UNIVERSITY OF
EXETER

The Role of a Working Memory Training Program in Reducing Repetitive Negative Thinking in Older Adults.

Submitted by Jodie Rawlings to the University of Exeter,

as a thesis for the degree of Doctor of Clinical Psychology, April 2017

I certify that all material in this thesis which is not my own work has been identified and that no material has previously been submitted and approved for the award of a degree by this or any other University.

Signature: …………………………………………………………..

For Mina.

I love you sweetheart,

Mum x

Table of Contents

List of Tables

List of Figures

SCHOOL OF PSYCHOLOGY

DOCTORATE IN CLINICAL PSYCHOLOGY

LITERATURE REVIEW

What Evidence Exists of a Role for Working Memory in Depression and Anxiety within the General Older Adult Population?: A Systematic Review

Trainee Name:	**Jodie Rawlings**
Primary Research Supervisor:	**Dr Jenny Limond**
	Senior Lecturer and Co-Director of Research
Secondary Research Supervisor:	**Dr Philip Yates**
	Research Tutor
Target Journal:	Clinical Gerontologist
Word Count:	*(5869-1998 in tables, figures etc.)* = 3871

Submitted in partial fulfilment of requirements for the Doctorate Degree in Clinical Psychology, University of Exeter

Abstract

Background: There is increasing evidence within general adult populations of a link between working memory (WM), and depression and anxiety. At the same time, research indicates a decline in WM with aging.

Objectives: This review summarises and synthesises the literature investigating working memory and its relationship to depression and anxiety within the older adult population.

Method: Systematic review of all peer-reviewed literature to date using AgeLine, Ovid, and PubMed databases alongside searches of the grey literature. The research is synthesised using narrative discussion.

Results: Sixteen relevant papers were included, consisting of cross-sectional and prospective studies.

Conclusions: The included studies broadly supported a relationship between poor WM and depression and anxiety. However, there appears to be no difference in WM impairment between individuals with early- and late-onset depression and no amelioration in WM impairment with effective treatment of depression. The direction of relationship between mood and WM impairment is hard to ascertain, given the paucity of experimental designs within the field.

Clinical Implications: Older adults presenting to services with depression are likely to have some degree of deficit in WM.

Keywords: depression, anxiety, older adult, working memory, cognition

Introduction

Rationale

   Working memory. Working memory (WM) has been defined as "actively maintaining… or manipulating information across a short delay" (Snyder, 2013, p.82). Baddeley and Hitch's (1974) model of WM comprises a central executive, which controls and regulates information within two slave systems: the visuospatial sketchpad, responsible for the storage of visual data, and phonological loop, which stores auditory information. Baddeley's (2000) revision of the model (Figure 1) results in the inclusion of a third slave system, the episodic buffer. This system is proposed to represent an intermediary stage between short- and long-term memory, in which information is integrated across time and space.
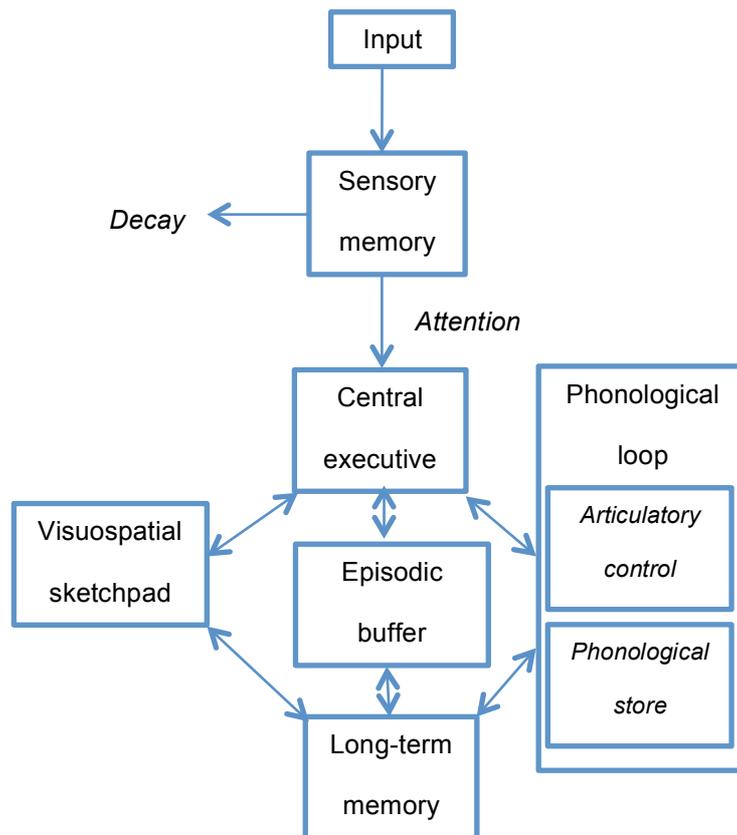
Figure 1: Baddeley 's (2000) model of WM

