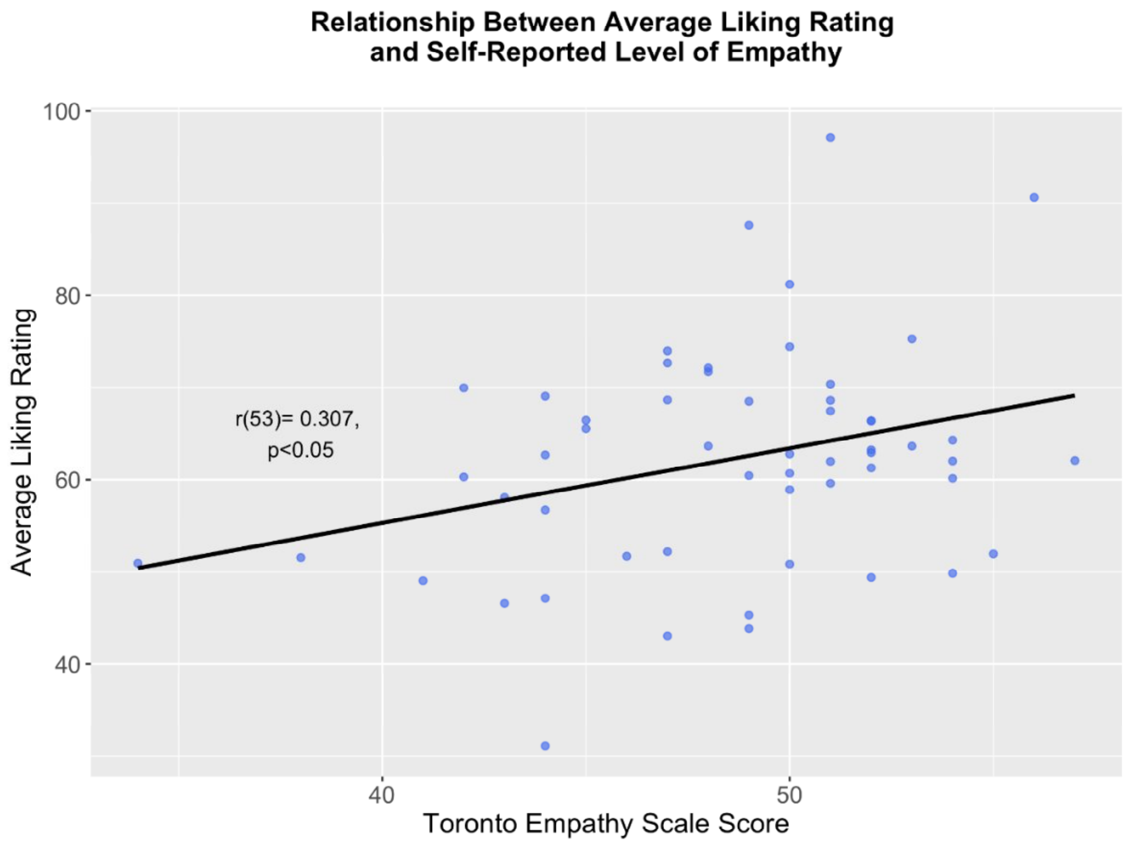
Supplementary Figure S.3: The bar plot shows average recognition of the intended emotion for each individual clip in the full McNorm Library from Experiment 1. This ranged from 0% (S12, F5, F11) to 68% (S8). Neutral clips are shown in grey, happy clips are shown in yellow, sad clips are shown in blue, anger in red, and fear in green. Highlighted bars indicate clips from the full McNorm library which were selected for further study in Experiment 2. 4 clips from each emotion category (with the highest recognition rates for the target emotion) were selected. These individual stimuli clips are N2, N5, N9, and N15 for neutral (grey); H5, H11, H12 and H13 for happy (yellow); S1, S2, S8 and S9 for sad (blue); A3, A4, A10 and A11 for anger (red); F3, F4, F10 and F13 for fear (green).



