A single case experimental design study of a reminder app for supporting adherence to personalised treatment goals in Parkinson's Disease

Georgina Rayment, (BSc, MSc)

Submitted in partial fulfilment of the requirements for the degree of

Doctorate in Clinical Psychology

Institute of Health and Wellbeing

College of Medical, Veterinary and Life Sciences

University of Glasgow

July, 2022

Table of Contents

List of Tables

List of Figures

Acknowledgements

Firstly, I would like to thank my supervisor Professor Jonathan Evans for his expertise, guidance and encouragement throughout the research process. It has been a pleasure to learn from your invaluable knowledge and experience.

Secondly, I would like to thank my field supervisor Dr Salma Iqbal for support throughout the project and her persistence with recruitment. My thanks also to Dr Victoria Marshall from Greater Glasgow and Clyde's Neurology team for your assistance with recruitment. Thank you to Sophie Jolly for her contribution to my systematic review and much needed encouragement.

Most importantly, I would like to thank all the participants and their partners who took the time to take part in this research. I am extremely grateful for your commitment to this project.

To my fellow doctorate friends, thank you for the endless entertainment and support, even in the height of a pandemic. As always, thank you to my family for your unconditional love support over the years.

Sonny, thank you for your unwavering love and patience for which there aren't enough words. Without you this wouldn't have been possible.

Finally, I'd like the thank the individuals along the way who helped me to truly understand that neurodivergence is not a flaw. When life gives you melons and all that jazz.

Chapter 1

A systematic review of the association between performance on objective neuropsychological assessment measures of executive functioning and subjective measures of executive functioning in everyday life in Parkinson's disease

Prepared in accordance with the author requirements for *The Clinical Neuropsychologist*

https://www.tandfonline.com/action/authorSubmission?show=instructions&journalCode=ntcn20#word-limits

Abstract

Background: Executive dysfunction is a common clinical feature in people with Parkinson's disease (PwPD). The relationship between neuropsychological tests of executive function (EF) and everyday executive functioning is unknown. **Objective:** This review aimed to investigate the relationship between objective deficits in EF, as measured by performance on neuropsychological tests, and everyday executive functioning, as measured by subjective (self and other) reports. **Methods:** Five databases were systematically searched. Observational studies which utilised both neuropsychological tests and self/informant report measures of EF were included if they reported statistical analysis of the association between measures. Eleven studies were identified and appraised for risk of bias using the structured Appraisal tool for Cross-Sectional Studies (AXIS tool). **Results:** A narrative synthesis of findings indicated both an objective and subjective deficit in EF was reported in the majority of included studies. No clear association was found between neuropsychological tests of EF and self-reported everyday executive functioning. **Conclusions:** There is no obvious relationship between commonly used neuropsychological tests of EF and the impact of EF on daily life measured through self-report for PwPD. However, with heterogeneity in both the samples and measures used, it is difficult to draw any strong conclusions. Future research should consider subjective reports of EF across disease severity and its association with neuropsychological tests, clarifying the nature of any relationship and informing clinical practice. Findings highlight the need for clinicians to use multiple types of measure, considering real-world application, for a comprehensive assessment of EF difficulties.

Keywords: Parkinson's disease, executive function, neuropsychological tests, everyday functioning, self-report

Introduction

Affecting around one in 500 people, Parkinson's disease (PD) is a progressive neurodegenerative disease predominantly associated with motor symptoms (NHS, 2019). Some people with Parkinson's disease (PwPD) also develop a specific profile of cognitive deficits in the early stages, characterised by difficulties with executive function (EF), episodic memory and visual-spatial abilities, and are at increased risk of developing dementia (Aarsland et al., 2017). EF has been found to be a common clinical feature in PwPD and PD-related dementia (Emre, 2003). Definitions of EF vary, but generally refer to cognitive processes that are necessary for goal-directed behaviour. These processes include inhibition, initiation, cognitive flexibility, planning, problem-solving and reasoning (Sira & Mateer, 2014). The domains of working memory and attention are often viewed as interlinked with EF (Kudlicka, Clare & Hindle, 2011). Kudlicka et al. (2011) conducted a systematic review of EF in PwPD, providing consistent evidence of executive deficits in the early stages of the disease via meta-analyses from five commonly used neuropsychological tests (Verbal Fluency, Trail Making Test B, Wisconsin Card Sorting Test, Stroop Word-Colour Interference, Digit Span).

Neuropsychological tests allow for detailed analysis of an individual's cognitive strengths and weaknesses in specific cognitive domains within a controlled environment, improving the reliability of findings (Lanni et al., 2014). A core aim of neuropsychological assessment, beyond diagnosis, is to aid the management of patient care by considering the impact of cognitive deficits on everyday function (Lezak, 2012). However, in their review, Kudlicka et al. (2011) note that despite evidence of an objective EF impairment, deficits found via neuropsychological testing may not significantly impact everyday life.

Burgess et al. (2006) raised questions over the generalisability and representativeness of traditional neuropsychological tests for predicting difficulties in daily life. It has been argued that the highly structured nature of some neuropsychological tests of EF prevents them from capturing the full range of executive functioning and hence may not reflect the demanding nature of everyday activities (Sobreira et al., 2008). To this end, Koerts et al. (2012) investigated whether reports of EF deficits in the daily lives of PwPD corresponded with scores on neuropsychological tests, finding that reported deficits in EF by PD patients and

informants did not correspond with neuropsychological tests. More than a decade later, following a multitude of further research, the relationship between neuropsychological tests and PwPD's subjective experiences remains unclear.

Self and informant report, whether through standardised questionnaires or clinical interview, is often used to understand how EF deficits (and other cognitive deficits) may impact on daily life (Puente, Cohen, Aita & Brandt, 2016). However, self-report measures are subjective in nature and therefore open to the influence of external factors (Domensino, Evans & van Heugten, 2022). For example, anosognosia (lack of awareness of deficit) is commonly associated with impairment in EF (and with frontal lobe damage), highlighting the problematic nature of reliance on self-report data for measuring everyday EF (Kudlicka, Clare & Hindle, 2013). Research into awareness of cognitive difficulties in PwPD is mixed. Koerts et al. (2012) found good agreement between self and informants reports of EF. Whereas McKinlay et al. (2008) found that PwPD reported more difficulties than informants. However, these studies did not consider the association between subjective reports and objective deficits based on neuropsychological tests.

Rationale

Deficits in EF have been evidenced even in the early stages of PD on neuropsychological tests and these deficits characterise PD-related dementia, but the relationship to daily executive functioning is less clear. A key task for neuropsychological assessment is to make predictions about how a person's deficits (identified via standardised neuropsychological tests) will impact them in everyday life (Lezak, 2012). The impact of deficits in executive functioning in everyday life is most typically measured by the use of questionnaires that ask for the view of the patient, or a significant other, on how they function (Domensino et al., 2022). There are limitations to this, particularly for self-report, given the possibility of impairments of awareness that come with EF deficits. Therefore, to test whether neuropsychological tests of EF are useful at predicting everyday functioning in PwPD, there is a need to examine whether performance on neuropsychological tests predicts scores on subjective measures of everyday executive functioning.

Objectives

The current review aimed to investigate the relationship between objective deficits in EF, as measured by performance on neuropsychological tests of executive functions, and everyday executive functioning, as measured by subjective (self and other) reports.

Review questions

What is the impact of PD on executive functioning in everyday life when measured by subjective measures?

What is the association between subjective and objective assessment of EF in PD?

What is the methodological quality of the available evidence?

Method

This review has been reported in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses 2020 updated guidelines (PRISMA; Page et al., 2021). A protocol for this review is registered on Prospero (CRD42022310586).

Eligibility criteria

All studies were screened based on the following criteria:

Inclusion

- Studies which included participants with a formal diagnosis of PD, at any age or stage, without dementia.
- Studies which investigated EF, as measured by objective neuropsychological tests and as defined by the authors. Objective neuropsychological tests were defined as those which have been standardised, validated and have normative data (Lezak, 2012).
- Included studies had to report subjective measures of executive functioning in everyday life. These could be self-report, informant-report or both.
- Studies were required to report data investigating the statistical association between objective assessments and subjective reports of EF.
- Studies were published in English in a peer reviewed journal.
- Only quantitative, observational design studies were included, such as cross-sectional, cohort and case-control studies.

Exclusion

- Studies with participants with a diagnosis of dementia were excluded unless used as a comparator group.
- Intervention studies and qualitative designs, editorials, reviews, protocols, conference papers, theses and dissertations, and books were excluded.

Information sources

The following electronic databases were searched from their inception on 14th May 2022;

Via Ebscohost:

- PsycINFO
- PsycARTICLES
- MEDLINE
- CINAHL

Via Ovid:

- Embase

Grey literature was not searched as part of this review. Searching of grey literature can be difficult to systematically replicate in the same way as peer reviewed literature and has difficulties with sensitivity and specificity of the search strategy (Adams et al., 2016).

Search Strategy

A systematic search strategy was created using the eligibility criteria for included studies and keywords from related research. The sensitivity and specificity of the search strategy was examined and amended where necessary by conducting a scoping review of results and ensuring key papers were included. The search strategy was then adapted for use with each database and reviewed by a University of Glasgow librarian. See appendix 1.1 for an example search strategy.

Selection process

Search results were exported to EndNote X9, where de-duplication was completed. Results were then uploaded to Rayyan systematic reviewing software, where titles and abstracts were screened against inclusion criteria using screening tool developed for this review (appendix 1.2). A second reviewer (SJ) screened 10% of papers (N=537) to ensure reliability of the screening tool. There was 99.6% agreement between the two raters, suggesting good inter-rater reliability. Studies included at this stage were read in full and screened against eligibility criteria. Reference lists of included studies were screened to highlight any studies that were not identified by the initial search strategy.

Data Collection Process

The remaining studies were included within the review and data extracted using an extraction template on Microsoft Excel which included all relevant variables related to the eligibility criteria and can be found in Tables 1.1, 1.2, 1.3, 1.4.

Rating of Methodological Quality

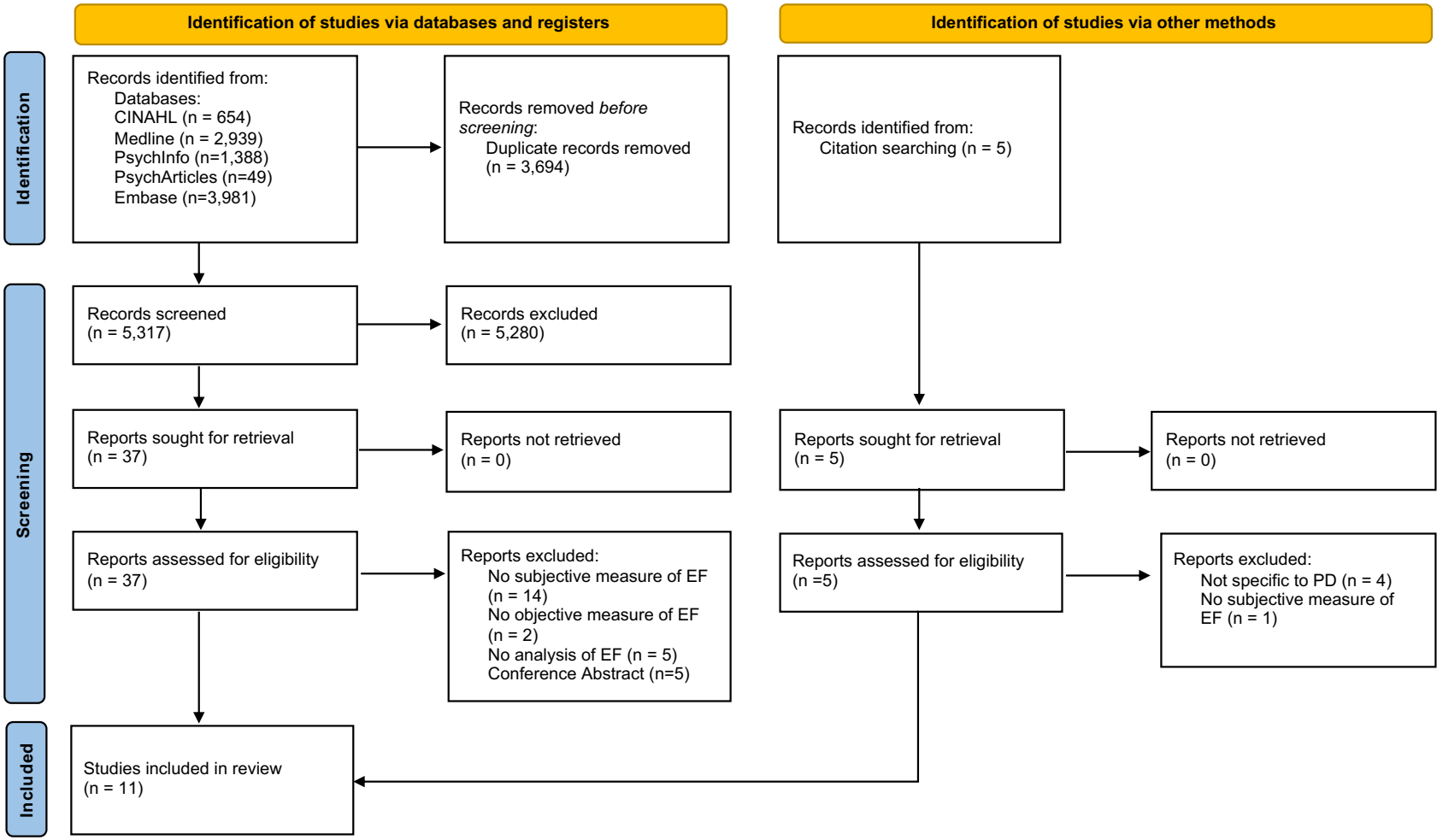
Prior to synthesis of findings, each study was appraised for methodological quality and risk of bias using the structured Appraisal tool for Cross-Sectional Studies (AXIS tool, Downes, Brennan, Williams & Dean., 2016, Appendix 1.3). The AXIS tool does not provide an overall score but instead provides questions related to three subdomains revelant for appriasing observational studies. These subdomains are quality of reporting, study design quality and introduction of bias. The second reviewer (SJ) reviewed 5/11 studies on the AXIS tool to assess the methodological quality. Across all items there was 93% agreement indicating good inter-rater reliability, with discrepancies resolved through discussion.

Results

Study selection

A total of 9,011 records were identified with the search strategy from electronic databases. Following removal of duplicates, 5,317 articles were screened by title and abstract against inclusion criteria. Thirty-five studies were selected for full text screening with a further five articles identified and screened via reference list screening. In total 11 papers were included within this review (see figure 1.1).

Figure 1.1 Flowchart of search strategy (PRISMA, 2020)



Study characteristics

A detailed description of study and participant characteristics can be found in Tables 1.1 and 1.2. Overall, studies examined 785 PwPD and 354 healthy controls (HC) were used as a comparison group. Seven of the 11 studies included a control group for comparison (Koerts et al., 2011a; Koerts et al., 2012; Kudlicka et al., 2013; Lanni et al., 2014; Lovstad et al., 2016; Vlagsma et al., 2017). Eight studies also compared participant and informant data (Copeland, Lieberman, Oravivattanakul & Tröster, 2016; Koerts et al., 2011a; Koerts et al., 2012; Kudlicka et al., 2013; Lanni et al., 2014; Lovstad et al., 2016; Siquier & Andres, 2021; Vlagsma et al., 2017).

Four studies were conducted in the USA (Copeland et al., 2016; Lanni et al., 2014; Mills et al., 2020; Puente et al.,2016), three in the Netherlands (Koerts et al., 2011a; Koerts et al., 2012; Vlagsma et al., 2017) and one each in the UK (Kudlicka et al., 2013), Spain (Siquer & Andres, 2021), Norway (Lovstad et al., 2016) and Brazil (Sobreira et al., 2008). The majority of studies recruited from movement disorder/neurology clinics. Three studies drew data from existing databases, including a medical trial (Mills et al., 2020) and pre assessment for Deep Brain Stimulation surgery (Lovstad et al., 2016; Puente et al.,2016). Only Lanni et al. (2014) recruited from a variety of sites, recruiting participants from movement disorder clinics, PD support groups, senior centres and veterans' organisations.

Risk of bias

All studies were evaluated against the AXIS tool, which covers quality of reporting, quality of study design and consideration of the introduction of bias (Downes et al., 2016, Table 1.4). Quality appraisal was coded with 'Low' corresponding to low of risk of bias and good quality and 'High' corresponding to high risk of bias and poorer quality, in all subdomains. Studies were not excluded based on quality appraisal.

Generally, most studies displayed a good quality of reporting (n=10), with all studies clearly describing aims, target population and methods and internally consistent results (n=11). Only two studies were deemed low risk in the introduction of bias subdomain due to undertaking measures to address and categorise non-responders (Kudlicka et al., 2013; Puente et al., 2016). Within this subdomain, most studies used appropriate well-validated measures (n=9). However, two studies were

deemed to have a high risk from the introduction of bias due to a non-representative sample selection process (Lovstad et al., 2016) and inappropriate measures (Copeland et al., 2016). The majority of studies had a good study design quality with a ‘Low’ risk of bias rating in seven cases. However, all included studies failed to justify their sample size, with the only mentions of this arising within the limitations section. Only one study was deemed to have a poor study design quality due to a non-representative sample (Lovstad et al., 2016).

Six studies were deemed to show low risk of bias overall (Kudlicka et al., 2016; Lanni et al., 2014; Mills et al., 2020; Puente et al., 2016; Siquier & Andres, 202; Vlagsma et al., 2017). Only one study was deemed to be high risk of bias overall as a result of a small, non-representative sample and unclear reporting (Lovstad et al., 2016). However, the level of bias in three studies was deemed unclear and so risk of bias is possible (Koerts et al., 2011a; Koerts et al., 2012; Sobreira et al., 2008). See appendix 1.3 for full quality appraisal.