Working memory in older adults. There is clear evidence of a decline in WM with the aging process (Braver & West, 2007). In a discussion on the history of the field, Salthouse (1994) concludes that a slowed speed of processing is sufficient to explain this decline, such that more of the WM capacity is given over to the processing of information, decreasing capacity for storage. This is supported by findings which indicate that performance on complex span tasks - those, such as reading span, which require the manipulation or processing of information held online - is particularly impaired in older adults (OAs), while basic maintenance tasks show little age related decline (see Dobbs & Rule, 1989; Bopp & Vergaehen, 2005). Engle (2002) has highlighted the clear relationship between the 'central executive' within

Baddeley's model, and other functions of executive control[1], suggesting that

decline in the processing functions of WM may be reflective of a primary decline

in executive function. Daselaar and Cabeza (2013) suggest that decline in the

structural integrity of medial temporal lobe and pre-frontal cortex may be the

cause of WM decline in OAs.

Working memory and mood.

Depression. Research has consistently demonstrated a relationship

between depression[2] and WM. Studies have predominantly used a case-

comparison approach, in which individuals with and without depression are

compared on WM tasks. These reports have demonstrated poorer WM

performance in individuals with depression (Rose & Ebmeier, 2006). A

thorough meta-analysis from Snyder (2013) demonstrated impairments for

depressed patients on both visuospatial and verbal working memory tasks, with

reported composite effect sizes for both at d = .45. While some studies have

demonstrated impairment across both the phonological loop and visuospatial

sketchpad (Christopher & MacDonald, 2005), others have failed to support this

finding (Channon, Baker, & Robertson, 1993). Channon and colleagues instead

found impairment on tasks tapping in to central executive functions. These

findings have since been replicated (Harvey et al., 2004), suggesting that it may

be aspects of updating and executive facets of WM which are most affected in

depression.

Harvey and colleagues (2004) reported negative correlations between n-

back task performance - a measure of WM - and both the number of

---

[1] Executive function is an umbrella term for a wide range of cognitive control functions, including

[2] Depression can be defined as a condition characterised by "depressed mood, loss of interest and enjoyment, and reduced energy" (p.100, World Health Organisation, 1992).

hospitalisations and longitudinal course of depression. Gohier et al. (2009) also demonstrated a correlation between depression severity and WM, such that increased depression related to poorer WM. The relationship between low mood and working memory is maintained even at non-clinical levels, with Owens, Koster and Derakshan (2012) finding low WM capacity in dysphoric compared to non-dysphoric student participants.

Anxiety. The relationship between anxiety and WM has been less widely studied. In Christopher and MacDonald's (2005) study on WM, individuals with a diagnosis of generalised anxiety disorder (GAD) formed a control group for the individuals with depression. The authors found that while the anxious participants had performance comparable to healthy controls on tasks requiring the visuospatial sketchpad and phonological loop, their performance was in line with depressed participants on WM tasks making use of the central executive (i.e. updating tasks). This suggests that there may also be an executive level WM impairment in anxiety as well as depression. Walters (2016) investigated updating tasks of WM in individuals with anxiety, finding a clear trend within the small sample of slower WM updating in individuals with high levels of anxiety compared to low anxiety peers. Further research is warranted.