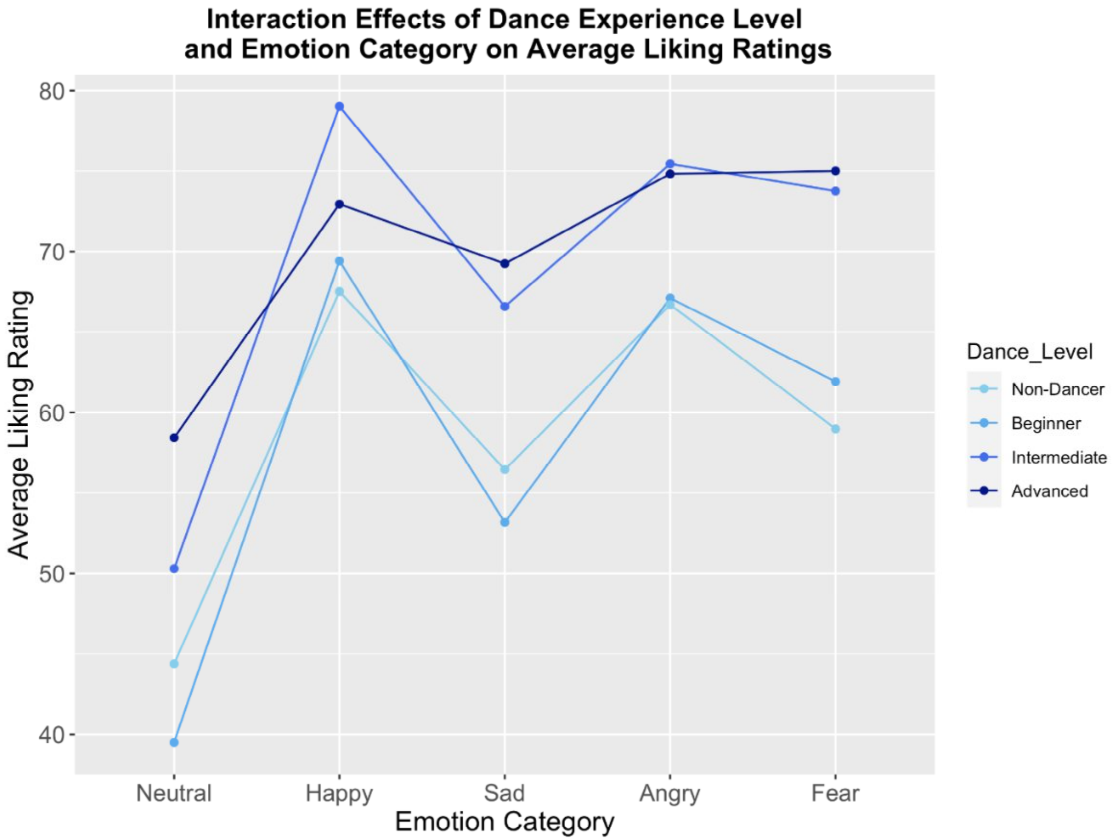
Supplementary Figure S.4: A Visualisation of the relationship between trait empathy (as measured by the Toronto Empathy Scale), and average beauty ratings provided by observers.



Supplementary Figure S.5: A visualisation of the interaction between self-reported level of previous dance experience and emotional expressivity of the clip on average beauty judgements provided by participants in this experiment.



Supplementary Figure S.6: Visualisation of the relationship between trait empathy (as measured by the Toronto Empathy Scale), and average liking ratings provided by observers.



Supplementary Figure S.7: A visualisation of the interaction between self-reported level of previous dance experience and emotional expressivity of the clip on average liking judgements provided by participants in this experiment.

SECTION B:

EXTRACTION OF COMPUTATIONAL IMAGE PROPERTIES

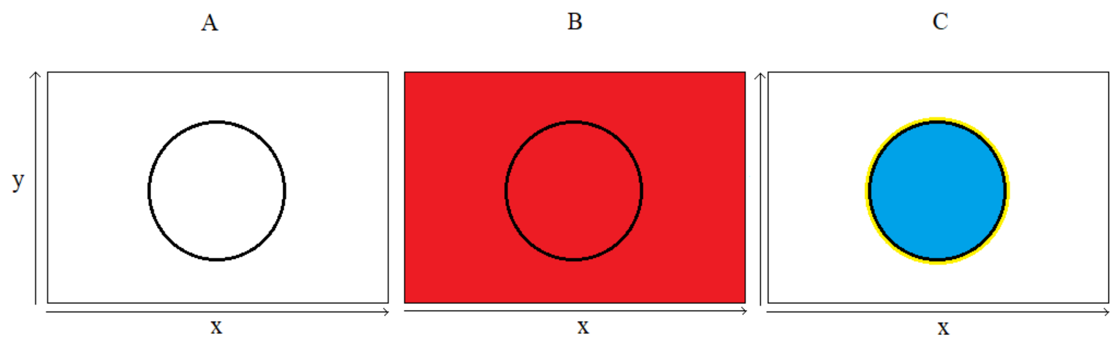
Due to the nature of the image creation process (see Experiment 1 Methods section in Chapter 4), we have access to two kinds of data; the final continuous line drawing images themselves (denoted by *i*), and the original movement coordinate position data (denoted by *pd*) they were created from.