Table 1.1 Study characteristics

Study, Design	**Aims**	**Tests of EF**	**Findings on Objective and Subjective tests**	**Association between objective and subjective measures**
1) Copeland et al. (2016) - Cross-sectional study - PwPD PD-MCI and informants	1) The accuracy of PD-MCI participant and informant subjective EF reports compared to objective deficits across all cognitive domains 2) Agreement between participant and informant reports	Objective: WCST, SWCT, TMT B Subjective: Opinion on changes in EF. Responses recorded yes/no. Specific examples given. – measure developed for the study	Objective deficits in EF were observed in more than half of PD-MCI participants (67% 28/42). EF was the domain least likely to be endorsed as impaired PD-MCI participant (14% 6/42) consistent with informant report (5/42). There was agreement between PD-MCI and informants K = .48, P < .01.	Poor agreement between PD-MCI subjective reports and objective deficits in EF K = -.00 P = ns. Poor agreement between informants and objective deficits in EF K = .05 P = ns.
2) Koerts et al. (2011a) - Observational, Cross-sectional study - PD patients vs HCs	1) Whether objective measures of EF reflect difficulties in daily life and vice versa	Objective: FAB Subjective: DEX	PwPD showed worse scores on the FAB (t= -5.6; p=<.001, d=1.19) and on the DEX-self (t=2.2, p=.03, d=.47). PwPD did not differ from informants on EF deficits in daily life (t=1.3; p = 0.20; d = 0.19)	The DEX-self total score was not explained by the scores on the subtests of the FAB (F = 0.96; p = 0.45; R2 = 0.13). The total score on the FAB was not explained by the subscales of the DEX-self (F = 1.23; p = 0.31; R2 = 0.06).
3) Koerts et al. (2012) - Cross-	1) Whether people with mild-moderate PD display EF deficits	Objective: SCWT, TMT B, OMO, Odd Man Out test, Zoo Maps BADS,	PwPD showed higher scores on the DEX-self compared to HC's (t= 2.1; p =0.04; d= 0.5) but no differences	No significant correlations were found between the DEX-self, DEX-other and EF test performance for PD

sectional study - PwPD vs HCs	in daily life 2) The association between subjective EF reports with informants and objective EF measurement	Semantic Verbal Fluency, Phonemic Verbal Fluency, Digit Span (WMS-R) Subjective: DEX-self & other	were found between groups for the DEX-other (t= 1.9; p= ns; d= 0.4). PwPD and HCs reported on average the same number of problems as their relatives (respectively, t=-1.6; p= ns; d= 0.2 and t= -1.6; p= ns; d= 0.5). The DEX-self scores differed from the DEX-other within PwPD in H&Y stage 2.5-3 (t= -3.6; p= .002; d= 0.6). PwPD also showed lower scores on some tests of EF than HC's (TMT z=-2.5 p=.01; OMO z=-3.5 p =<.001; Word fluency professions z=2.2 p= .03).	patients (correlations ranging from -0.13 to 0.18 and -0.17 to 0.13).
4) Kudlicka et al. (2013) - Cross-sectional study - PwPD (with & without EF deficits) vs HCs	1) The accuracy of PD patients in assessing overall EF and performance on EF tasks compared to HCs. 2) The correlates of decreased awareness	Objective: DKEFS: TMT, DKEFS: CWI Subjective: BRIEF-A (discrepancy scores), self-ratings of performances on Likert scale	No differences between groups on BRIEF-A self (H[2] = 3.33, p=.189) or BRIEF-A informant (H[2]=3.87, p=144). Self- and informant BRIEF-A ratings (compared within each study group) were similar for controls and PwPD with EF deficits. PwPD without EF deficits reported higher BRIEF-A ratings than did their informants (t=66.50, p=.001). PwPD with EF deficits were less accurate than those without deficits and HC, overestimating their	PwPD with EF deficits BRIEF-A self-ratings were negatively related to performance on CWI 3, with poorer performance on CWI 3 related to fewer difficulties reported (rs= -.439, p<.05, n=21). In PD without EF deficits both self- and informant BRIEF-A ratings were positively related to performance on CWI 4, with poorer performance associated with more

			performance on rating of performance (TMT: F[2,68]= 12.03, p<.000; CWI t[32]=4.02, p<.000)	difficulties reported Self: rs=.390, p<.05, n=40; inf: rs=.423, p<.05, n=30). TMT showed no significant correlations to subjective tests.
5) Lanni et al. (2014) - Cross-sectional study - PwPD, informants vs HC	The relationship between self and informant reports of EF in daily life and the contribution of specific factors	Objective: DKEFS: CWI, D-KEFS Verbal Fluency Test, D-KEFS Category Fluency subtest, D-KEFS Category Switching subtest, Digit Span (WAIS III), Symbol Digit Modalities Test TIADL (not specific to EF) Subjective: BRIEF-A: BRI, MI subsets	There was a difference between PwPD and HCs on the MI domain (t(83)=2.00, p=.048). Comparisons between self and informants also found higher self scores than informant reports on the MI domain (t(83)=2.34, p=.020). Higher scores indicate more impairment. No difference was found between raters on the BRI domain (t(83)=0.25, p=.565). On objective measures, PwPD showed poorer scores on the DKEFS CWI (f(1,82)=5.11, P=.026) and the SDMT (F(1, 82) = 13.58, p < .001) but no differences on Category Fluency, Category Switching, or Digit Span. PwPD showed slower completion time on the TIADL (F(1, 82) = 5.08, p = .027)	The CWI and Category Switching subtests did not predict EF in daily life. The SDMT was associated with MI in PwPD (β = –.33, t(10) = –2.44, p = .019), explaining a significant portion of the variance(R^2 = .15, F(1, 44) = 8.03, p = .007) suggesting decreased processing speed is associated with subjective EF reports. Processing speed was also associated with informant report with (β = –.60, t(9) = –3.36, p = .002) explaining a significant portion of the variance, (R^2 = .11, F(1, 44) = 4.55, p = .040). With Category Fluency and Semantic Fluency also associated with informant

				report (β = –.45, t(7) = 2.77, p = .009; R^2 change= .14, F(1, 43) = 6.65, p = .014)
6) Lovstad et al. (2016) - Cross-sectional study - PwPD with severe motor problems vs HCs	1) The effect of EF on daily living in neurological and neuropsychiatric samples. 2) The association between BRIEF-A and self-reported psychological problems	Objective: DKEFS: CWI, Letter-Number Sequencing (WAIS-III), EF index calculated. Subjective: BRIEF-A self & informant, GEC, BRI MI subdomains	The PwPD did not differ from the HCs on the BRIEF-A GEC and BRI but did on the MI (t(83)=2.00, p=.048; p-values<.001–.02) No significant differences found between self and informant on BRIEF-A.	There was a positive correlation between CWIT1 and BRI in the PD group (r = .39; p < .01; R^2 = .15). No significant associations were found between informant BRIEF-A and the EF index.
7) Mills et al. (2020). - Longitudinal cohort study - Early-stage PwPD	1) Association between MoCa domains and subjective cognitive impairment 2) whether Neuro-QoL domains are predictive of future PD-MCI	Objective: MoCA - visuospatial/executive domain score Subjective: Neuro-QoL -“Applied Cognition - Executive Function” domain	No group level deficits in EF reported by PwPD on Neuro-QoL-EF (t-scores M 53.6 SD = 6.43) Average EF scores suggest no group level deficits on testing by MoCA: visuo-spatial - executive domain: raw scores M = 4.6, SD = 0.63	There was no correlation between Neuro-Qol EF and MoCA EF scores. None of the baseline MoCA domain scores predicted the degree of Neuro-QoL change when adjusted for demographic and clinical variables.
8) Puente et al. (2016) - Cross-sectional study -	Whether informant ratings of EF and EF test scores predict performance on activities of daily living in PD	Objective: TMT B, DKEFS The Tower subtest, Brixton Spatial Anticipation Test, Letter–word fluency Subjective:	PwPD MCI compared to PwPD had worse TMT B-A (96s vs. 43s; p < .001) and greater reported EF difficulties on the FrSBe (T-score, 68 vs. 55; p = .015)	The TMT B-A was positively correlated to all FrSBe scales except Disinhibition scale. (Total: r= .268, p<.01; Apathy: r= .273, p<.01.; EF: r= .251, p<.05; disinhibition: r= .124, p= ns)

PwPD, Informants		FrSBe total (informant): current functioning - subdomains: Apathy, disinhibition, Executive dysfunction Lawton IADLs and Physical Self-Maintenance Scales-rater form (not specific to EF)	In the PD MCI sample, the EF tests accounted for a significant amount of variance in IADLs (adjusted R^2 = .371), F(4, 23) = 4.974, p = .005). The significant predictors were TMT B-A (β = .404, p < .05) and Brixton SAT (β = .364, p < .05).	Letter–word fluency was only related to the FrSBe Disinhibition scale (r=.180, p<.05). Neither the Brixton nor the D-KEFS Tower Test was related to any FrSBe scale. IADLS were positively correlated with TMT-b-A (r=.324, p<.01), Brixton (r=.252, p<.05) and the FrSBE (total: r= .512, p<.01, apathy: r=.532, p<.01; disinhibition r=.467, p<.01.; EF r=.441, p<.01).
9) Siquer & Andres (2021) - Cross-sectional study - PwPD vs HCs	The extent that inhibition and cognitive flexibility are affected in cognitively intact PD and its effect on daily life through performance and subjective measures	Objective: The Trails Test Hayling Test Subjective: BRIEF-A (discrepancy scores between self and other), subdomains. QUIP-RS.	PwPD displayed more reported difficulties in EF on the BREIF-A on total score and subdomains ([GEC, Mann–Whitney U = 181, p = 0.004], [BRI index, Mann– Whitney U = 164, p = 0.033] [MI index, Mann–Whitney U = 189, p = 0.001]) and higher scores on the QUIP-RS compared to controls ([t(27) = 2.26, p = < 0.05, d = 0.839]). No differences between self- and informant-report for either group or index.	The Hayling Test positively correlated with BREIF-A and QUIP-RS (GEC: r = 0.572, p < 0.001; BRI: r = 0.546, p < 0.01; MI: r = 0.580, p < 0.01; QUIP-RS; r = 0.360, p < 0.05).

			On objective measures there was a difference between groups on both tests, with PwPD performing worse than controls. (Hayling: [F(1, 28) = 43.32, p < 0.001, ηp2 = 0.607]. Trails Test: [F(1, 28) = 10.56, p < 0.01, ηp2 = 0.273]).	
10) Sobreira et al. (2008) - Cross-sectional study - PwPD	Performance of tests of EF in PD and explore the relationships between simple and complex bedside tests	Objective: WCST, FAB, Semantic Verbal Fluency, Digit Span - inverse order (WAIS III) Subjective: MDRS Attention,MDRS Initiation/reservation, MDRS conceptualisation,SCOPA-COG - attention and EF subscale	Low mean scores on objective and subjective measures	The FAB was positively correlated with MDRS C (rs=.814, p=.001), MDRS I/P (rs=.601, p=.001), MDRS A (rs=.814, p=.002) and the Scopa-cog-ef (rs=.541, p=.002). The WCST (preservative errors) was negatively correlated with MDRS C (rs= -.445, p=.01) MDRS I/P (rs=-.407, p=.019) but not with MDRS A or the SCOPA-COG-EF. The WCST (categories completed) was positively correlated with MDRS I/P (rs= .382, p=.028) but no other subjective measure. Verbal Fluency was positively correlated with MDRS C (rs=.501, p=.003),

				MDRS I/P (rs=.529, p=.002) and the scopa-cog-ef (rs=.371, p=.04), but not MDRS A. Digit Span was positively correlated to MDRS I/P (rs=.358, p.044) and SCOPA-COF-EF (rs=.507, p=.004) but not MDRS C or MDRS A.
11) Vlagsma et al. (2017) - Cross-sectional study - PwPD vs HCs	Whether level of participation and QoL is predicted by impairments in EF, measured by subjective and objective measures	Objective: SCWT, TMT B, Zoo Maps (BADS), Phonemic Verbal Fluency, Visual Elevator Test subtest (Test of Everyday Attention) Subjective: DEX-self/other, BAFQ, BDEFS	PwPD scored worse on TMT-B (χ^2 =7.61, P=.006), Visual Elevator Test (χ^2 =4.09, p=.043) and the BADS Zoo Map subtest (χ^2 =6.33, p=.012). PwPD reported more problems with EF in daily life on all subjective measures than controls (DEX, F=9.57, p=.002, d=1.20; BAFQ, F=7.34, p=.008, d=1.18; BDEFS-TM, F=22.52, p=<.001, d=1.54). No significant differences between self and informants on subjective ratings other than BDEF-TM, where PwPD reported more problems with time management than informants (t=2.82, p=.008, d=36).	No significant relationship was found between objective test performance and self-reported EF deficits

Key: WCST: Wisconsin Card Sorting Test; SCWT: Stroop Colour-Word Interference test; TMT B: The Trail Making Test part B-part A; FAB: Frontal Assessment Battery; DEX: Dysexecutive questionnaire; OMO: Odd Man Out test; BADS: Behavioural Assessment of the Dysexecutive Syndrome; WMS-R: Wechsler Memory Scale-Revised; DKEFS: Trail Making Test (TMT4/TMT2); DKEFS CWI:

Color-Word Interference Test; BRIEF-A: Behavior Rating Inventory of Executive Function for Adults (GEC: Global Executive Composite, subdomains BRI: Behavioural Regulation Index, MI: Metacognition index); WAIS III: Wechsler Adult Intelligence Scale – third edition; TIADL: Timed Instrumental Activities of Daily Living; MoCA: Montreal Cognitive Assessment, FrSBe: Frontal Systems Behavior Scale; IADL: Instrumental Activities of daily living QUIP-RS: The Spanish version of the Questionnaire for Impulsive- Compulsive Disorders in Parkinson's Disease-Rating Scale; MDRS: Mattis Dementia Rating Scale, SCOPA-COG: The SCales for Outcomes in PArkinson's disease-COGnition; BAFQ brock adaptive functioning questionnaire; BDEFS-TM "time management" scale of the Barkley Deficits in Executive Functioning Scale

Table 1.2 Participant characteristics

Study	Recruitment site	Participants N	Mean age (SD)	Sex (M/F)	Education (years) Mean (SD)	Cognitive screen	Disease duration (years)	Medication (LEDD)	H&Y,	UPDRS part III
Copeland et al. (2016)	Outpatient clinics, USA	PD 42 (IN 42)	67.31 (5.88)	28/14	15.19 (2.59)	DRS: 134.67 (4.90)	n = 41 8.36 (5.71)			
Koerts et al. (2011a)	MDCs, The Netherlands	PD 39	63.5 (8.5)	22/17	*5.2 (0.9)	27.5 (1.4)	4.6 (3.7)	562.7 (446.6)	2.2 (0.6)	24.2(8.4)
		HC 24	63.0 (11.7)	10/14	*4.8 (0.8)	27.5 (1.1)				
Koerts et al. (2012)	MDCs, The Netherlands	PD 43 (IN)	63.7 (8.6)	24/19	*5.2 (1.1)	MMSE 27.5 (1.4)	5.1 (4.1)	561.7 (435.3)	2.2 (0.6)	24.6 (8.8)
		HC 25 (In)	62.8 (11.5)	11/14	*4.8 (0.7)	27.6 (1.2)				
Kudlicka et al. (2013)	MDCs, NHS Wales, UK	PD 65 (EF- 23, EF+ 42) (IN)	70.11 (8.92) EF- 72.91 (7.25) EF+ 68.57 (9.44)	30/35 EF- 10/13 EF+ 20/22	12.97 (2.98) EF- 12.41 (2.78) EF+ 13.27 (3.07)	MMSE: 29.48 (0.92) EF- 29.30 (0.88) EF+ 29.57 (0.41)	***71.97 (50.42), EF- 81.93 (61.60) EF+ 66.51 (42.96)	579.19 (556.35) EF- 685.84 (690.38) EF+ 523.32	1.34 (0.57) EF- 1.53 (0.55) EF+ 1.33 (0.57)	
		43 (IN)	72.02 (6.05)	8/25	13.98 (2.15)	28.63 (1.02)				
Lanni et al. (2014)	MDCs, senior centres, PD support groups,	PD 51 (IN)	66.6 (5.6)	25/ 26	16.0 (2.8)	MMSE: 28.6 (1.7)	5.4 (4.2)	589.10 (402.20)	2.2 (0.6)	22.2 (14.8)
		HC 38 (IN)	65.8 (6.3)	19/19	16.9 (3.0)	28.6 (1.3)				

	veteran's orgs, USA									
Lovstad et al. (2016)	data from pre DBS surgery, Norway	PD 45 (IN 21)	59.8 (6.5)	31/9	13.5 (3.1)	DRS, 140.1 (3.3)	12.1 (3.8)		2 (0 to 3)	****
		HC 115 (IN 46)	31.3 (11.2)	49/66	13.2 (2.6)					
Mills et al. (2020)	57 Parkinson Study Group sites, USA	323	61.8 (9.11)	223/ 100		MoCA: 28.1 (1.4)		Not on medication	1.7 SD (0.49)	17.2 (7.0)
Puente et al. (2016)	John-Hopkins Hospital, USA	85	65.1 (7.7)	57/28	15.3 (3.0)	DRS-2: 136.2 (5.7;	**9.4 (5.4; 1–35)	levodopa % 96.4		17.6 (10.1)
Siquer & Andres (2021)	Neurology department, Spain	PD 15	67.3 (9.7)	14/1	13.4 (4.6)	MoCA: 26.5 (2.4)	**6.87 (4.61)	729.53 (298.23)	1.77 (0.37)	17.13 (10.20)
		HC 15	67.1 (5.64)	13/2	14.1 (3.08)	27.6 (1.2)				
Sobreira et al. (2008)	MDC, Brazil	35	63.1 (12.4)	21/14	5.5 (4.1)	MMSE: 24.8 (3.0)	7.0 (4.3)		2 (0.6)	11.7 (6.6)
Vlagsma et al. (2017)	3 departments of neurology, The Netherlands	PD 42 (IN 39)	60.8 (9.9)	27/15	*5.6 (1.1)			718.3 (622.3), 0-3020	2.1 (0.6), 1-3	20.3 (9.5), 8-59
		94	58.5 (6.8)	43/51	*5.3 (0.9)					