Working memory and mood in older adults. Given the evidence of both a decline in WM with age, and a likely role for WM in both anxiety and depression, it is sensible to consider to what degree WM may be relevant to the mood disorders of OAs (Mental Health Foundation, 2015). The prediction can be made that the relationship between WM and mood seen in adults of working age would be replicated in older adults. Von Hippel and colleagues (2008)

compared OA with late (>60yrs at first episode) and early (<60yrs) onset

experiences of depression. OAs with late-onset depression were more impaired

on tasks of executive function (EF) – including a simple task of WM – than their

early-onset peers, leading the authors to the suggestion that late-onset

depression is more closely linked to the neurological decline that causes WM

impairment than early-onset depression. WM decline associated with

depression in OA was linked to the reorganization of relevant brain networks,

including frontal and parietal areas (Dumas & Newhouse, 2015).

With regard to anxiety, Butters and colleagues (2011) compared OA with

generalised anxiety disorder (GAD) to healthy OA controls using the digit span

task as part of a neuropsychological battery. They found poorer WM in the GAD

group. Following psychopharmacological intervention (escitalopram),

participants with GAD demonstrated small but significant improvements in WM

alongside clinical improvements in anxiety symptoms. The authors suggest that

effective intervention for anxiety in OA might therefore take into account EF as

a treatment target, as improvements in WM may be causal in improving anxiety

symptoms.

Objectives

The objective of this systematic review (SR) is to bring together the

burgeoning literature exploring the relationship between WM and mood in OAs.

The research question is:

> What evidence exists of a role for working memory in depression and anxiety in the general older adult population?

Given this is a developing field, any evidence of the relationship, regardless of direction, was considered to be relevant.

In addressing the question, the following PICO criteria were used:

Table 1

PICO Systematic Review Inclusion Criteria

Participants	Adults over the age of 60 years (without a diagnosis of dementia)
Intervention	Changes in working memory
Comparison	Various comparisons possible to investigate question e.g.: • Diagnosis of clinical depression/anxiety • High vs. low scores on measures of mood • High vs. low scores on measures of working memory • Age of depression/anxiety onset
Outcomes	Repetitive negative thought (both rumination and/or worry), depression or anxiety

These criteria were operationalised to develop the search terms:

Table 2

Operationalisation of Search Terms

Terms	Operationalisation
Relating to older age	Study populations must be compromised exclusively of adults over the age of 60 years. Study population should be comprised of individuals who have not been selected for inclusion on the basis of a pre-existing condition other than those relevant to the study i.e. dementia, heart disease, diabetes.
Relating to executive functioning	Study must measure working memory using a valid measure e.g. complex span tasks, n-back tasks, Wechsler memory scale, updating tasks, binding tasks. Measures will be deemed appropriate where there exists a separate peer reviewed paper describing the creation of the measure itself, or the article presents a rational justification for its use. Studies will be included even if the paper does not refer to the test as a measure of WM, assuming it is evidenced as such according to the criteria above.
Relating to mental health	Study must include a standardised measure of mood or repetitive negative thinking, including for example quantitative measures (e.g. BDI, BAI, HADS, PHQ9, GAD-7, RRS, etc.) or standard diagnostic interviews (e.g. SCID).

Note: BDI – Beck Depression Inventory; BAI – Beck Anxiety Inventory; HADS – Hospital Anxiety and Depression Scale; PHQ-9 – Patient Health Questionnaire-9; GAD-7 – Generalised Anxiety Disorder Scale; RRS – Rumination Response Scale; SCID – Structured Clinical Interview for DSM-5

Methods of Review

Protocol and Registration

The review was undertaken according to the PRISMA guidelines (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009). These guidelines ensure consistent and accurate reporting of SR data.

Eligibility Criteria

Studies were eligible for inclusion if they were peer-reviewed journal articles of quantitative data, published or accessible online any time up to 21st October 2016. Articles were included if they contained data exploring the relationship between WM and depression or anxiety in an older adult (>60 years) population.

Inclusion criteria:

1) Peer reviewed journal article
2) Quantitative study design and analysis in which the relationship between working memory and mood is *directly* explored. This includes for example classic experimental studies, RCTs, single case experimental design
3) Includes a well-validated measure of working memory (i.e. there exists a separate peer reviewed paper describing the creation of the measure itself, or the article presents a rational justification for its use) e.g. *n*-back tasks, complex span tasks, updating tasks.
4) Includes either/both of:
   a. Appropriate standardised measure of depression/anxiety or repetitive negative thinking e.g. BDI, BAI, HADS, PHQ9, GAD7, RSS
   b. Clinical population of individuals with a diagnosis of depression/anxiety, as ascertained by full diagnostic interview.
5) Sample of individuals above the age of 60 years

Exclusion criteria:

1) Selection to study on basis of diagnosis other than depression/anxiety/WM impairment e.g. coronary heart disease, diabetes
2) Includes participants with diagnosis of dementia
3) Includes any participants under the age of 60 years
4) Article not available in English
5) Non-empirical studies e.g. theoretical or review articles
6) Qualitative study design
7) Measure of working memory or mood is exclusively obtained using imaging methods

Figure 2: Eligibility criteria used for the systematic review.