An important caveat for the use of the pd data is that when the dancer's speed varied during the movement performances, the number of points in the resultant *pd* data varies too (i.e., faster movements will have less sequential points in the *pd* data than slower movements). To compensate for this imbalance we used cubic spline interpolation to resample the sequential data at intervals that are spaced equally (at intervals of 0.03 units in the *pd*).

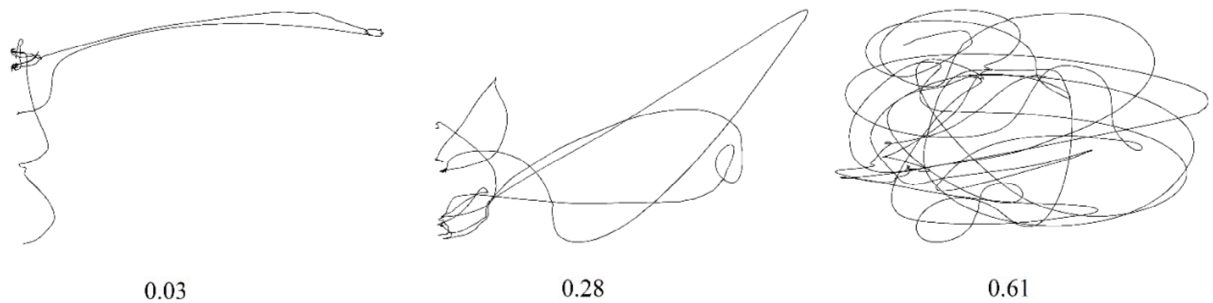
For some of the image features, we required the *pd* as the sequential ordering could not be inferred from the *i* alone. Below we describe how each of these computational features were extracted in more detail. These features will then be explored in relation to the emotion classifications and aesthetic value judgements (beauty and liking) provided by participants in the previous experiment.

1. Size Related Features

Drawing size (*i*, 0-1 ratio) is calculated by extracting the perimeter value (see Section 3.2. below), counting how many pixels fall inside this perimeter, and dividing this count by the total number of pixels in the image. Therefore, drawing size is a ratio with values ranging between 0 (no black pixels in the image) and 1 (where the perimeter of the line drawing covers the full image). We note here, that “empty” circles (i.e., only the outline of the circle), for example, will have the same drawing size as “filled” circles; as we only consider perimeter in this calculation, rather than ink size (see Section 3.1. below). A simplified example of which values are extracted during this process can be found in Supplementary Figure S.8, and the drawing size values assigned to a sample of our images can be found in Supplementary Figure S.9.



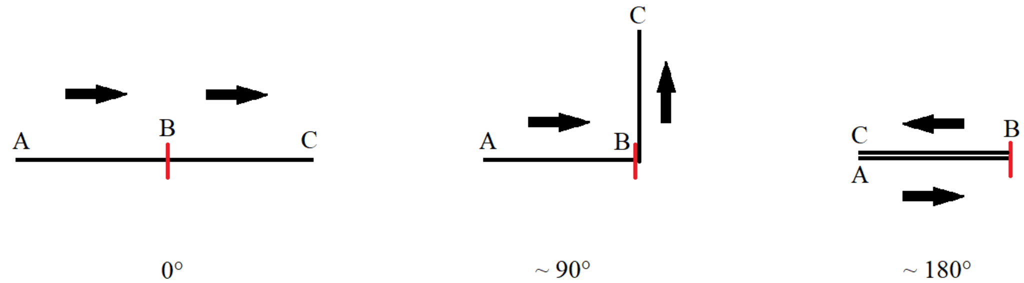
Supplementary Figure 8: Here, A represents the target drawing (note, rectangles around the circles indicate the boundaries of the image; i.e., width x, height y). As before, the perimeter of the drawing is represented by the yellow area in C. The “drawing size” feature requires extracting the total number of pixels that fall within the perimeter (i.e., all pixels within the yellow perimeter area in image C; this includes pixels in the blue area, and pixels from the original black circle). This total is then divided by the total number of pixels in the image.