*Key: N: sample size, PD: Parkinson's disease, IN: Informant, HC: Healthy control, SD: standard deviation, MDC: Movement Disorder Clinic LEDD: Levodopa daily equivalent, H&Y:Hoehn and Yahr stage, UPDRS part III: Unified Parkinson;s Disease Rating Scale motor examination, DRS: Dementia rating scale, MMSE: Mini Mental State examination, EF-: deficits in EF, EF+: No deficits in EF, MoCA: Montreal cognitive assessment, * Education rated on scale 1 (elementary school not finished) to 7 (university degree) ** Years with symptoms *** Months **** inclusion more than 20 scoring*

Table 1.3 – Table of objective and subjective measures used in included studies

	Copeland et al. (2016)	Koerts et al. (2011 a)	Koerts et al. (2012)	Kudlicka et al. (2013)	Lanni et al. (2014)	Lovstad et al. (2016)	Mills et al. (2020)	Puente et al (2016	Siquier & Andres (2021)	Sobreira et al. (2008)	Vlagsma et al. (2017)
Objective measures											
WCST	x									x	
SCWT	x		x								x
DKEFS: TMT				x							
TMT B	x		x					x			x
The Trails Test									x		
FAB		x								x	
OMO			x								
BADS Zoo maps			x								x
DKEFS: CWI				x	x	x					
D-KEFS Verbal Fluency Test					x						
D-KEFS Category fluency subtest					x						

D-KEFS category Switching subtest,					x						
DKEFS The Tower subtest								x			
Hayling test									x		
Brixton Test								x			
Semantic verbal fluency			x							x	
Phonemic verbal fluency			x								x
Letter–word fluency								x			
Letter-Number Sequencing (WAIS-III)						x					
MOCA - visuospatial/executive domain score							x				
Digit span (WMS-R)			x								
Digit span (WAIS III)					x						
Digit span - inverse order (WAIS III)										x	
Symbol digit modalities test					x						

Visual Elevator Test subtest (Test of Everyday Attention)											x
Subjective measures											
DEX-self		x	x								x
DEX-other		x	x								x
DEX-Discrepancy score											
BRIEF-A (discrepancy scores)				x					x		
BRIEF-A-GEC						x			x		
BRIEF-A-BRI					x	x			x		
BRIEF-A-MI					x	x			x		
Informant GEC						x			x		
Informant BRI						x			x		
Informant MI						x			x		
FrSBe total (informant)								x			
FrSBe Apathy								x			
FrSBe disinhibition								x			

FrSBe Executive dysfunction								x			
MDRS Attention										x	
MDRS Initiation/ preservation										x	
MDRS conceptualisation										x	
SCOPA-COG-EF										x	
QUIP-RS									x		
BAFQ											x
BDEFS-TM											x
Neuro-QoL- EF							x				
Performance on Likert scale				x							
Subjective opinion on EF changes: Yes/no.	x										

See Key for table 1.1

Table 1.4 Summary quality appraisal for each study

AXIS Domains	Copeland et al. (2016)	Koerts et al. (2011 a)	Koerts et al. (2012)	Kudlicka et al. (2013)	Lanni et al. (2014)	Lovstad et al. (2016)	Mills et al. (2020)	Puente et al. (2016)	Siquier & Andres (2021)	Sobreira et al. (2008)	Vlagsma et al. (2017)
Quality of reporting	L	L	L	L	L	L	L	L	L	H	L
Study design quality	L	U	U	L	L	H	L	L	L	U	L
Introduction of bias	H	U	U	L	U	H	U	L	U	U	U

L = Low risk of bias (good quality), U = Unclear risk of bias, H = High risk of bias (poor quality)

Objective measures of EF

Overall, deficits in EF on objective measures were reported for PwPD on at least one test in nine studies (Copeland et al., 2016; Koerts et al., 2011a; Koerts et al., 2012; Kudlicka et al., 2013; Lanni et al., 2014; Puente et al., 2016; Siquier & Andres, 2021; Sobreira et al., 2008, Vlagsma et al., 2017). Two studies did not report on the scores of objective measures (Lovstad et al., 2016, Mills et al., 2020).

There was a large variation in neuropsychological tests of EF used, with 24 completed in total (see Table 1.3). Versions of the CWI and the TMT were most commonly used across studies, both being utilised in six studies (SCWT: Copeland et al., 2016, Koerts et al., 2011a, Vlagsma et al., 2017, DKEFS CWI: Kudlicka et al., 2013, Lanni et al., 2014., Lovestad et al., 2016; DKEFS TMT: Kudlick et al., 2013, TMT-B/A: Copeland et al., 2016, Koerts et al., 2012 Puente et al., 2016, Vlagsma et al., 2017, TESen: Siquier & Andrews, 2021). Studies found an objective deficit in EF on TMT tests, whereas objective deficits based on CWI were less clear. Lanni et al. (2014) found a significant objective deficit compared to controls when using the CWI. Whilst Koerts et al. (2012) and Vlagsma et al. (2017) found no significant difference between PwPD and controls on CWI tests, Copeland et al. (2016), and Lovestad et al. (2016) did not conduct specific analysis on the CWI to be able to report this. Kudlicka et al. (2013) used the CWI (and TMT) to group participants into those with and without EF deficits, suggesting some PwPD did display objective deficits on the CWI. All other tests are listed in Table 1.3 and objective deficits reported on Table 1.1.

Subjective measures of EF

Overall, subjective deficits in EF were reported in seven studies, suggesting a significant impact of EF on daily life (Koerts et al., 2011a; Koerts et al., 2012; Lanni et al., 2014; Lovstad et al., 2016; Puente et al., 2016; Siquier & Andres, 2021; Vlagsma et al., 2017). Three studies found no deficits in EF (Copeland et al, 2016; Kudlicka et al., 2013; Mills et al., 2020). However, Copeland et al. (2016) was found to be high risk from the introduction of bias due to a lack of a standardised measure of subjective EF, only asking dichotomous yes/no questions during clinical interview. Additionally, these questions were asked of participants and informants in the presence of each other, raising further issues of bias. Mills et al. (2020) reported T scores for subjective

measures which suggest EF was in the average range, however, there was no discussion of this result. Finally, with poor quality of reporting, Sobreira et al.'s (2008) findings of subjective EF were also unclear, noting low mean scores but no discussion of whether this implies a subjective EF deficit. Of note, all three studies which reported no or unclear subjective deficit did not include control groups, further reducing the reliability of their findings as direct comparison could not be made.

Only two subjective measures were used across multiple studies, the BRIEF-A and the DEX. The BRIEF-A was the most used subjective measure of EF utilised by four papers (Kudlicka et al., 2013; Lanni et al., 2014, Lovstad et al., 2016; Siquier & Andres, 2021). Siquier & Andres (2021) used discrepancy scores between self and informant ratings, finding a significant difference between PwPD and controls. Lanni et al. (2014) and Lovstad et al. (2016) also found a significant difference on the BREIF-A between PwPD and controls, however, only in the MI subdomain. Whereas Kudlica et al. (2013) found no significant differences between groups.

The DEX was used as the subjective measure in three studies, all reporting a significant deficit in subjective EF compared to controls on the DEX-self (Koerts et al., 2011a; Koerts et al., 2012; Vlagsma et al., 2017). All other subjective measures used are listed in Table 1.3. Notably, the remaining studies were those noted to have found no deficits or have unclear results. Additionally, they did not utilise HCs in their design.

Relationship between informant and self-report on subjective measures

Eight studies investigated the association between informant and self-report on subjective measures of EF (Copeland et al., 2016; Koerts et al., 2011a; Koerts et al., 2012; Kudlicka et al ., 2013; Lanni et al., 2014; Lovstad et al., 2016; Siquier & Andres, 2021; Vlagsma et al., 2017). Seven studies found no significant differences between self and informant report, suggesting high levels of agreement and indicating that PwPD have good insight into the impact of EF on daily life. Kudlicka et al. (2013) also found no significant differences between self and informants in those with EF deficits, however, in those without objective EF deficits PwPD reported significantly more EF difficulties daily life than informants. This finding may suggest that PwPD without objective EF deficits may identify subtle internal changes that are not yet observable to informants or objective measures.

Association between objective and subjective measures

All included studies investigated the association between objective and subjective measures for PwPD. Five studies found no significant association between objective and subjective measures (Copeland et al., 2016; Koerts et al., 2011a; Koerts et al., 2012; Mills et al., 2020; Vlagsma et al., 2017).

Within a sample of PD-MCI, Copeland et al. (2016) found very little agreement between objective and subjective measures of EF for both participants and informants. However, the study was found to be high risk from the introduction of bias, relating to use of a non-validated measure of subjective EF. Mills et al. (2020) also found no associations between MoCA subdomain EF scores and Neuro-Qol EF domain scores. This was the only study to use a cognitive screening measure (MoCA) rather than neuropsychological tests of EF when investigating the association between subjective and objective measures. However, as it is designed as a global measure, the MoCA may not be sufficient for domain specificity. Additionally, the sample was taken from a medical trial with strict inclusion criteria around newly diagnosed PD and thus lacks generalisability.

In comparison, Koerts et al. (2011a) examined whether objective measures of EF reflect difficulties in daily life using a standardised subjective measure (DEX). Analysis between objective and subjective measures showed that the DEX-self could not be explained by subtest scores on the FAB, alternately, the FAB total score could not be explained by the DEX-self subscales. Taken together with significantly worse scores on both the FAB and DEX and good agreement with informants, this suggests that not all PwPD who show objective deficits in EF report subjective decline and not all those who report difficulties with EF show objective deficits. Additionally, due to the global nature of the FAB it may not be the most appropriate neuropsychological test for detecting EF in PwPD.

In further research, Koerts et al. (2012) addressed the limitations of the previous study by utilising a larger sample and including multiple neuropsychological assessment measures of EF (TMT, CWI, OMO, BADS zoo maps). However, no significant correlations were found between DEX-self, DEX-other and EF objective test performance. Therefore, they argue the low ecological validity of neuropsychological

tests for detecting everyday EF difficulties may explain the lack of association between objective and subjective measures.

Vlagsma et al. (2017) went one step further, investigating the relationship between multiple measures of objective EF (SCWT, TMT B-A, BADS Zoo Maps, Phonemic Verbal Fluency, Visual Elevator Test subtest) and multiple subjective EF measures (DEX-self/other, BAFQ; BDEFS-TM) in PwPD and controls. However, despite PwPD displaying significantly worse performance on objective measures (TMT-B, visual elevator test, BADs Zoo Map) and reporting more subjective problems in daily life (DEX, BAFQ, BDEFS-TM), no significant relationship was found between objective test performance and self-reported EF deficits. Vlagsma et al. (2017) evidenced good quality of reporting and study design with a low risk of bias overall, recruiting from multiple sites, using HCs and multiple measures of EF. Koerts et al. (2012) and Vlagsma et al. (2017) were the only two studies with overlap on both objective measures (CWI, TMT) and subjective measures (DEX), giving more weight to their findings.

Only one study found a clear association between objective and subjective measures (Siquier & Andres, 2021), finding that the Hayling test positively correlated with BREIF-A and QUIP-RS. This was the only study to use the Hayling test as a measure of objective deficit, though with small sample sizes results may lack generalisability. Despite small sample sizes, the use of controls, quality of reporting and study design mean that the study was deemed low risk of bias overall.

All other studies describe a varied picture with some tests being associated and others not (Kudlicka et al., 2013; Lanni et al., 2014; Lovstad et al., 2016; Puente et al., 2016; Sobreira et al., 2008, see Table 1.1). The inconsistency between studies highlights that more commonly used tests of objective EF may not correspond with subjective everyday reports and overall, there does not appear to be an association between objective and subjective measure

Discussion

The aim of this review was to understand the relationship between objective deficits in EF, as measured by performance on neuropsychological tests and everyday executive functioning, as measured by subjective (self and other) reports.

Most studies reported an objective deficit on tests of EF in PwPD, in line with previous research (Kudlicka et al., 2011). Similar to Kudlicka et al. (2011), many different neuropsychological tests were used across studies, which is likely due to the overarching nature of EF as a higher-order cognitive process (Sira & Mateer, 2014).

What is the impact of PD on executive functioning in everyday life when measured by subjective measures?

The majority of studies reported a subjective deficit in EF in PwPD. Those which did not report a clear subjective deficit were found to be methodologically weaker. This suggests that there is likely a significant impact of EF on daily life in PwPD. Overall, there were fewer subjective tests compared to the number of objective tests used, though this may be because the questionnaire measures covered a wide range of aspects of executive functioning in everyday life (Puente et al., 2016). Despite the BRIEF-A being utilised most often, the DEX appears to be the subjective measure most consistently detecting a deficit in PwPD, with the studies utilising it finding a significant deficit compared to controls. Within subjective measures, all studies which investigated the discrepancy between self and informant report found good agreement. This indicates that PwPD have good insight into their own difficulties and accurately report the impact of EF in their daily lives. However, reviewing self and informant reported deficits on subjective measures alone does not necessarily mean these measures are effective at detecting EF impairment in everyday life.

What is the association between subjective and objective assessment of EF in PD?

Overall, the results suggest that there is no clear association between objective and subjective measures of EF, as measured by neuropsychological tests and self-report questionnaires. Five studies found no relationship, even when multiple objective and

subjective measures were utilised across different areas of EF. A further five studies showed a mixed picture, with no clear association between objective and subjective measures, but found associations between some specific subtests of neuropsychological tests and subdomains of self-report measures. Only one study found a clear association between an objective and subjective measure, finding that the Hayling test positively correlated with BREIF-A and QUIP-RS (Siquier & Andres, 2021). This was the only study to use the Hayling test as a measure of objective deficit and with small sample sizes, results may not be representative for all PwPD and cannot be generalised to all objective measures. Therefore, this indicates that there is no obvious relationship between commonly used neuropsychological tests of EF and the impact of EF on daily life through self-report measures for PwPD.

There are a few possible explanations for this finding. Firstly, the ecological validity of neuropsychological tests of EF is highly debated, with traditional tests often displaying poor ecological validity (Manchester, Priestley & Jackson, 2004). Ecological validity can be defined as the degree to which these tests within a controlled environment relate to performance in everyday life (Domensino et al., 2022). The neuropsychological tests used in most of the included studies are highly structured and often testing specific components of EF, whereas self-report measures typically cover a broad range of behavioural and emotional scenarios. This may explain why some studies found association between subsections of objective and subjective measures. Additionally, unlike the format of objective measures, the situations that make demands on EF difficulties in daily life will be unstructured and rely on other cognitive domains as well as EF (Koerts et al., 2011a).

Two key areas of focus when considering the ecological validity of neuropsychological tests is verisimilitude and veridicality (Domensino et al., 2022). Verisimilitude refers to the extent to which tests resemble tasks in everyday life, resulting in high face validity. Whereas veridicality refers to the extent to which a test predicts performances in everyday life. The EF deficits captured on neuropsychological tests within an artificial environment may have low veridicality and thus may not be useful in predicting difficulties in daily functioning (Chaytor & Schmitter-Edgecombe, 2003). This may account for the lack of agreement between types of measurement found in this review. With a primary aim of neuropsychological tests being the prediction of the impact of cognitive deficits on everyday functioning,

this raises questions over the utility of these measures for PwPD. It may be more helpful to consider tests which aim to reflect EF within daily life (verisimilitude), such as planning, problem solving and task management. For example, Koerts et al. (2011b) found that PD patients used compensatory strategies on the Cognitive Effort Test, which measures initiation, planning and task management, performing differently but not worse than controls, adjusting for their impairments. Compensatory strategies may be another factor which can potentially explain the lack of association between types of test, with traditional neuropsychological tests of EF not allowing for the use of compensatory strategies, whereas, these may be considered within self-reported views of daily functioning (Chaytor & Schmitter-Edgecombe, 2003).

However, Vlagsma et al. (2017) and Koerts et al. (2012) utilised the BADS Zoo Maps as an objective measure which resembles tasks of everyday life and is noted to have good ecological validity (Norris & Tate, 2000), but still found no association with a subjective measure. Therefore, as most studies evidenced both an objective and subjective deficit in EF in PwPD but no association between them, it is likely that objective and subjective measures are examining different aspects of EF. Specifically, neuropsychological tests measure deficit at the bodily functions level, whereas self-report questionnaires aim to measure the impact of deficits on performance in daily life, which is open to the influence of emotional, behavioural and environmental factors (Domensino et al., 2022). This would suggest that using both types of measure, as part of a multidimensional assessment of EF, would be most beneficial for understanding EF in PwPD in daily life (Domensino et al., 2022). This is supported by Puente et al. (2016) who found that neuropsychological tests were a significant predictor of IADLS in PD-MCI but the subjective self-report measure added incremental validity to neuropsychological tests for both PD-MCI and cognitively intact PwPD in explaining IADL performance. This raises implications for clinical practise as using both types of measure will assist clinicians to provide a more detailed neuropsychological assessment and clinical opinion to predict functioning (Vlagsma et al). Furthermore, as part of a multidimensional assessment, it may allow for more person-centred cognitive rehabilitation and support (Kudlicka, 2011; Domensino et al., 2022).

What is the methodological quality of the available evidence?

There are a few key limitations that should be kept in mind when considering these findings. Firstly, no included study justified its sample size, with none reporting a power calculation. Therefore, some studies may be underpowered, impairing their ability to detect an effect. Secondly, with the many different measures and analyses used it was not possible for meta-analysis to be conducted. As previously noted, this likely relates to the overarching nature of EF. A key issue for research within EF and investigation into the ecological validity of testing is the lack of agreement around the construct of EF, which is reflected in the broad range of tests used within included studies (Chaytor & Schmitter-Edgecombe, 2003). Bearing this in mind, future research may consider using a specific battery of tests recommended for use in PD, for example the MDS level II suggested tests (Dubois et al., 2007). Additionally, nearly all studies failed to use self-report measures specific to PD. With a distinctive cognitive profile characterised by dysexecutive features (Dubois et al., 2007) and with the unique interplay between motor and non-motor symptoms in PwPD, the development or use of measures specific to this population would be valuable.