Note: BDI – Beck Depression Inventory; BAI – Beck Anxiety Inventory; HADS – Hospital Anxiety and Depression Scale; PHQ-9 – Patient Health Questionnaire-9; GAD-7 – Generalised Anxiety Disorder Scale; RRS – Rumination Response Scale

Information Sources

Studies were identified using the online electronic databases PubMed,

Ovid[3] and AgeLine. These databases were searched from the beginning of their

records up to 21st October 2016. The reference lists of key articles were also

searched for further relevant literature. Key authors in the literature (e.g. Nebes)

were also contacted to obtain additional relevant literature.

Search

All databases were searched using the following string, with inbuilt

Boolean logic. Given the emphasis upon dementias within OA literature, the

decision was taken to specifically exclude papers referencing dementia within

the title/abstract.

---

older adult OR 60 years OR old age OR elderly OR geriatric OR pensioner OR third age OR later life

executive function* OR executive process* OR central executive OR dysexecutive OR working memory OR attentional control OR cognitive control OR executive control

rumination OR worry OR anx* OR depress* OR low mood OR repetitive negative th* OR GAD OR generalized anxiety

NOT dementia

The search string was set to apply to titles and abstracts only.

Study Selection

The review author (JR) screened titles/abstracts of identified articles, with

articles that did not meet eligibility criteria excluded. Remaining articles were

---

[3] Search using the Ovid database also included PsycARTICLES, Embase, Global Health, Ovid MEDLINE and PsycINFO.

read in full, and were screened against eligibility criteria to ascertain inclusion to the study.

Data Collection Process and Data Items

Data was assessed and extracted from articles by the study author only. Relevant information included 1) participants, 2) study design, 3) measure of WM used, 4) method of assessing mood used, and 5) relevant findings (Table 3).

Risk of Bias in Individual Studies

All articles were assessed using the Quality Assessment Tool for Quantitative Studies (QATQS; Effective Public Health Practice Project, 2009; Appendix A). This tool allows for the assessment of quality across a range of areas, including selection bias, study design, and data collection method. The QATQS has excellent inter-rater reliability and sound psychometric properties (Armijo-Olivo, Stiles, Hagen, Biondo, & Cummings, 2012).

The information from this assessment was used to guide narrative synthesis of the findings, such that more robust studies were afforded greater emphasis.

Summary Measures and Synthesis of Results

Quantitative summary methods were not used within this SR. This is due to the complexity of the bi-directional relationship between WM and mood. As a result, the studies comprised a wide variety of primarily correlational designs, which could not usefully be synthesised. Results of the SR were therefore narratively synthesized, in line with the ESRC narrative synthesis guidance (Popay et al., 2006).

Risk of Bias across Studies

The possible role of risk of bias across studies (e.g. publication bias) was noted and reflected within the narrative synthesis.