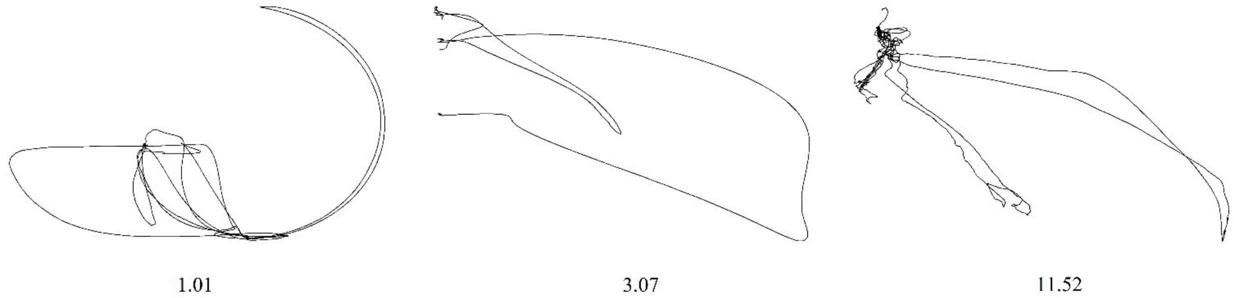
Supplementary Figure S.9: Images from our stimuli set with small, median, and large values for the factor Drawing Size.

2. Curvature Related Features

Curvature (*pd,* in °), in this instance means something different computationally than it would in everyday conversation, as higher values for this factor actually represent a greater degree of "jaggedness", rather than roundedness of contours. The possible values for this feature range between 0° (which would represent a continuous straight line; denoted by 0° in Figure 10) and 180° (the sharpest possible U-turn; denoted by 180° in Figure S.10). This value for each image was calculated by examining the angle of the turn between each of the interpolated coordinate points, and averaging these angles over the entire sequence. Higher values on this feature will correspond to a greater degree of jaggedness in general. See Supplementary Figure S.10 for a visualisation, and Supplementary Figure S.11 for the curvature values assigned to a sample of our image stimuli.



Supplementary Figure S.10: A visualisation of lines (from left to right) with the minimum (0°), median (90°), and near maximum (~180°) values for the curvature factor.



Supplementary Figure S.11: A visualisation of trajectory images from our stimuli set with low, median and large values on the curvature factor. Lower values (e.g., 1.01 for the leftmost image) indicate lower curvature, translating to lower "jaggedness", and higher values (e.g., 11.52 for the rightmost image) indicate higher curvature, which translates to higher "jaggedness".

3. Complexity Related Dimensions

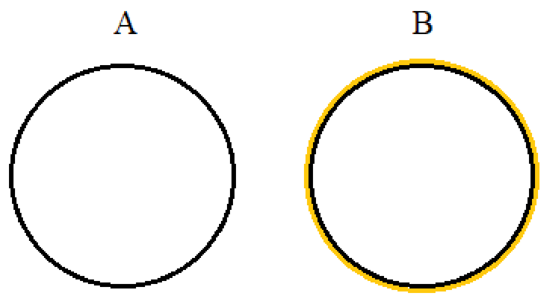
We have implemented, and adapted where necessary, the features below from existing literature on image complexity analysis (Chikhman et al., 2012; Näsänen et al., 1993).

3.1. Ink Size

Ink size (*i*, in pixels) refers to the total number of black pixels which make up the line drawing. A visualisation of what is meant by ink size can be observed in shape A of Supplementary Figure S.12.

3.2. Perimeter Size

Perimeter size (*i*, in pixels) was calculated by summing the total number of white pixels which bordered the outermost black pixels of the line drawing. A visualisation of this can also be observed in Supplementary Figure S.12 (shape B).



Supplementary Figure S.12: A illustration of how the perimeter and ink size factors would be calculated for a simple circle. Ink size is represented by the number of black pixels which make up circle A (which is the "target image" in this example). The count of yellow pixels outlining circle B represents the perimeter value.