Furthermore, the heterogeneity of the sample makes it difficult to draw strong conclusions and may have a significant impact on results. For example, many studies had inclusion criteria around mild to moderate disease severity, excluding those at H&Y stages 4+. Whereas, because of recruiting from a database for DBS one study had inclusion criteria of at least five years disease duration and severe motor difficulties. Koerts et al. (2012) found that those at moderate disease severity reported more problems on subjective EF measures than relatives, compared to mild PD where good agreement was found. Additionally, whilst many studies excluded patients based on global cognitive impairment scores, one study specifically focused on those with PD-MCI. Future studies should consider investigating differences in the association between objective and subjective measures and awareness of EF difficulties at both different H&Y stages and across a range of cognitive difficulties. This would allow for greater understanding of the mechanisms behind this relationship and inform clinical practise.

A limitation of this review was that only 10% of abstracts were screened by a second reviewer. With full text screening and data extraction only completed by the first reviewer. Therefore, there is the potential that studies or data were excluded in

error. Along these lines, intervention study designs were excluded, so intervention studies which reported an association between neuropsychological test scores and subjective reports may have been missed. Additionally, only five of the included studies were appraised for methodological quality by both reviewers. However, inter-rater reliability was high.

Conclusions

This review aimed to investigate the relationship between objective neuropsychological tests of EF and subjective report measures of EF difficulties experienced in everyday life in PwPD. Despite deficits evidenced on both objective and subjective measures, no clear association was found between types of tests suggesting that these may be measuring different aspects of EF in PwPD. However, with heterogeneity in both the sample and measures used, it is difficult to draw strong conclusions. Future research into subjective reports of EF across disease severity and its association to neuropsychological tests is necessary to further clarify the nature of this relationship and inform clinical practise. Findings still have important clinical implications highlighting the need for clinicians to use multiple types of measure, with real-world applicability in mind, for a more comprehensive assessment of EF difficulties. Doing so would ensure appropriate support is provided.

Reference list

Aarsland, D., Creese, B., Politis, M., Chaudhuri, K., Weintraub, D., & Ballard, C. (2017). Cognitive decline in Parkinson disease. *Nature Reviews Neurology*, *13*(4), 217-231.

Adams, J., Hillier-Brown, F. C., Moore, H. J., Lake, A. A., Araujo-Soares, V., White, M., & Summerbell, C. (2016). Searching and synthesising 'grey literature' and 'grey information' in public health: critical reflections on three case studies. *Systematic reviews*, *5*(1), 1-11.

Burgess, P. W., Alderman, N., Forbes, C., Costello, A., Laure, M. C., Dawson, D. R., ... & Channon, S. (2006). The case for the development and use of "ecologically valid" measures of executive function in experimental and clinical neuropsychology. *Journal of the international neuropsychological society*, *12*(2), 194-209.

Chaytor, N., & Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. Neuropsychology review, 13(4), 181-197.

Copeland, J. N., Lieberman, A., Oravivattanakul, S., & Tröster, A. I. (2016). Accuracy of patient and care partner identification of cognitive impairments in Parkinson's disease–mild cognitive impairment. *Movement Disorders*, *31*(5), 693-698.

Domensino, A. F., Evans, J., & van Heugten, C. (2022). From word list learning to successful shopping: The neuropsychological assessment continuum from cognitive tests to cognition in everyday life. Applied Neuropsychology: Adult, 1-8.

Downes, M. J., Brennan, M. L., Williams, H. C., & Dean, R. S. (2016). Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). *BMJ open*, *6*(12), e011458.

Dubois, B., Burn, D., Goetz, C., Aarsland, D., Brown, R. G., Broe, G. A., ... & Emre, M. (2007). Diagnostic procedures for Parkinson's disease dementia: recommendations from the movement disorder society task force. Movement disorders, 22(16), 2314-2324.

Emre, M. (2003). Dementia associated with Parkinson's disease. *Lancet Neurology, 2*(4), 229-237.

Koerts, J., Tucha, L., Leenders, K. L., van Beilen, M., Brouwer, W. H., & Tucha, O. (2011a). Subjective and objective assessment of executive functions in Parkinson's disease. Journal of the neurological sciences, 310(1), 172-175.

Koerts, J., van Beilen, M., Leenders, K. L., Brouwer, W. H., Tucha, L., & Tucha, O. (2012). Complaints about impairments in executive functions in Parkinson's disease: The association with neuropsychological assessment. Parkinsonism & Related Disorders, 18(2), 194-197.

Koerts, J., Van Beilen, M., Tucha, O., Leenders, K. L., & Brouwer, W. H. (2011b). Executive functioning in daily life in Parkinson's disease: initiative, planning and multi-task performance. PloS one, 6(12), e29254.

Kudlicka, A., Clare, L., & Hindle, J. V. (2011). Executive functions in Parkinson's disease: Systematic review and meta-analysis. *Movement disorders*, *26*(13), 2305-2315.

Kudlicka, A., Clare, L., & Hindle, J. V. (2013). Awareness of executive deficits in people with Parkinson's disease. Journal of the International Neuropsychological Society : JINS, 19(5), 559-570.

Lanni, K. E., Ross, J. M., Higginson, C. I., Dressler, E. M., Sigvardt, K. A., Zhang, L., . . . Disbrow, E. A. (2014). Perceived and performance-based executive dysfunction in Parkinson's disease. Journal of clinical and experimental neuropsychology, 36(4), 342-355.

Lezak, M. D. (2012). Neuropsychological assessment (5th ed.). Oxford University Press, USA Pg 4-7

Løvstad, M., Sigurdardottir, S., Andersson, S., Grane, V. A., Moberget, T., Stubberud, J., & Solbakk, A. K. (2016). Behavior Rating Inventory of Executive Function Adult Version in Patients with Neurological and Neuropsychiatric Conditions: Symptom Levels and Relationship to Emotional Distress. Journal of the International Neuropsychological Society : JINS, 22(6), 682-694.

Manchester, D., Priestley, N., & Jackson, H. (2004). The assessment of executive functions: Coming out of the office. Brain injury, 18(11), 1067-1081.

McKinlay, A., Grace, R. C., Dalrymple-Alford, J. C., Anderson, T. J., Fink, J., & Roger, D. (2008). Neuropsychiatric problems in Parkinson's disease: comparisons between self and caregiver report. *Aging and Mental Health*, *12*(5), 647-653.

Mills, K. A., Schneider, R. B., Saint-Hilaire, M., Ross, G. W., Hauser, R. A., Lang, A. E., . . . Marras, C. (2020). Cognitive impairment in Parkinson's disease: Associations between subjective and objective cognitive decline in a large longitudinal study. Parkinsonism & Related Disorders, 80, 127-132.

NHS (2019) Parkinson's Disease. Retrieved from: https://www.nhs.uk/conditions/parkinsons-disease/

Norris, G., & Tate, R. L. (2000). The Behavioural Assessment of the Dysexecutive Syndrome (BADS): Ecological, concurrent and construct validity. Neuropsychological rehabilitation, 10(1), 33-45.

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Systematic reviews*, *10*(1), 1-11.

Puente, A. N., Cohen, M. L., Aita, S., & Brandt, J. (2016). Behavioral ratings of executive functioning explain instrumental activities of daily living beyond test scores in Parkinson's disease. *The Clinical Neuropsychologist*, *30*(1), 95-106.

Siquier, A., & Andres, P. (2021). Cognitive and Behavioral Inhibition Deficits in Parkinson's Disease: The Hayling Test as a Reliable Marker. Frontiers in aging neuroscience, 12, 621603.

Sira, C. S., & Mateer, C. A. (2014). Executive Function. Encyclopedia of the Neurological Sciences, 239–242.

Sobreira, E. S. T., Pena, M. C. S., Filho, J. H. S., Souza, C. P., Oliveira, G. N., Tumas, V., & de Assis Carvalho do Vale, F. (2008). Executive cognitive tests for the evaluation of patients with Parkinson's disease. Dementia & neuropsychologia, 2(3), 206-210.

Vlagsma, T. T., Koerts, J., Tucha, O., Dijkstra, H. T., Duits, A. A., van Laar, T., & Spikman, J. M. (2017). Objective Versus Subjective Measures of Executive Functions: Predictors of Participation and Quality of Life in Parkinson Disease? Archives of Physical Medicine & Rehabilitation, 98(11), 2181-2187.

Chapter 2

A single case experimental design study of a reminder app for supporting adherence to personalised treatment goals in Parkinson's disease

Prepared in accordance with the author requirements for *The Clinical Neuropsychologist*

https://www.tandfonline.com/action/authorSubmission?show=instructions&journalCode=ntcn20#word-limits

Plain Language summary

Title

A single case experimental design study of a reminder app for supporting adherence to personalised treatment goals in Parkinson's disease

Background

Parkinson's disease (PD) is a progressive condition which affects the nervous system causing tremors and slowed movement. Other symptoms include problems with balance, sleep, memory and anxiety. Treatment aims to improve symptoms and quality of life. Poor memory in PD has been shown to reduce a person's ability to manage activities of daily living and treatments, such a taking medications. However, patients with PD have been shown to benefit greatly from external reminders.

Aims and Questions

The aim of this research project was to investigate whether the use of a reminder app improves adherence to treatment goals in people with PD.

Methods

Patients seen by the movement disorder clinic with a diagnosis of PD and a partner who was able to monitor progress were invited to take part in the study. Written consent to take part was acquired from each participant. Measures of memory, emotional wellbeing and quality of life were completed to build a profile of each participant. Participants had a 'baseline' phase at the start where they carried on as usual, but their partner recorded completion of daily goal-related tasks. The baseline phase varied in length for each person. Participants created specific and personalised treatment goals at the beginning of the baseline phase. Goals were relevant to the management and treatment of PD. Then, an intervention phase was completed with the ApplTree app (Jamieson et al., 2020) introduced to remind participants of treatment goals. This design is known as a single case experimental design. Partners monitored their goal adherence, via a weekly monitoring form in both phases.

Results

Visual analysis of graphs revealed that most but not all participants showed improvement from baseline to intervention phase, though only two participants showed statistically significant change between phases using statistical analysis.

Conclusions

This research provided evidence that a smartphone reminder app can improve adherence to goal-related tasks for some people with PD but not all, enabling patients to take a more active role in self-management of PD. Some participants showed high task completion in the baseline phase. It may be that the goal setting session and interaction with the researcher motivated participants to completed daily goal-related tasks before introduction of the reminder app. Included participants measures of memory showed they had no difficulties; future research could look at the effect of the reminder app in people with PD who have memory difficulties.

References

Jamieson, M., Cullen, B., Lennon, M., Brewster, S., & Evans, J. (2020). Designing ApplTree: Usable scheduling software for people with cognitive impairments. *Disability and Rehabilitation: Assistive Technology*, 1-11.

Abstract

Parkinson's Disease (PD) is a progressive neurological condition with a specific profile of cognitive deficits, including executive function and episodic memory. Treatment management is often complex and challenging for individuals with PD. Deficits in prospective memory (the ability to recall intended actions after a delay) in PD are associated with poorer functioning and reduced autonomy, including self-management of treatments. PD patients have more difficulty with internally cued behaviour and benefit significantly from external cues. The aim of this research was to investigate whether the use of a reminder app, providing external cues, improves completion of goal-related tasks in PD patients. A single-case experimental design was utilised with a randomised multiple baseline phase prior to an intervention phase (AB). Seven participants were recruited from a movement disorder service, six completing the study. Measures of cognitive impairment, apathy, prospective memory, PD-related quality of life and mental wellbeing were completed to characterise participants. Task adherence was recorded via a weekly monitoring form by a nominated person. Visual analysis of graphs revealed that most, but not all, participants showed improvement from baseline to intervention phase, though only two participants showed statistically significant change between phases using Tau-U analysis. This research provided evidence that a smartphone reminder app can improve treatment-related goal adherence for some people with PD, enabling patients to take a more active role in self-management of PD. Future research should consider investigating the impact of a reminder app in those with specific deficits, considering the most objective ways to record this data.

Keywords: Parkinson's disease, prospective memory, apathy, self-management, reminder app, smartphone, adherence

Introduction

Parkinson's disease (PD) is a progressive neurological condition affecting around one in every 500 people, with most people developing symptoms after 50 years of age (NHS, 2019). PD affects people differently, but three main motor symptoms include tremors, slowed movement and stiff muscles (NHS, 2019). Other symptoms can include balance and gait disturbance and non-motor symptoms such as anxiety, insomnia and cognitive impairment. There is no cure for PD, but treatments aim to help reduce the main symptoms and improve quality of life. Often non-motor symptoms can have the greatest impact on day-to-day functioning, with depression, psychosis, cognitive impairment, apathy and sleep disorders being found to be the most common independent determinants of health-related quality of life and increased disability (Raggi et al., 2012).

Apathy, characterised by a reduction in goal-orientated thoughts and behaviour, is a common symptom in PD, with a meta-analysis stating a prevalence rate of 40% (den Brok et al., 2015). Analysis revealed that apathy was highly associated with cognitive impairment, comorbid depression, disease severity and disability in PD. This is likely due to the overlap in symptoms between depression, apathy and cognitive impairment in PD (den Brok et al., 2015). When comorbid depression and cognitive impairment were controlled for, apathy was still present in 50% of patients, confirming ideas that it is a separate clinical symptom present in PD.

Evidence suggests that people with PD may develop a specific profile of cognitive deficits, particularly in executive function (EF), episodic memory and visual-spatial abilities, in the early stages of the disease and are at greater risk than the general population of developing dementia (Aarsland et al., 2017). In a review of the cognitive profile of PD and PD-related dementia, Emre (2003) notes that impairment in EF, which can be defined as an ability to plan, organise and regulate goal-directed behaviour, is a common clinical feature in patients with PD. As a result of this dysexecutive syndrome, PD patients have considerable problems with internally cued behaviour, due to difficulties shifting attention, and benefit significantly from external cues (Emre, 2003). A number of studies have now evidenced the effectiveness of visual, auditory and cutaneous external cues to aid initiation of behaviour in PD (Butterfield et al., 2017).

Prospective memory (PM) relates to the ability to carry out intended actions after a delay, at a particular time (e.g. attending an appointment at 10am) or in association with a particular event (e.g. posting a letter on the way to work). A systematic review of literature investigating PM functioning in PD revealed that both EF and episodic memory difficulties are associated with PM performance in PD (Costa, Caltagirone & Carlesimo, 2018). Costa et al. (2018) also report that cognitive interventions may be effective in improving PM function in PD patients. Additionally, studies have shown that patients with Parkinson's related dementia have impaired free recall but benefit greatly from semantic cueing, suggesting new information is stored but not easily accessed (Emre, 2003). Deficits in PM in PD have been shown to be significantly associated with poorer functional ability and reduced autonomy, including the self-management of treatments and activities of daily living, such as medication management (Pirogovsky, Woods, Vincent Filoteo, & Gilbert, 2012).

Medication is often used to improve symptoms of PD, such as tremors and movement problems, as well as manage side effects of treatment and comorbidities (Parkinson's UK, 2019). The number of medications and the frequency with which they are taken typically increases with disease progression (Schapira et al., 2009). As a result, self-management and treatment adherence is often complex and challenging (Lakshminarayana et al., 2017). It has been demonstrated that withdrawal of dopamine medical treatments (e.g. Levodopa) in PD, leading to characteristic 'off states', results in significantly poorer PM performance (Costa et al., 2008). Therefore, it is likely poor medication adherence could be both a contributor to and consequence of deficits in PM and could lead to reduction in adherence to other goals and tasks. To add to this challenge, clinicians often have limited time available for face-to-face consultations, making it difficult to fully assess and address issues with self-management and non-adherence (Lakshminarayana et al., 2017).

Non-adherence to prescribed therapy in PD ranges between 10-67% (Malek & Grosset, 2015). Difficulty with adherence to treatment in PD is significantly associated with poorer motor scores, more daily 'off' time where PD symptoms increase between medication doses and worse mobility compared to patients with satisfactory compliance (Grosset et al., 2009). Additionally, non-compliance in PD is associated with higher rates of depression and reduced quality of life (Grosset, Bone & Grosset, 2005). A range of non-pharmacological treatments for management of PD are

recommended in the NICE guidelines (2017) alongside medication including physiotherapy, speech and language therapy, occupational therapy, psychological interventions and nutrition. However, most studies to date have focused solely on medication adherence and its consequences.

Providing patient-centred care has been shown to increase treatment adherence in patients with PD, with self-management support identified as a key way to support and empower these individuals (van der Eijk et al., 2011). Self-management relates to increasing patients' participation in goal setting, treatment planning, and independent management of a range of aspects of their treatment/care, increasing the experience of control over their lives (Kralik, Koch, Price & Howard, 2004).

A recent systematic review into mobile apps for medication adherence found that people who use reminder apps are significantly more likely to adhere to medication regimes than those who do not (Armitage, Kassayou & Sutton, 2020). However, findings for six of the nine studies included in the meta-analysis were based on self-reported measures of adherence, so they state results should be interpreted with caution. Grosset et al. (2006) found that, due to significant under-reporting by patients, self-report measures are insensitive in detecting sub-optimal PD medication intake; therefore, objective methods should be used. Additionally, Armitage et al. (2020) found that studies with interventions matched to patients' specific needs, beyond simple medication regime reminders, such as interventions focused on wider treatment goals, mood, cognitive impairment, or symptom control, had the largest effect sizes. This is in line with Cabrera-Martos et al. (2018) who found that people with PD show greater improvement in goal attainment if treatment is focused on a specified tailored set of goals.

The formulation of goals helps to initiate goal-directed thoughts and behaviour through motivation, planning and intention development (Locke & Latham, 2002), all of which have been evidenced as impaired in PD. Research into retrospective and prospective memory suggests that PD patients are unable to internally initiate encoding strategies related to intention formation (Foster, Rose, McDaniel & Rendell, 2013). A RCT found that strategies supporting encoding of PM cues and associated behaviour, such as creating implementation intentions, improved performance in PD patients within a laboratory setting (Foster, McDaniel & Rendell, 2017). However, they

state that future research should investigate implementing strategies into patients' everyday lives to improve patients' clinical care, daily functioning and quality of life.