3.3. JPEG File Size

JPEG bytes (*i*, in bytes) refers to the file size of the image, compressed to JPEG format. The compression software requires a *quality factor* where 0 is the worst quality and highest compression, and 95 is the best quality and least compression. We used a *quality factor* of 75 for all images (the default value which did not substantially decrease the original image quality), and this feature is the size of the resulting compressed image (in bytes).

Previously, researchers have suggested that JPEG bytes of an image reflect the computational complexity (Chikhman et al., 2012). Applying this rationale, higher file sizes (JPEG bytes) for images with the same dimensions corresponds to greater complexity.

3.4. Perimeter to Ink Ratio

Perimeter to ink ratio (*i*, 0-1 ratio) is calculated by dividing the perimeter size factor by the ink size factor (see Supplementary Figure S.12). This ratio is unitless, yields a positive number, and serves as a proxy for pixel density in an image. Here, larger values would indicate images which are less pixel dense (i.e., less complex), while smaller values indicate drawings which are more pixel dense (i.e., more complex).

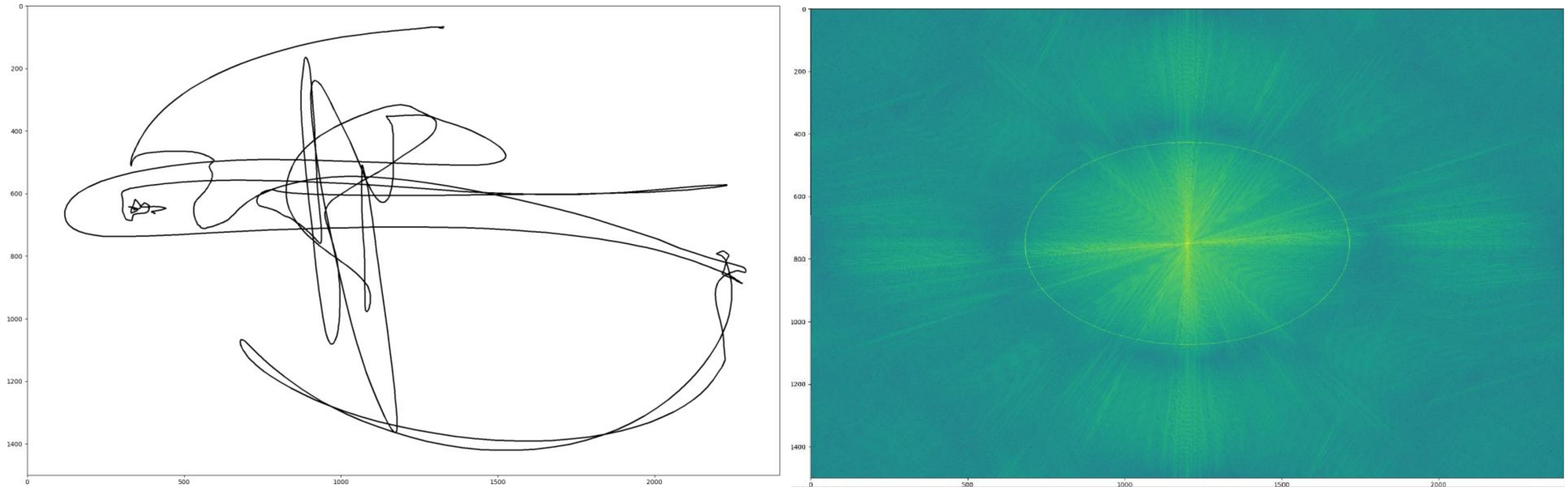
3.5. Spectral Complexity

The spectral complexity features included in this analysis are based on the work of Chikman and colleagues (2012), and Näsänen and colleagues (1993).

3.5.1. 2-Dimensional Spectral Complexity

2D spectral complexity: It has been suggested that the complexity of images can be estimating using the median of the spatial frequency distribution (higher median means more complexity; Chikhman et al., 2012; Näsänen et al., 1993). Specifically, they propose to use the square of the spectral median, multiplied by image size.