Deficits in initiation (internal self-generation) are a well-documented feature of PD which affects multiple domains, including motor and cognitive (Butterfield et al., 2017). These deficits have shown to be improved by external cues, for example, visual cues have been shown to improve stride length and auditory cues improve gait, length and rhythm of strides in PD patients (Butterfield et al., 2017). Pagni et al. (2011) found that even PD patients in the early stages who have not yet received dopamine medication performed worse on the prospective component of event-based tasks compared to controls but not the retrospective component. This suggests that it is the initiation phase of tasks that is most impaired and, as evidenced, external cues as aids would be helpful.

A recent review of mobile apps for self-management in PD highlights the use smartphone apps as a relatively new area for PD interventions and notes that most research has focused on self-monitoring of symptoms rather than management (Lee et al., 2022). Despite the push for person-centred self-management, no study so far has investigated the use of a reminder app for personalised treatment goals in PD, with the aim of aiding self-management. Lee et al. (2022) highlight that the relatively older age of those with PD and the impact of motor symptoms may impact on smartphone use for treatment management. Additionally, research into health apps suggests people prefer those which are simple and straightforward to use (Peng, Kanthawala & Yuan, 2016). ApplTree is a reminding app designed to be used by people with cognitive impairments related to acquired brain injury and as a result has been developed with simplicity in mind (Jamieson et al., 2020). ApplTree has a narrow/deep user interface design which presents a small amount of information at a time to minimise the cognitive burden on people when setting reminders (Jamieson et al., 2020). Therefore, as people with PD (PwPD) often display a specific cognitive profile, with deficits resulting in difficulties with initiation and internally cued behaviour, a reminder app which is simple to use may be most effective at improving adherence to personalised treatment goals.

Aims

The aim of this study was to investigate whether the use of a recently developed reminder app, ApplTree (Jamieson et al., 2020), improved adherence to personalised treatment goals in people with PD. It was hoped that the formulation of specific, measurable, achievable, relevant and time-limited (SMART; Bovend'Eerdt, Botell, & Wade, 2009) treatment goals in collaboration with the person with PD would increase the encoding of goals, whilst the associated external visual and auditory reminders would both cue participants to PM tasks and increase initiation of goal-directed behaviours.

Hypothesis

The introduction of the reminder app will significantly improve completion of goal-related tasks.

Methods

Research has been reported in accordance with Single-Case Reporting Guidelines in Behavioural Interventions (SCIBE, Tate et al., 2016). The protocol for this study was registered on ClinicalTrials.gov (NCT05106985) prior to recruitment.

Design

The study design was developed with reference to the methodological quality criteria for single-case experimental designs and n-of-1 trials (Risk of Bias in N of 1 trials – RoBiN-T, Tate et al., 2013) A single case experimental design (SCED) was utilised with multiple baselines across participants with randomisation of the onset of the intervention phase for each participant.

A two phase (AB) design was used with participants randomised to different lengths of baseline. Blinding of participants and researcher was not possible and a withdrawal design (e.g. ABA) was not deemed ethical due to the nature of the study. Quantifying datapoints required at least one reminder per day and hence a two-day period represented one datapoint. The data collection phase for each participant was based on completing at least five data points within both the baseline and intervention phase. Participants were randomly assigned to a baseline period of 10, 16 or 22 days using an online randomiser programme (http://randomiszer.org). All participants then completed an intervention period of 22 days where the ApplTree app was introduced. Study length was therefore due to range from 31-43 days for each participant, however, due to extraneous circumstances study length ranged from 22-53. Variations in planned design are discussed within individual participant results. The study was run in batches of three participants, with replication of the design across the second set of three to ensure at least three demonstrations of the treatment effect in each participant set (Tate et al., 2013). However, baseline phases were not concurrent due to recruitment challenges.

Participants

Participants with a clinical diagnosis of PD were recruited via NHS Greater Glasgow and Clyde's Movement Disorders Team, through the Neurology Service's Movement

Disorder Clinics or by the Older People's Psychology Service. Patients who clinicians felt would benefit from the reminder app to assist with treatment goals were invited to participate. Participants were required to own a smart phone and have a partner or significant other who was able to, and consented to, monitor and support the participant's goal attainment as the nominated person. Potential participants and their carers/partners (nominated person) were sent study invitation and participant information sheets by post or were given them by a member of their clinical team during a clinic appointment (Appendix 2.1, 2.2). Potential participants were able to indicate their willingness to discuss participation in the research by returning a consent for contact form (Appendix 2.3) or informing the clinical team member at their clinic appointment of their agreement for contact. Participants and nominated persons who consented to contact received a phone call to discuss the study in more detail and, if ready, decide on whether they would like to participate.

Exclusion criteria

- A diagnosis of dementia
- Pre-existing neurological or severe and enduring psychiatric disorder
- Sensory deficits preventing the use of a smart device
- Did not own a smart phone capable of downloading apps
- Were involved in any other research study
- Lack capacity to consent

Seven participants were recruited initially. One participant (P04) withdrew prior to commencement of the intervention phase, noting difficulties keeping up with life demands. A second participant (P07) stopped recording data partway through the intervention phase due to going on holiday. Both participants' data collected prior to withdrawal has been included. One participant (P01) was unavailable to commence the intervention phase on the agreed date, and as a result an extended baseline was continued until they were available to commence intervention. This data has also been included. Participant characteristics, including pre-intervention measures, are reported in Tables 2.1, 2.2.

The following measures were used to characterise each participant. The below measures are validated for PD and recommended by the Movement Disorder Society (MDS, 2021) unless specified:

- The Montreal Cognitive Assessment (MoCA, Nasreddine et al., 2005).
- Parkinson's Disease Quality of Life Questionnaire (PDQ-39 Peto, Jenkinson & Fitzpatrick, 1998)
- Prospective and Retrospective Memory Questionnaire (PRMQ, Smith et al., 2000)
- Apathy Scale (AS, Starkstein et al., 1992) self-report version, 14-item questionnaire.

Currently, no measure of anxiety is recommended by the MDS, therefore, the Generalised Anxiety Disorder Severity Index(GAD-7; Spitzer et al., 2006) and the Patient Health Questionnaire (PHQ-9; Spitzer, Kroenke & Williams, 1999) were used to assess emotional wellbeing. In an internal systematic review and audit by NHS Greater Glasgow and Clyde Older People's Psychology Service, results indicated that the PHQ-9 and GAD-7 were the most suitable outcome measures due to strong psychometric properties in their use with older people (NHS GG&C, 2019). These measures are the routine outcomes measures for NHS Greater Glasgow and Clyde Older People's Psychology Service.

The following measures were completed by the nominated person:

- PDQ-39-Carer (Peto, Jenkinson & Fitzpatrick, 1998)
- PRMQ-Proxy (Crawford, Henry, Ward & Blake, 2006).
- Apathy Scale – proxy version (Adapted from Starkstein et al., 1992)

Ethics

Ethical approval was obtained from the West of Scotland Research Ethics Committee 3 (Appendix 2.4, 2.5) and management approval obtained from NHS Greater Glasgow and Clyde (Appendix 2.6, 2.7). Informed consent was obtained from all participants and their nominated persons (Appendix 2.8, 2.9).

Procedure

Participants met with the researcher initially to read through and sign the consent forms and complete pre intervention measures. All meetings were held in the participants' own homes or remotely via MS Teams. Random allocation to baseline phase (10, 16 or 22 days) was completed following the first meeting. Participants met with the researcher a second time to formulate specific and personalised treatment goals, prior to starting the baseline phase. Discussion around goals was informed by the Smart goal worksheet (Appendix 2.10), information from involved clinicians, pre-interventions measures, partners and carers but focused primarily on participants' views. Goals were relevant to the management and treatment of PD and associated symptoms and had a PM and/or motivational component. See Appendix 2.11 for goals for each participant.

The baseline phase then commenced with the nominated person completing a monitoring form to record completion of intended goals (Appendix 2.12). The nominated person received a reminder SMS text message daily to prompt completion of the monitoring record. At the end of the baseline phase a further meeting with the researcher was held to set up the app with specific goals and provide guidance on its use. Nominated persons continued to monitor goal adherence through the intervention phase (22 days) with reminder texts continuing. Following completion of the intervention phase, a final debrief meeting was completed. Those wishing to continue using a reminder app were assisted to do this at the end of the study.

Measures and Materials

A weekly/daily monitoring form was used by the nominated person during both baseline and intervention phase to record completion of intended goal-related tasks (e.g. taking medication on time, completing prescribed exercises, nutritional snack, attending support group).

A smart goal worksheet was developed to ensure goals were relevant to the management and treatment of PD and associated symptoms and had a PM or motivational component such as reminders for medication regime or physical exercises.

ApplTree is a reminding app designed to be used by people with cognitive impairments as a result of acquired brain injury (Jamieson et al., 2020). Each participant downloaded the ApplTree app onto their own smart device. The app is free to use. ApplTree is owned by the University of Glasgow and is maintained by Glasgow University Software Services (GUSS).

Data analysis

The primary outcome measure was calculated as the percentage of goal-related tasks carried out independently (i.e. without prompting from the nominated person) per two-day period. The primary outcome is measured via the daily monitoring form, completed by each participant's nominated person. Some participants provided additional data outwith the specified baseline and intervention phases, and this has been included in analysis.

In line with recommendations for analysis of SCED data (Tate et al., 2013), the data was analysed through visual analysis of graphs with reference to visual inspection guidelines described by Lane and Gast (2014). Tau-U analysis was also conducted to determine whether a significant improvement was found between baseline and intervention phase (Parker, Vannest, Davis, & Sauber, 2011).

Results

Participant characteristics and cognitive measures are described in Table 2.1 and measure of emotional wellbeing in Table 2.2. The Apathy Proxy Rating Scale is not reported for two participants, one participant withdrew prior to return of the measure and the other nominated person was unavailable to complete the measure. All participants were recruited through their consultant neurologist at the movement disorders clinic. All participants decided to use the app for goals around increasing activity and exercise and nearly all used the app to assist with taking regular medications. For example, taking medications at the same time each day, every four hours at specified times or completing exercises once day. See appendix 2.11 for goal-related tasks for each participant.

Table 2.1 Participant characteristics and cognitive measures

	P01	**P02**	**P03**	**P04**	**P05**	**P06**	**P07**
Age	60	63	63	46	72	65	49
Sex	M	M	M	F	M	F	M
Nominated person	Partner	Partner	Partner	Partner	Partner	Partner	Partner
Time since diagnosis	14 years	4 years	3 years	4 years	7 years	3 years	5 years
Levodopa treatment	Yes	Yes	Yes	No	Yes	Yes	Yes
MoCa	27 Normal	28 Normal	26 Normal	28 Normal	29 Normal	28 Normal	Blind 20/22 Normal
PRMQ T score (CI 95%)	66 (58-70) Superior	47 (41-53) Average	62 (55-66) High average	38 (33-45) Low average	59 (52-63) High average	49 (43-55) Average	66 (58-70) Superior
PRMQ Prospective T score (CI 95%)	62 (54-67) High average	42 (37-50) Low average	56 (49-62) Average	32 (29-42) Borderline	60 (52-65) High average	50 (44-57) Average	70 (61-74) V. superior
PRMQ Retrospective T score (CI 95%)	65 (55-69) Superior	51 (44-58) Average	65 (55-69) Superior	45 (39-53) Average	55 (47-61) Average	47 (41-55) Average	57 (49-63) Average
PRMQ – Proxy T score (CI 95%)	61 (55-65) High Average	58 (53-63) High Average	61 (55-65) High Average	47 (42-52) Average	61 (55-65) High Average	50 (45-55) Average	65 (59-69) Superior
PRMQ P -prospective T score (CI 95%)	58 (51-63) High average	58 (51-63) High Average	60 (53-65) High Average	40 (35-48) Low average	62 (55-67) High Average	49 (43-55) Average	67 (59-71) Superior
PRMQ p-Retrospective T score (CI 95%)	62 (53-67) High Average	57 (49-63) Average	60 (51-65) High Average	55 (47-61) Average	57 (49-63) Average	51 (44-57) Average	62 (53-67) High Average

Table 2.2 Participant emotional wellbeing measures

	P01	**P02**	**P03**	**P04**	**P05**	**P06**	**P07**
GAD-7 (Anxiety)	1 Minimal	5 Mild	8 Mild	10 Moderate	7 Mild	8 Mild	3 Minimal
PHQ-9 (Depression)	2 None	2 None	15 Moderately severe	9 Mild	6 Mild	3 None	2 None
Apathy scale ≥14	14 Apathetic	17 Apathetic	32 Apathetic	21 Apathetic	10 Not Apathetic	10 Not Apathetic	8 Not Apathetic
Apathy scale proxy	-	26 Apathetic	23 Apathetic	-	7 Not Apathetic	10 Not Apathetic	10 Not Apathetic
PDQ- SI	23	16	45	19	22	36	29
Mobility	40	8	70	15	10	45	8
ADLS	71	17	88	38	21	17	25
Emotional wellbeing	21	21	46	33	42	42	42
Stigma	0	13	69	13	25	25	25
Social support	0	0	0	0	17	25	33
Cognitions	19	25	44	31	19	38	25
Communication	0	17	33	17	8	50	58
Bodily Discomfort	33	25	8	8	33	52	17
PDQ-C-SI	22	9	68	3	6	23	9
Social & Personal	35	6	65	2	8	15	13
Anxiety & Depression	25	17	75	4	8	33	4
Self-care	15	5	70	0	5	20	5
Stress	13	4	63	4	4	25	13

The below graph displays the overall percentage of completed goal-related tasks with the baseline (Phase A) and intervention phase (Phase B).

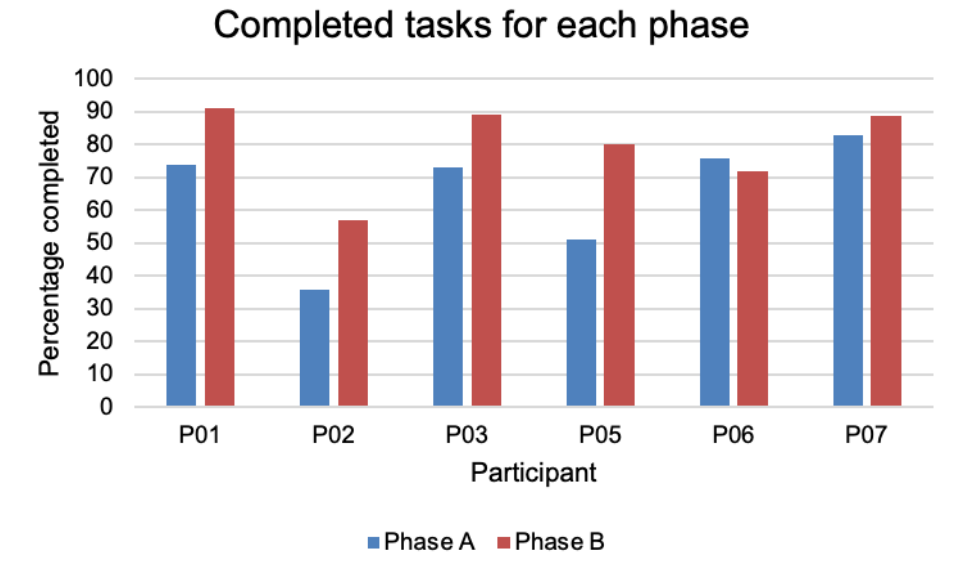


Figure 2.1 – Percentage of completed tasks per phase for each participant

Visual analysis

To investigate the effect of introducing the ApplTree app on goal related tasks, visual analysis of the percentage of completed tasks across both phases was completed. Visual inspection included the stability, level and trend of data within and between conditions as described by Lane and Gast (2014) The below graphs (Figure 2.2 & Figure 2.3) display the percentage of completed tasks per two-day datapoint within the baseline and intervention phases, with varying lengths of baseline phase.

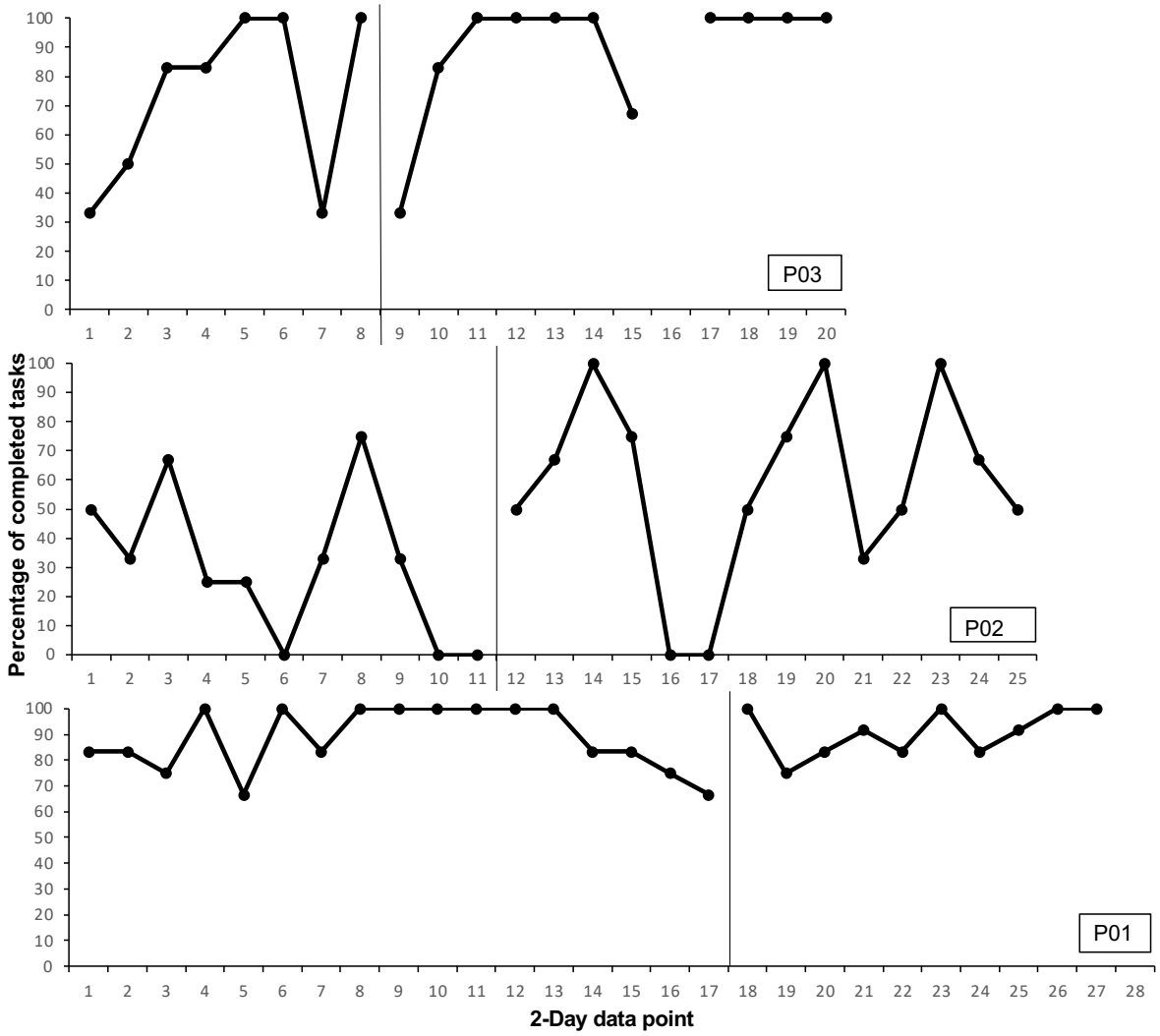


Figure 2.2 Participants percentage of completed goal-related tasks per two-day period, presented in order of baseline length (16, 21 & 33 days)

Participant 1

Participant 1 completed an extended baseline beyond the randomised 10 days he was allocated to, due to personal circumstances meaning he was unable to meet to download the app. Participant 1's final datapoint of the intervention phase was missing (point 28).

Participant 1's performance completing goal-related tasks showed an improvement from 74% (147/198) during baseline to 91% (109/120) during the intervention phase, with data being stable in both conditions. Split-middle method of trend estimation was completed, suggesting there was an accelerating therapeutic

trend within both phases, with data considered stable following the application of a stability envelope to trend lines (see appendix 2.13).

Between-condition analysis showed that mean, median and absolute level change measures suggested an improvement across conditions. However, Tau-U analysis into performance change between phases found that changes were not significant, suggesting no significant improvement with the introduction of the reminder app (Tau-U A vs B = 0.09, 90% CI [-0.30-0.47], p = 0.71).

Participant 2

Participant 2 was randomly allocated a baseline period of 22 days. Participant 2 had missing data within datapoint 11, so the datapoint was calculated with the remaining recorded data. Participant 2 also recorded additional datapoints at the end of the intervention phase, which were included in analysis.

Participant 2's goal performance showed an overall improvement completing 36% (13/36) during baseline and 57% (28/49) during intervention, however, data was variable during baseline and intervention phases. Evaluation of level change within conditions indicated performance was deteriorating during baseline and improving during intervention. Split-middle method of trend analysis was utilised and indicated that there was a zero-celerating trend during baseline and an accelerating therapeutic trend during intervention. However, data was considered variable following the application of a stability envelope to trend lines. Evaluation of trend between conditions, suggests a change in performance from a zero-celerating trend in baseline to accelerating improving trend during intervention. All level change measures suggest a positive change across conditions.

Tau-U analysis was conducted, indicating that behaviour change between phases was significant (Tau-U A vs B = 0.51, 90% CI [0.12-0.90], p = 0.03). Therefore, introduction of the reminder app had a significant positive effect on goal performance.

Participant 3

Participant 3 was allocated to a baseline period of 16 days. Participant 3 had datapoint 16 removed from analysis during intervention phase, due to being unwell and therefore unable to complete goal-related tasks. Participant 3 also recorded an additional datapoint at the end of the intervention phase, which was included in analysis.

Participant 3 showed an overall increase in goal-related tasks with introduction of the app, completing 73% (35/48) during baseline and 89% (59/66) during intervention. Evaluation of each phase suggested data was variable within the baseline and stable within the intervention phase. Evaluation of level change within-conditions indicated performance was improving during baseline and intervention phases. Split-middle method of trend estimation indicated there was an accelerating therapeutic trend during baseline and a zero-celerating trend during intervention, with data considered stable in both phases with application of the stability envelope. However, level change measures suggested a mixed picture, with absolute level change indicating a deteriorating performance, likely because of the initial drop at the start of the intervention phase. Whereas, mean and median level change both suggest a positive (improving) performance across phases. Tau-U analysis suggested that behaviour change between phases was not significant (Tau-U A vs B = 0.36, 90% CI [0.09-0.82], p = 0.19).

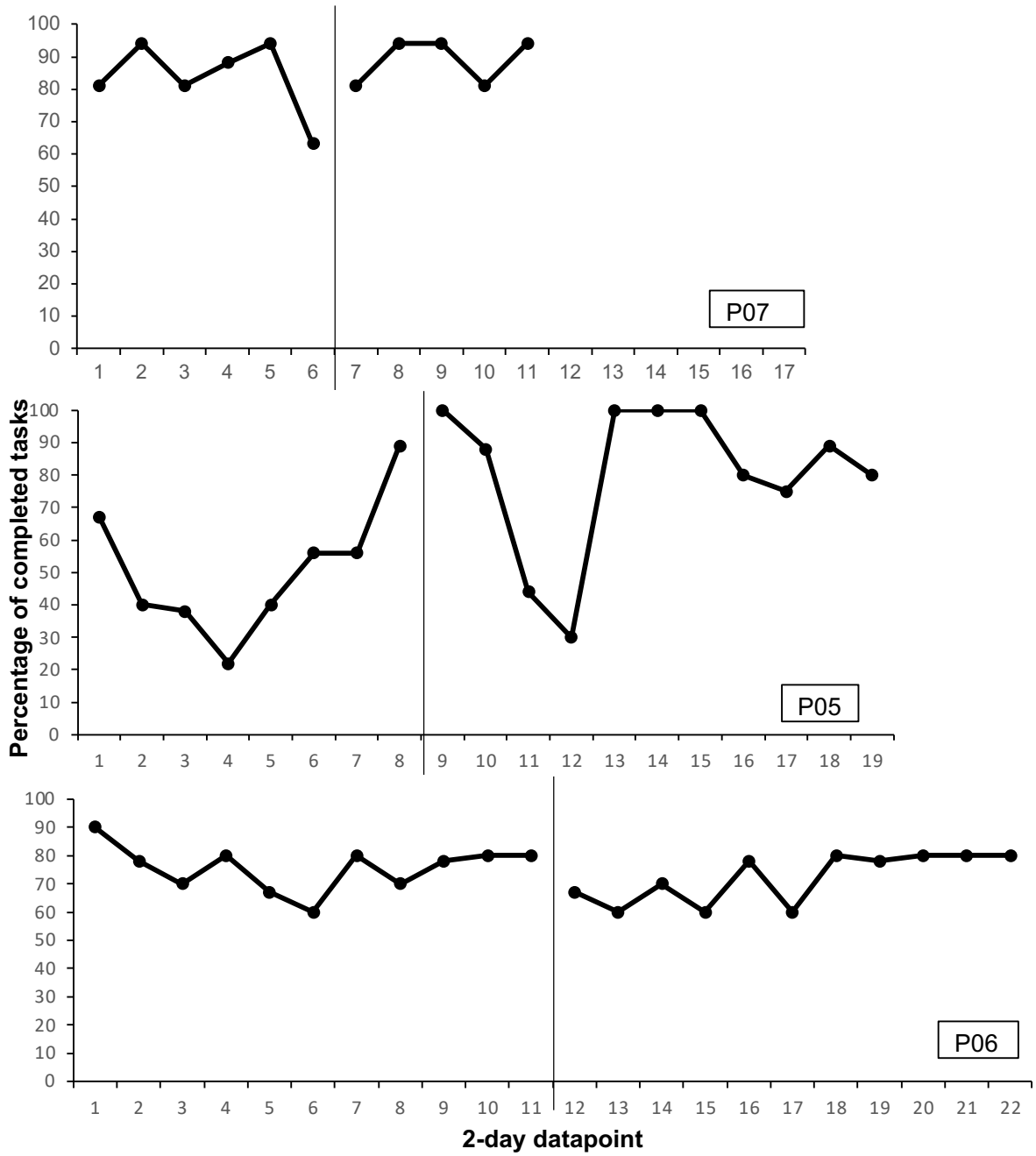


Figure 2.3 Participants' percentage of completed goal-related tasks per two-day period, presented in order of baseline length (12, 16, 22 days)

Participant 5

Participant 5 was randomly allocated to a baseline period of 16 days. Participant 5 showed an overall improvement in performance on goal related tasks from 51% (37/73) during baseline to 80% (81/101) during intervention, however data was

variable in both phases. Within-condition analysis of level change suggested an improving performance during baseline and a deteriorating performance during intervention. Split-middle method of trend estimation also indicated an accelerating therapeutic trend during baseline and decelerating contra-therapeutic trend during intervention. However, data was considered variable following application of the envelope of stability to trend lines. Whereas, between-condition of level change indicated a positive (improving) performance across conditions. Tau-U analysis indicated that behaviour change between phases was significant (Tau-U A vs B = 0.65, 90% CI [0.20-1], p = 0.02), indicating there was a significant positive improvement in goal-related tasks with introduction of the reminder app.

Participant 6

Participant 6 was allocated to a baseline period of 22 days. Participant 6's performance on goal-related tasks was relatively high during both baseline and intervention phases but displayed an overall decrease from 76% during baseline (81/107) to 72% (77/107) during intervention. Evaluation of each phase suggests data was stable across both phases. Within-condition analysis of level change indicated an improving performance within baseline and intervention phases. Split-middle method of trend analysis was utilised and indicated that there was a slightly accelerating therapeutic trend within the baseline phase and an accelerating therapeutic trend in the intervention phase, with data deemed stable following application of a stability envelope to trend lines. However the mean, median and relative level change measures suggested a deterioration between phases. Tau-U analysis suggested that behaviour change between phases was not significant (Tau-U A vs B = -0.2, CI [-0.62-0.2], p = 0.41, 90%).

Participant 7

Participant 7 was allocated to a baseline period of 10 days. However, he recorded an additional datapoint at the start of the baseline phase, which was included in analysis. Participant 7 discontinued data collection 10 days into the intervention phase, providing only 5 datapoints within phase B.

Participant 7 performance was also relatively high during both phases, displaying an overall increase from 83% (80/96) to 89% (71/80) with introduction of the app. Within-condition analysis showed that data was stable during baseline and intervention phases. Evaluation of level change within-conditions was unclear but indicated no change in performance during either phase. Split-middle method of trend estimation suggested that there was an accelerating therapeutic trend during baseline to a zero-celerating trend during intervention, with data considered stable following application of the envelope of stability to trend lines. Between-condition analysis suggested a change in trend across phases from accelerating to zero-celerating, however all level change measures suggest a positive (improving) change from baseline to intervention. Tau-U analysis suggested that behaviour change between phases was not significant (Tau-U A vs B = 0.27, p = 0.47, 90% CI [-0.33-0.87]).

Discussion

Analysis suggest that the introduction of the reminder app led to a significant improvement in goal-related tasks for two participants (P02, P05). Nearly all participants, except P06, displayed an improvement between baseline and interventions phases with the introduction of the reminder app, however this improvement was only significant for two participants.

Four participants had high levels of completed task (above 70%) within baselines which limited the ability to detect a significant result in the intervention phase. There are a few possible explanations for this. Firstly, a goal-setting session was provided prior to the baseline phase to ensure goals were person-centred and ensure daily tasks were within a SMART format (Bovend'Eerdt et al., 2009). To add to this most participants introduced at least one new activity/goal to their daily tasks at the start of the baseline phase, following the goal setting session. Therefore, it could be argued that this was a distinct intervention within itself that would have influenced participant motivation and goal attainment. This would be in line with Cabrera-Martos et al. (2018) who found that people with PD show greater improvement in goal attainment if treatment is focused on a specified tailored set of goals.

Secondly, in an attempt to resolve reporting limitations of previous research into the ApplTree reminder app, the monitoring form was developed to give the option for the nominated person to tick whether they had prompted the participant, should the task not be completed spontaneously (Wilson, 2021). To this end, a mixture of the goal-setting session and prompts from the nominated person may have assisted the participant to establish new routines within the baseline phase. This is supported by findings that both development of implementation intentions and repetition together were most helpful for increasing PM performance (Foster et al., 2017). However, within Foster et al's. (2017) RCT, this effect was largest for non-repeated event-related PM tasks compared to time-related PM tasks, such as taking medications at specific times. Research suggests that PwPD have more difficulty with time-related PM tasks (Raskin et al., 2011). Therefore, encoding strategies alone are likely not sufficient for improving time-related PM task performance, which is in line with findings of this study showing overall improvement with the use of a reminder app.

Thirdly, another possible contributing factor may be the interaction between participants, nominated person and the researcher during the baseline phase and the novelty of taking part in a study, which has been raised in previous reminder studies (Jamieson et al., 2019). Daily reminder texts were sent to the nominated person to complete the monitoring form and participants were aware of further meetings with the researcher to download the app ahead of the intervention phase, both of which may have impacted behaviour.

The study is the first of its kind to investigate the use of reminder apps for person-centred tailored treatment goals above and beyond medication reminders in PwPD and is the only study to monitor effectiveness through use of a nominated person (Lee et al., 2022). Using the nominated person rather than participants to record completion of daily tasks alleviates some of the difficulties found with self-report in PwPD, with Grosset et al. (2006) finding that self-report measures are insensitive at detecting sub-optimal adherence due to significant under-reporting of medication intake. However, daily monitoring by the nominated person is still classed as a subjective self-report measurement and is therefore more open to risk of bias than more objective measures. This a key limitation of the study. The monitoring form required the nominated person to note "yes or no" to the question of whether goals were completed and the same for whether they prompted the participant. Therefore, this is open in interpretation and value judgements. For example, a goal to take medication at 8am may be marked "yes" by the nominated person despite participant taking medication at 8.30am. It would therefore be helpful to ensure monitoring forms require more objective descriptions, such as the time medication was taken to reduce subjective judgements required. Future research using monitoring forms may benefit from further operationalising measures prior to baseline use, as well as assessing inter-rater reliability for nominated persons beforehand.

Despite measures taken to ensure informant report was suitable within each phase (e.g. no extended periods away from the participants), three participants reported periods where the nominated person was not with them for multiple days. As a result, these periods were less objective as they relied on participant self-report. Two participants anecdotally noted that it would be helpful if the app had a function to report when they had completed the task, as well as simply turning off the alarm. This may be a more effective way to monitor goal adherence than separate self-report or

informant report measures. Future studies should consider whether there are more objective ways of monitoring goal adherence, though this is clearly challenging in relation to everyday tasks.

A core strength of the research was the study design, with both the development and write up guided by RoBiNT recommendations for improving external and internal validity in SCED studies (Tate et al., 2013). In accordance with best practise for increasing interval validity, the study design allowed for three demonstrations of the intervention effect in each participant set; there was randomisation of the onset of phases and the assessor was independent of the researcher. Additionally, despite one participant discontinuing midway through the intervention phase, there were at least five datapoints within each phase for each participant. Additionally, a replication of the full design was completed with the second set of three participants, increasing the external validity and ability to interpret the data.

The key rationale behind the study hypothesis, that introduction of the reminder app would improve goal adherence, arose from research suggesting that PwPD display a specific cognitive profile with deficits in EF and episodic memory (Aarsland et al., 2017), resulting in difficulties with initiation and internally cued behaviour (Emre,2003). Related to this, PwPD have been shown to have increased prevalence of apathy and prospective memory deficits, both of which impact on everyday life and the management of their disease (den Brok et al., 2015; Pirogovsky et al., 2012). However, despite clinicians recruiting participants they felt would benefit from the use of the reminder app, baseline characteristics suggested no participants had global cognitive impairment or difficulties with PM, which may explain the non-significance of results. However, apathy was reported in half of participants and may explain their difficulties with adherence. To this end, the study provides tentative evidence that a reminder app may be beneficial for those with apathy. Despite no self-reported PM difficulties, results indicate an improvement with the introduction of the reminder app. Previous SCED research into a reminder app for those with dementia found significant benefit of this as a memory aid (McGoldrick, Crawford, & Evans, 2021). Additionally, previous research has found external cues helpful in those with dysexecutive syndrome in PD (Emre, 2003). Therefore, future research may consider using deficits in PM, EF and/or apathy specifically, whether measured by neuropsychological test or self-report, as inclusion criteria for participants.

Most participants expressed positive feedback and all participants except one continued to utilise a reminder app for treatment related goals past the end of the study. However, the study did not investigate the acceptability and usability of the ApplTree app for this population. Future research may wish to investigate the acceptability of reminder apps in PwPD and consider long-term follow-up to investigate whether use of these apps is maintained.

In conclusion, the aim of this study was to investigate whether the use of a recently developed reminder app improved adherence to personalised treatment goals in PwPD. Most, but not all, participants showed improvement from baseline to intervention phase, with two participants showing statistically significant change between phases. Although this suggests that the formulation of treatment goals increased the encoding of goals and that the associated external visual and auditory reminders increased initiation of goal-directed behaviours, the hypothesis that the introduction of a reminder app would significantly improve adherence to goal-related tasks could not be confirmed for all participants. Future research should consider investigating the impact of a reminder app in those more severe deficits, consider the most objective ways to record this data and consider the acceptability and long-term maintenance of app use in PD.

Reference List

Aarsland, D., Creese, B., Politis, M., Chaudhuri, K., Weintraub, D., & Ballard, C. (2017). Cognitive decline in Parkinson disease. *Nature Reviews Neurology*, *13*(4),217-231.