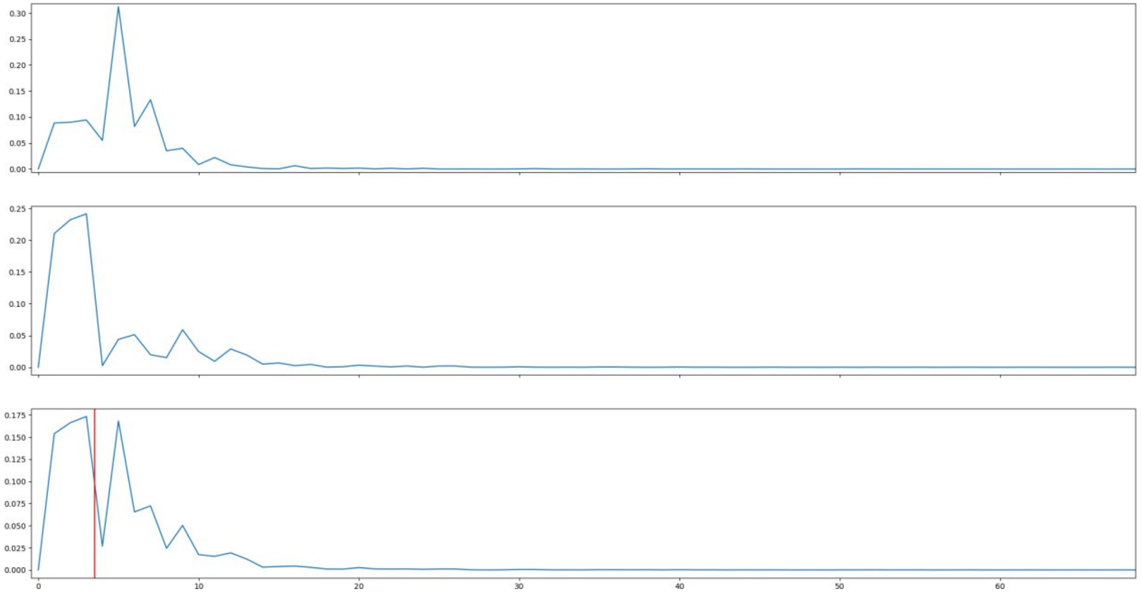
Given the 2-dimensional Discrete Fourier Transformation (DFT) of a drawing, the median is defined by Näsänen and colleagues as the spatial frequency such that 50% of the spectral energy is above and 50% below. This definition is also taken by Chikhman and colleagues. Unfortunately, it fails to address the fact that a 2-dimensional DFT is defined on a set of spatial frequency pairs, so there are infinitely many ways of doing a 50/50 split. To overcome this, we propose to use a decision boundary, such that all frequency pairs with the same (fvert2+fhoriz2) are connected (see image below). This yields a centred elliptic boundary, and naturally extends the idea of the spectral median to 2D, because a smaller ellipse will always contain strictly lower frequencies. The task then, given a 2D DFT, is simply to find the right size of a centred ellipse such that half of the DFT energy is inside, and half outside. The normalized size of the ellipse is in the 0 to √2 range, and we take this value as the 2D spectral complexity, where larger means more complexity.



Supplementary Figure S.13: This figure represents the 2D spectrogram (right) of one of our line drawing stimuli (left), where the centre contains the zero-frequency (static) component, and frequencies increase as we get close to the borders. The outlined ellipse separates 2 regions with equal spectral energy. Larger ellipses are interpreted here as a proxy for higher drawing complexity.

3.5.2. 1-Dimensional Spectral Complexity

1D spectral complexity: In addition to the 2D spectral complexity described above, we propose this 1-dimensional variant. We do so because unlike previous research on images, the data for our experiments consists of 1-dimensional lines embedded in a 2-dimensional space. For that reason, we can also compute the 1D spectrum of the sequences, and the corresponding median can also be used as a proxy for complexity, by squaring the median and multiplying it by the image size (analogously to the 2D counterpart). In this case, the values are normalized to be between 0 (the lowest frequency) and 1 (the highest frequency). As with the 2D spectral complexity feature, higher values represent images with greater complexity.



Supplementary Figure S.14: The top (and middle) plots show the 1D spectrograms of the horizontal (and vertical) positions of the drawing. The bottom plot represents the "joint spectrogram" as a result of adding both spectrograms. The vertical line on the bottom represents the median, such that half of the joint spectrogram energy is to its left, and half to its right. A line shifted to the right corresponds to a higher median, and represents more complexity.