Armitage, L. C., Kassavou, A., & Sutton, S. (2020). Do mobile device apps designed to support medication adherence demonstrate efficacy? A systematic review of randomised controlled trials, with meta-analysis. *British Medical Journal Open; 10*(1), e032045.

Bovend'Eerdt, T. J., Botell, R. E., & Wade, D. T. (2009). Writing SMART rehabilitation goals and achieving goal attainment scaling: a practical guide. *Clinical Rehabilitation*, *23*(4), 352–361.

Butterfield, L.C., Cimino, C.R., Salazar, R., Sanchez-Ramos, J., Bowers, D., & Okun, M. S. (2017). The Parkinson's Active Living (PAL) Program. *Journal of Geriatric Psychiatry and Neurology, 30*(1), 11-25.

Cabrera-Martos, I., Ortiz-Rubio, A., Torres-Sánchez, I., Rodríguez-Torres, J., López, L., & Valenza, M. (2018). A randomized controlled study of whether setting specific goals improves the effectiveness of therapy in people with Parkinson's disease. *Clinical Rehabilitation., 33*(3), 465-472.

Costa, A., Caltagirone, C., & Carlesimo, G. (2018). Prospective memory functioning in individuals with Parkinson's disease: A systematic review. *The Clinical Neuropsychologist: Prospective Memory in Clinical Populations, 32*(5), 937-959.

Costa, A., Peppe, A., Brusa, L., Caltagirone, C., Gatto, I., & Carlesimo, G. A. (2008). Levodopa improves time-based prospective memory in Parkinson's disease. *Journal of the International Neuropsychological Society*, *14*(4), 601-610.

Crawford, J.R., Henry, J.D., Ward, A.L. and Blake, J. (2006), The Prospective and Retrospective Memory Questionnaire (PRMQ): Latent structure, normative data and discrepancy analysis for proxy-ratings. *British Journal of Clinical Psychology*, *45*: 83-104.

den Brok, M. G., van Dalen, J. W., van Gool, W. A., Moll van Charante, E. P., de Bie, R. M., & Richard, E. (2015). Apathy in Parkinson's disease: a systematic review and meta-analysis. *Movement Disorders*, *30*(6), 759-769.

Emre, M. (2003). Dementia associated with Parkinson's disease. *Lancet Neurology, 2*(4), 229-237.

Foster, E. R., Rose, N. S., McDaniel, M. A., & Rendell, P. G. (2013). Prospective memory in Parkinson disease during a virtual week: Effects of both prospective and retrospective demands. *Neuropsychology*, *27*(2), 170.

Foster, Erin R, McDaniel, Mark A, & Rendell, Peter G. (2017). Improving Prospective Memory in Persons With Parkinson Disease: A Randomized Controlled Trial. *Neurorehabilitation and Neural Repair, 31*(5), 451-461.

Grosset, D., Antonini, A., Canesi, M., Pezzoli, G., Lees, A., Shaw, K., ... & Grosset, K. (2009). Adherence to antiparkinson medication in a multicenter European study. *Movement disorders: official journal of the Movement Disorder Society*, *24*(6), 826-832.

Grosset, K. A., Bone, I., Reid, J. L., & Grosset, D. (2006). Measuring therapy adherence in Parkinson's disease: a comparison of methods. *Journal of Neurology, Neurosurgery & Psychiatry*, *77*(2), 249-251.

Grosset, K., Bone, I., & Grosset D. (2005). Suboptimal medication adherence in Parkinson's disease. *Movement Disorders. 20*(11):1502–7.

Jamieson, M., Cullen, B., Lennon, M., Brewster, S., & Evans, J. (2020). Designing ApplTree: Usable scheduling software for people with cognitive impairments. *Disability and Rehabilitation: Assistive Technology*, 1-11.

Jamieson, M., Monastra, M., Gillies, G., Manolov, R., Cullen, B., McGee-Lennon, M., ... & Evans, J. (2019). The use of a smartwatch as a prompting device for people with acquired brain injury: a single case experimental design study. *Neuropsychological rehabilitation*, *29*(4), 513-533.

Kralik, D., Koch, T., Price, K. & Howard, N. (2004) Chronic illness self-management: taking action to create order. *Journal of Clinical Nursing, 13,* 259–267.

Lakshminarayana, R., Wang, D., Burn, D., Chaudhuri, K. R., Galtrey, C., Guzman, N. V., et al. (2017). Using a smartphone-based self-management platform to support medication adherence and clinical consultation in Parkinson's disease. *NPJ Parkinson's Disease, 3*(1), 2-10.

Lee, J., Yeom, I., Chung, M. L., Kim, Y., Yoo, S., & Kim, E. (2022). Use of Mobile Apps for Self-care in People With Parkinson Disease: Systematic Review. *JMIR mHealth and uHealth*, *10*(1), e33944.

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American psychologist*, *57*(9), 705.

Malek, N. & Grosset, D. G. (2015). Medication adherence in patients with Parkinson's disease. *CNS Drugs, 29,* 47–53.

McGoldrick, C., Crawford, S., & Evans, J. J. (2021). MindMate: A single case experimental design study of a reminder system for people with dementia. *Neuropsychological rehabilitation*, *31*(1), 18–38.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., ... & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695-699.

NHS (2019) Parkinson's Disease. Retrieved from: https://www.nhs.uk/conditions/parkinsons-disease/

NHS Greater Glasgow and Clyde, Older People's Psychology Service (2019) Routine Outcome Measures (ROMs) in Older People's Mental Health: Report on a pilot introduction of ROMs in East and West Dunbartonshire [Unpublished review]

Pagni, C., Frosini, D., Ceravolo, R., Giunti, G., Unti, E., Poletti, M., ... & Tognoni, G. (2011). Event-based prospective memory in newly diagnosed, drug-naive Parkinson's disease patients. *Journal of the International Neuropsychological Society*, *17*(6), 1158-1162.

Parker, R. I., Vannest, K. J., Davis, J. L., & Sauber, S. B. (2011). Combining nonoverlap and trend for single-case research: Tau-U. *Behavior Therapy, 42*(2), 284–299.

Parkinsons UK (2019) Parkinson's Drugs. Retrieved from: https://www.parkinsons.org.uk/information-and-support/parkinsons-drugs

Peng, W., Kanthawala, S., Yuan, S. *et al.* A qualitative study of user perceptions of mobile health apps. *BMC Public Health* 16, 1158 (2016).

Peto, V., Jenkinson, C., & Fitzpatrick, R. (1998). PDQ-39: a review of the development, validation and application of a Parkinson's disease quality of life questionnaire and its associated measures. *Journal of neurology*, *245*(1), 10-14.

Pirogovsky, E., Woods, S. P., Vincent Filoteo, J., & Gilbert, P. E. (2012). Prospective memory deficits are associated with poorer everyday functioning in Parkinson's disease. *Journal of the International Neuropsychological Society, 18*, 986–995.

Raggi, A., Leonardi, M., Covelli, V., Albanese, A., Soliveri, P., Carella, F., & Romito, L. (2012). Concordance between severity of disease, prevalence of nonmotor symptoms, patient-reported quality of life and disability and use of medication in Parkinson's disease. *Neurological sciences*, *33*(4), 847-853.

Raskin, S. A., Woods, S. P., Poquette, A. J., McTaggart, A. B., Sethna, J., Williams, R. C., & Tröster, A. I. (2011). A differential deficit in time-versus event-based prospective memory in Parkinson's disease. *Neuropsychology*, *25*(2), 201.

Schapira, A. H. *et al.* (2009). Perspectives on recent advances in the understanding and treatment of Parkinson's disease. *European Journal of Neurology,16,* 1090–1099.

Smith, G., Del Sala, S., Logie, R. H., & Maylor, E. A. (2000). Prospective and retrospective memory in normal ageing and dementia: A questionnaire study. *Memory, 8*(5), 311–321.

Spitzer, R. L., Kroenke, K., Williams, J. B., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: the GAD-7. *Archives of internal medicine*, *166*(10), 1092-1097.

Spitzer, R., Kroenke, K. & Williams, J. (1999) Validation and utility of a self-report version of PRIME-MD: the PHQ Primary Care Study. JAMA, 282, 1737–44.

Starkstein, S. E., Mayberg, H. S., Preziosi, T. J., Andrezejewski, P., Leiguarda, R., Robinson, R. G. (1992). Reliability, validity, and clinical correlates of apathy in Parkinson's disease. *Journal of Neuropsychiatry and Clinical Neuroscience, 4,* 134-139.

Tate, R. L., Perdices, M., Rosenkoetter, U., McDonald, S., Togher, L., Shadish, W., ... & Vohra, S. (2016). The Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016: Explanation and elaboration. *Archives of Scientific Psychology*, *4*(1), 10.

Tate, R. L., Perdices, M., Rosenkoetter, U., Wakim, D., Godbee, K., Togher, L., & McDonald, S. (2013). Revision of a method quality rating scale for single-case experimental designs and n-of-1 trials: The 15-item risk of bias in N-of-1 trials (RoBiNT) scale. *Neuropsychological Rehabilitation, 23*(5), 619–638.

van der Eijk, M., Faber, M. J., Al Shamma, S., Munneke, M. & Bloem, B. R. (2011). Moving towards patient-centered healthcare for patients with Parkinson's disease. *Parkinsonism and Related Disorders, 17,* 360–364.

Wilson, J. (2021). *ApplTree: A single case experimental design study of a smartphone reminding application with community-dwelling adults who have sustained a stroke* (Doctoral dissertation, University of Glasgow)

Appendices

Appendix 1.1 Search Strategy

Ebscohost: PsychInfo (Search strategy for APA PsychArticles is the same)

1	DE "Parkinson's Disease"
2	TI Parkinson* OR AB Parkinson*
3	TI "Parkinson's disease" OR AB "Parkinson's disease"
4	TI "parkinsons disease" OR AB "parkinsons disease"
5	TI "Idiopathic Parkinson*" OR AB "Idiopathic Parkinson*"
6	1 OR 2 OR 3 OR 4 OR 5
7	DE "Executive Function" OR DE "Cognitive Control" OR DE "Set Shifting" OR DE "Task Switching" OR DE "Dysexecutive Syndrome" OR DE "Executive Functioning Measures"
8	TI "executive function*" OR AB "executive function*"
9	TI "executive dysfunction*" OR AB "executive dysfunction*"
10	TI dysexecutive OR AB dysexecutive
11	7 OR 8 OR 9 OR 10
12	DE "Neuropsychological Assessment" OR DE "Cognitive Assessment" OR DE "Test Standardization" OR DE "Test Administration" OR DE "Test Scores" OR DE "Testing"
13	DE "Self-Report" OR DE "Self-Perception" OR DE "Questionnaires" OR DE "Surveys" OR DE "Functional Status" OR DE "Life Experiences" OR DE "Daily Activities" OR DE "Activities of Daily Living"
14	TI "neuropsychological assessment" OR AB "neuropsychological assessment"
15	TI "neuropsychological test*" OR AB "neuropsychological test*"
16	TI test* OR AB test*

17	TI measure* OR AB measure*
18	TI "cognitive test*" OR AB "cognitive test*"
19	TI "cognitive assessment*" OR AB "cognitive assessment*"
20	TI scor* OR AB scor*
21	TI instrument OR AB instrument
22	TI measurement OR AB measurement
23	TI subjective OR AB subjective
24	TI ("self report" OR Self-report) OR AB ("self report" OR Self-report)
25	TI behav* OR AB behav*
26	TI objective OR AB objective
27	TI informant OR AB informant
28	TI everyday OR AB everyday
29	TI function* OR AB function*
30	TI daily OR AB daily
31	TI activit* OR AB activit*
32	TI experience* OR AB experience*
33	TI perception* OR AB perception*
34	TI impact OR AB impact
35	S12 OR S13 OR S14 OR S15 OR S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25 OR S26 OR S27 OR S28 OR S29 OR S30 OR S31 OR S32 OR S33 OR S34
36	S6 AND S11 AND S35

Appendix 1.2 Screening tool

Stage 1: Citation, title and abstract screening

Include: *Articles looking at executive function in Parkinson's Disease (PD).* Studies that report measures of executive function in people with PD, both subjective and objective. *If unclear, use inclusion checklist below*
Title contains: Parkinson's disease / Parkinson / Idiopathic Parkinsons / parkinsons / Parkinsonism AND Executive Function / Executive Functioning Measures / executive dysfunction AND Neuropsychological Assessment / Neuropsychological test / Cognitivre assessment / Cognitive test / Test Standardization / Test Administration / Test Scores / Testing / Test / Self-Report / Daily Activities / Activities of Daily Living / Life Experiences / Functional Status / Surveys / Questionnaires / scores / instrument / everyday functioning / activities / perception / Impact
Exclude: Title contains: Qualitative, Intervention, Systematic review / review / meta – analysis

Inclusion checklist
Y = Yes **N =** No **U =** Unclear **E =** Exclude**.**
If a box with E is ticked/marked stop screening.
Please add any notes onto Rayyan.

Criteria	Y	N	U	Notes
Article characteristics				
1. Is the study published in English?		E		
2. Is the study published in a peer-reviewed journal?		E		
3. Is the study one of the following? *Editorials, literature reviews, systematic reviews, meta-analyses, protocols, conference abstracts, posters, theses and dissertations, methodological and epidemiological studies and letters.*	E			
4. Is the study a quantative or mixed methods design?		E		

Participants				
5. Does it involve humans?		E		
6. Does it involve children (under 16 years of age)?	E			
7. Does it include participants with a diagnosis of Parkinson's Disease?		E		
Intervention/Exposure				
8. Does the study include an objective measure of executive function? *Neuropsychological test examples include:* *Trail Making Test (TMT);Verbal Fluency Test (VFT) - F, A and S; VFT Animals category; Clock Drawing Test (CDT); Digits Forward and Backward subtests (WAIS-R or WAIS-III); Stroop Test; Wisconsin Card Sorting Test (WCST)* ** If executive function is not mentioned explicitly or it is unclear but a measure of cognition is included, then include in full text screening*		E		
Comparator				
9. Does it include a subjective measure of executive function? *Examples include:* *Questionnaires, Behavior Rating Inventory of Executive Function-Adult (BRIEF-A), BADS-DEX, subjective complaints, clinical interview* ** If executive function is not mentioned explicitly or it is unclear but a measure of subjective/self reported cognition is included, then include in full text screening*		E		
Additional informantion: Does it include a comparator population? *Healthy controls, people with Parkinson's disease without executive dysfunction, informant measures* **Please add a label if no comparator population.*				
Outcome Characteristics				
10. Does it include analysis (correlation/ association/ regression) looking at the association between objective assessments and subjective reports of executive function?		E		

Decision	**Action**
Include: if above criteria is either met (there is no E ticked) or unclear	Mark as included

Exclude: if an E is ticked	Mark as excluded *Please provide the reason for exclusion on rayyan

Appendix 1.3 AXIS Quality appraisal

AXIS Questions	Copeland et al. (2016)	Koerts et al. (2011a)	Koerts et al. (2012)	Kudlicka et al . (2013)	Lanni et al. (2014)	Lovstad et al. (2016)	Mills et al. (2020)	Puente et al (2016	Siquier & Andres (2021)	Sobreira et al. (2008)	Vlagsma et al. (2017)
Introduction											
1) Were the aims/objectives of the study clear?	Y	Y	Y	Y	Y	Y	Y	y	Y	Y	Y
Methods											
2) Was the study design appropriate for the stated aim(s)?	Y	Y	y	Y	Y	Y	Y	y	Y	Y	Y
3)Was the sample size justified?	N	N	N	N	N	N	N	N	N	N	N
4) Was the target/reference population clearly defined? (Is it clear who the research was about?)	Y	Y	y	Y	Y	Y	Y	Y	Y	Y	Y

5) Was the sample frame taken from an appropriate population base so that it closely represented the target/reference population under investigation?	Y	Y	y	Y	Y	N	Y	Y	Y	DK	Y
6) Was the selection process likely to select subjects/participants that were representative of the target/reference population under investigation?	Y	Y	y	Y	Y	N	Y	Y	Y	Y	Y
7) Were measures undertaken to address and categorise non-responders?	DK	DK	DK	Y	DK	N	DK	Y	DK	DK	DK
8) Were the risk factor and outcome variables measured appropriate to the aims of the study?	Y	Y	y	Y	Y	Y	Y	Y	Y	Y	Y

9) Were the risk factor and outcome variables measured correctly using instruments/measurements that had been trialled, piloted or published previously?	N	Y	y	Y	Y	Y	Y	Y	Y	DK	Y
10) Is it clear what was used to determined statistical significance and/or precision estimates? (e.g. p-values, confidence intervals)	Y	Y	y	Y	Y	Y	Y	Y	Y	Y	Y
11) Were the methods (including statistical methods) sufficiently described to enable them to be repeated?	Y	Y	y	Y	Y	Y	Y	Y	Y	Y	Y
Results											
12) Were the basic data adequately described?	Y	Y	y	Y	Y	Y	Y	Y	Y	N	Y
13) Does the response rate raise concerns about	DK	DK	DK	DK	DK	N	DK	N	DK	DK	DK

non-response bias? (reverse)*											
14) If appropriate, was information about non-responders described?	N/A	N/A	N/A	N/A	N/A	N	NA	Y	NA	NA	NA
15) Were the results internally consistent?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
16) Were the results presented for all the analyses described in the methods?	Y	N	Y	Y	Y	N	Y	Y	Y	N	Y
Discussion											
17) Were the authors' discussions and conclusions justified by the results?	Y	Y	y	Y	Y	Y	Y	Y	Y	DK	Y
18) Were the limitations of the study discussed?	Y	Y	y	Y	N	Y	Y	Y	Y	Y	Y
Other											
19) Were there any funding sources or conflicts of interest that may affect the authors’	N	DK	DK	N	N	N	N	N	N	DK	N

interpretation of the results? (reverse)											
20) Was ethical approval or consent of participants attained?	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y

AXIS Domains	Copeland et al. (2016)	Koerts et al. (2011 a)	Koerts et al. (2012)	Kudlicka et al . (2013)	Lanni et al. (2014)	Lovstad et al. (2016)	Mills et al. (2020)	Puente et al (2016)	Siquier & Andres (2021)	Sobreira et al. (2008)	Vlagsma et al. (2017)
Quality of reporting	L 7/7	L 6/7	L 7/7	L 7/7	L 6/7	L 6/7	L 7/7	L 7/7	L 7/7	H 5/7	L 7/7
Study design quality	L 6/7	U 5/7	U 5/7	L 6/7	L 6/7	H 5/7	L 6/7	L 6/7	L 6/7	U 3/7	L 6/7
Introduction of bias	H 2/5	U 3/5	U 3/5	L 4/5	U 3/5	H 3/5	U 3/5	L 7/7	U 3/5	U 2/5	U 3/5

Appendix 2.1 Participant Information sheet

https://osf.io/b2ny9/files/osfstorage/62e287e6c79a4c67679e6485

Appendix 2.2 Nominated Person Information sheet

https://osf.io/b2ny9/files/osfstorage/62e289c2588bb96c01b875e9

Appendix 2.3 Consent for contact form

https://osf.io/b2ny9/files/osfstorage/62e28aec588bb96c05b87e13

Appendix 2.7 Non-substantial amendment R&I approval

Sunday, June 19, 2022 at 16:27:47 British Summer Time

Subject: R&I Ref GN21NE440 Protocol V12 15/03/2022 NSA01 18/03/2022 – Cat A
Date: Friday, 18 March 2022 at 10:12:13 Greenwich Mean Time
From: Graham, Brittany
To: Jon Evans, Georgina Rayment (PGR)
CC: Colette Montgomery Sardar, Ross, Barbara, Surtees, Pamela

Dear Professor Evans and Georgie,

R&I Ref: GN21NE440 **Ethics Ref**: 21/WS/0154
Investigator and site(s): Professor Jonathan Evans (CI), Ms Georgina Rayment (PI, QEUH)
Project Title: A single case experimental design study of a reminder app for improving adherence to personalised treatment goals in Parkinson's Disease
Protocol Number: V12 15/03/2022
Amendment: Non-substantial Amendment 01 – Cat A
Sponsor: NHS GG&C

I am pleased to inform you that R&I have reviewed the above study's Amendment 01 (18/03/2022 – Cat A) and can confirm that Management Approval is still valid for this study.

Brief summary of documents reviewed by sponsor for submission:	Version	Dated
Protocol	V12	15/03/2022
Study Recruitment Information	V5	15/03/2022
Study Invitation Letter	V4	15/03/2022
Screening Tool	V3	15/03/2022
Information Sheet	V8	15/03/2022
OPPS HoD Email Agreement	N/A	14/03/2022

I wish you every success with this research project.

Yours sincerely,

Brittany

Brittany Graham
Senior Research Administrator
NHS Greater Glasgow & Clyde
Research & Innovation Department (R&I)
Ward 11 | Dykebar Hospital | Grahamston Road | Paisley | PA2 7DE

✉: Brittany.Graham@ggc.scot.nhs.uk
💻: Who Provides R&I Approval | Contact Us
💻: NHSGGC : Research & Development

I am currently working remotely and I am contactable by e-mail. My working hours are 8-4 Mon-Fri.

Appendix 2.8 Consent form

https://osf.io/b2ny9/files/osfstorage/62e28c8fc79a4c67779e5a22

Appendix 2.9 Nominated Person Consent form

https://osf.io/b2ny9/files/osfstorage/62e28ce21bb7a577d61f375a

Appendix 2.10 SMART goal worksheet

SMART Goal Worksheet

Today's Date: _____

Thinking about goals

Is there any aspect of your Parkinson's disease treatment that you're struggling to manage?

_____

_____

What, specifically, would you like to be different?

_____

_____

Do you ever forget to complete tasks that affect your Parkinson's disease?

_____

_____

Are there actions or tasks you should be completing for managing your Parkinson's disease that you struggle to motivate yourself to do?

_____

_____

Have you identified any areas/ goals that you would like to work on? Have you thought of events and tasks where the reminder app might help?

_____

_____

_____

Goal 1:

_____

_____

_____

Verify that your goal is SMART (Specific, Measurable, Achievable, Relevant, Timely)

Specific: What exactly do you want to accomplish? What are you going to do? How are you going to do it? Where are you going to do it? When are you going to do it? With whom are you going to do it?

_____

_____

_____

Measurable: How will you know when you have reached this goal?

_____

_____

_____

Achievable: Is achieving this goal realistic with effort and commitment? Have you got the resources to achieve this goal? If not, how will you get them?

_____

_____

_____

Relevant: Why is this goal significant to your life? How will it help with the management of Parkinson's disease?

_____

_____

_____

Timely: When will you achieve this goal? How many times will you complete it in a week?

_____

_____

_____

_____

This goal is important because:

_____

_____

The benefits of achieving this goal might be:

_____

_____

Potential Obstacles

_____

_____

_____

Potential Solutions

_____

_____

_____

_____

Goal 2:

_____

_____

_____

Verify that your goal is SMART (Specific, Measurable, Achievable, Relevant, Timely)

Specific: What exactly do you want to accomplish? What are you going to do? How are you going to do it? Where are you going to do it? When are you going to do it? With whom are you going to do it?

_____

_____

_____

Measurable: How will you know when you have reached this goal?

_____

_____

_____

Achievable: Is achieving this goal realistic with effort and commitment? Have you got the resources to achieve this goal? If not, how will you get them?

_____

_____

_____

Relevant: Why is this goal significant to your life? How will it help with the management of Parkinson's disease?

_____

_____

_____

Timely: When will you achieve this goal? How many times will you complete it in a week?

_____

_____

_____

_____

This goal is important because:

_____

_____

The benefits of achieving this goal might be:

_____

_____

Potential Obstacles

_____

_____

_____

Potential Solutions

_____

_____

_____

_____

Goal 3:

_____

_____

_____

Verify that your goal is SMART (Specific, Measurable, Achievable, Relevant, Timely)

Specific: What exactly do you want to accomplish? What are you going to do? How are you going to do it? Where are you going to do it? When are you going to do it? With whom are you going to do it?

_____

_____

_____

Measurable: How will you know when you have reached this goal?

_____

_____

_____

Achievable: Is achieving this goal realistic with effort and commitment? Have you got the resources to achieve this goal? If not, how will you get them?

_____

_____

_____

Relevant: Why is this goal significant to your life? How will it help with the management of Parkinson's disease?

_____

_____

_____

Timely: When will you achieve this goal? How many times will you complete it in a week?

_____

_____

_____

_____

This goal is important because:

_____

_____

The benefits of achieving this goal might be:

_____

_____

Potential Obstacles

_____

_____

_____

Potential Solutions

_____

_____

_____

_____

Appendix 2.11 Participant goal information

Participant	Goal 1	Goal 2	Goal 3
P01	Take medication at specified time – 4 times daily	Exercise twice a day at specific times (walking or exercise bike)	
P02	Exercise Morning (Physio stretches or walk)	Exercise Afternoon (Physio stretches or walk)	
P03	Take medication on time each morning	Exercise twice a day at specified times (walk, physio exercise, golf)	
P04	Take medication on time each morning	Exercise & Relaxation on specific days (walk, yoga or meditation)	Start sleep wind-down routine at specified time
P05	Take medication at specified time – 4 times daily	Exercise once a day (circuit training)	
P06	Take medication at specified time – 4 times daily	Exercise on specific days (Walk, dance or yoga)	
P07	Take medication at specified time – 4 times daily	Exercise once a day at specified time (Gym or walk)	Hydration – Drink water 3 times a day between medications – specified times

Appendix 2.12 Weekly monitoring form

Weekly Monitoring Form

Week Beginning Monday: / /2021

Please enter week commencing date above.

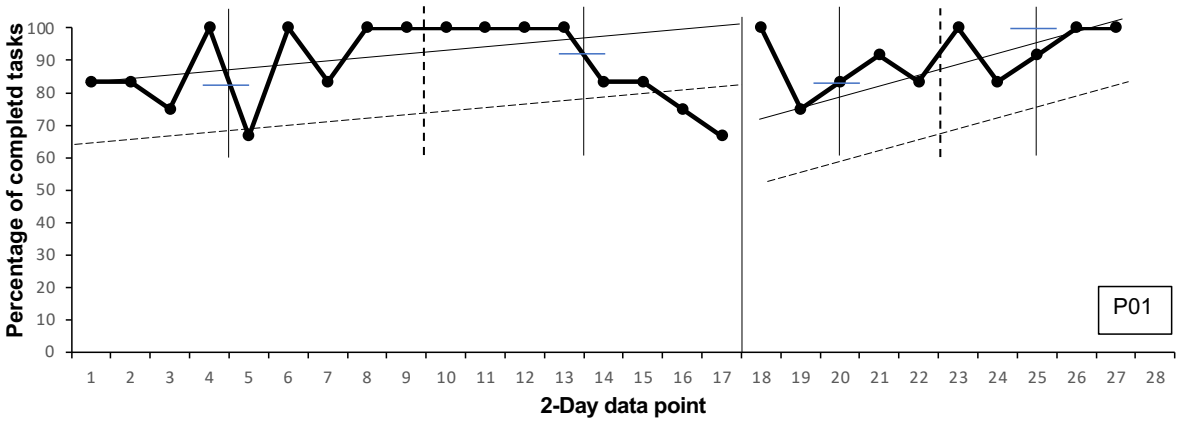
Please complete the form daily between DATE and DATE

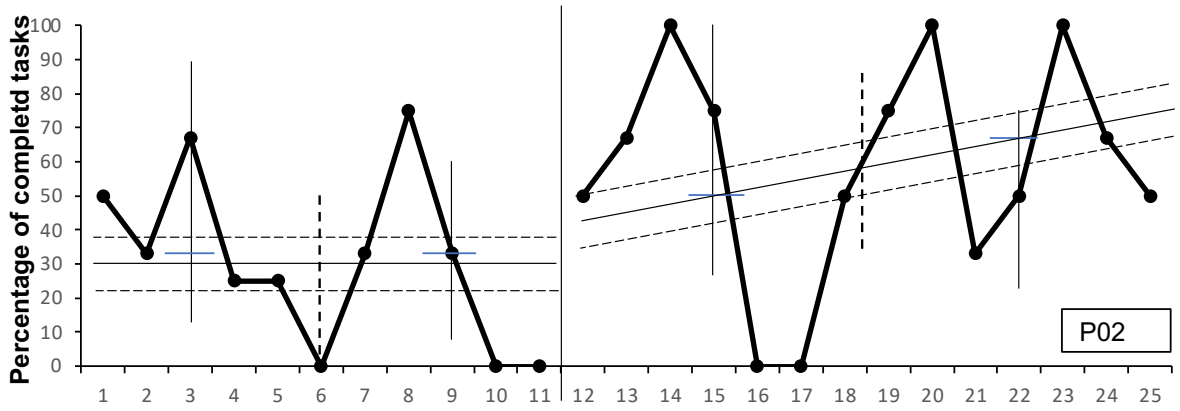
Please keep this form somewhere private where it cannot be used as a memory prompt for the participant.

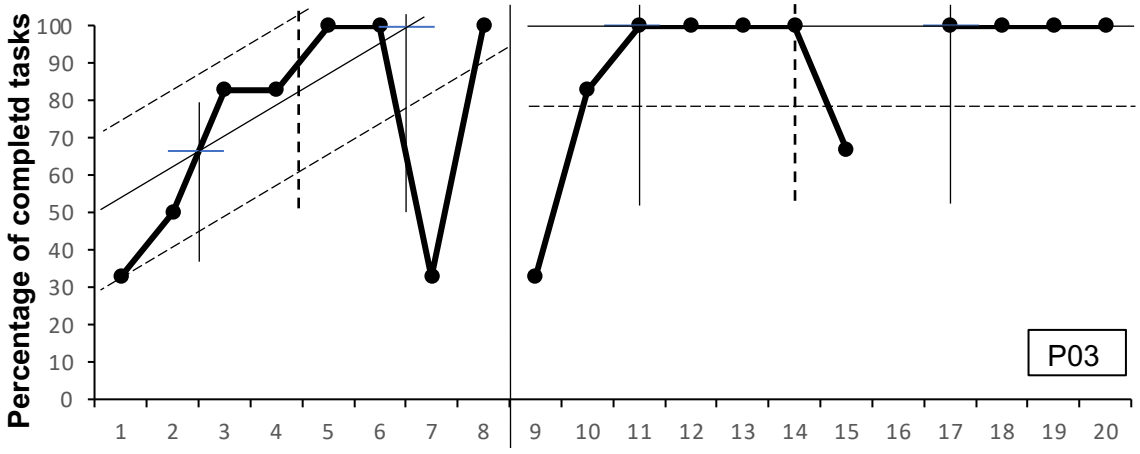
	Activity	**Was task completed Yes/No?**	**With a prompt? Yes/No**
Monday			
Tuesday			
Wednesday			

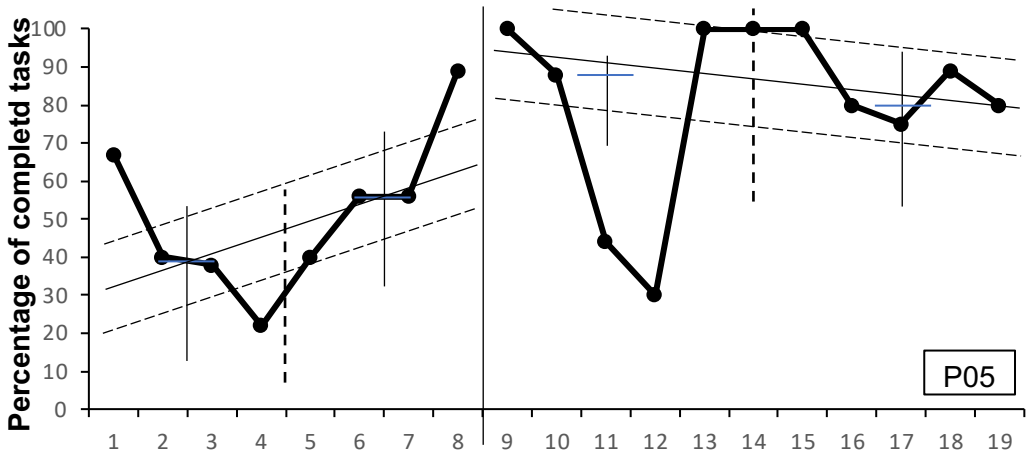
	Activity	**Was it completed Yes/No?**	**With a prompt? Yes/No**
Thursday			
Friday			
Saturday			
Sunday			

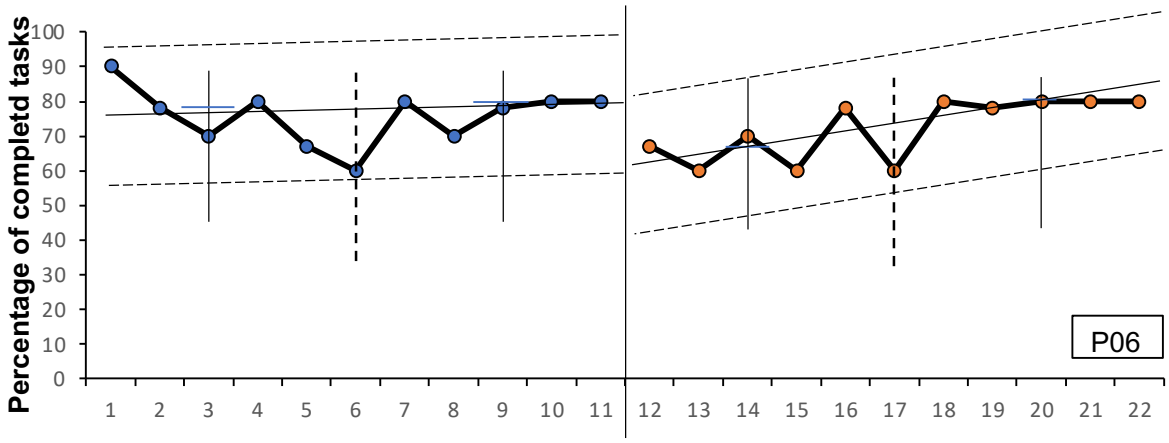
Appendix 2.13 Visual analysis of participants – Stability envelopes applied to data

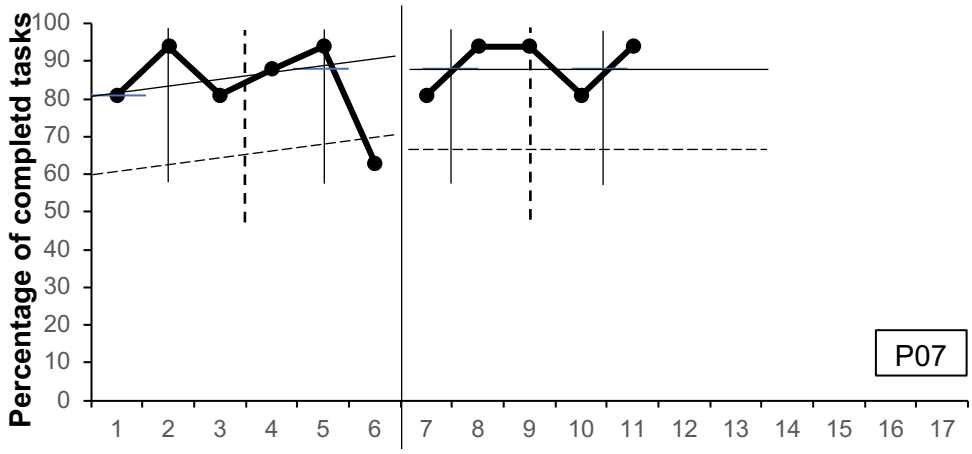







Percentage of completd tasks
100
90
80
70
60
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19
P05


Percentage of completd tasks
100
90
80
70
60
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22
P06


Percentage of completd tasks
100
90
80
70
60
50
40
30
20
10
0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
P07

Appendix 2.14 MRP Proposal

https://osf.io/b2ny9/files/osfstorage/62e2868227b74636f00ace05