Results

Study Selection

Across the three databases, 999 articles were identified. One hundred and eighty two of these were applicable for full text screening, of which 166 were not eligible for inclusion. The most common reasons for exclusion being that the relationship between WM and mood was not directly explored, and that measures of cognition did not appropriately identify WM as a distinct construct. Sixteen articles were identified which met full study criteria and which therefore form the data for the SR.
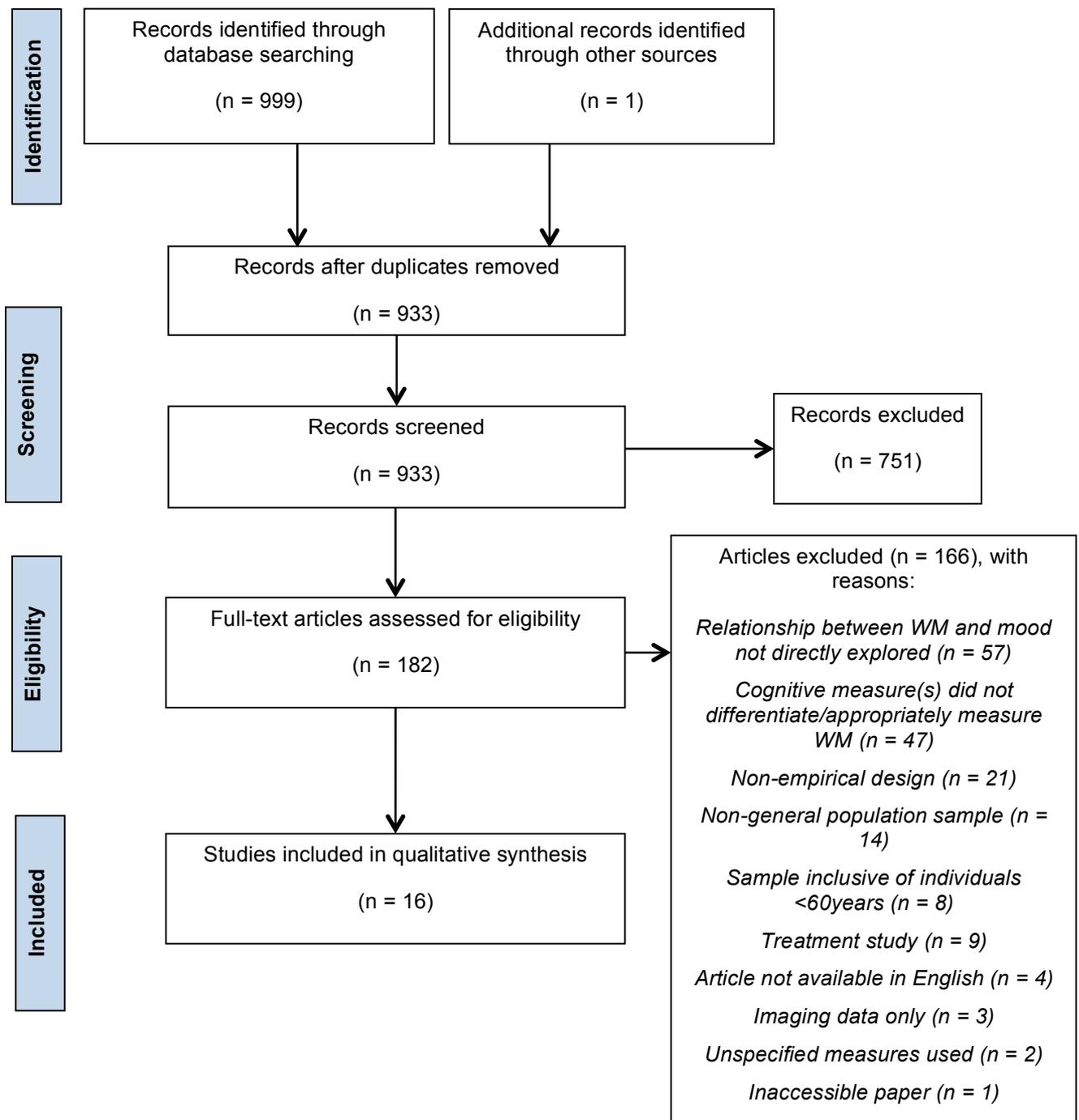
Figure 3: PRISMA flow diagram summarising study selection process

Study Characteristics

None of the 16 included studies were published before the year 2000.

Twelve of the studies employed a case-control design, with comparisons made

between: individuals with a diagnosis of depression and healthy controls

(studies 1, 4, 5, 6, 9, 14 and 16[4]); current depression compared with past

depressed and never-depressed (study 2); current diagnosable depression

compared with depressive symptoms and healthy controls (studies 3 and 7);

early- and late-onset depression (study 8); early-onset, late-onset depression

compared with controls (study 12). The four further studies (studies 10, 11, 13

and 15) were cross-sectional in design, with studies 10, 11 and 15 using older-

adult community participants and study 13 involving a population of individuals

with generalised anxiety disorder (GAD). Studies 6 and 9 were from the same

research study, and were possibly compromised with overlapping participants

(although this is not explicitly noted within the papers). Sample size ranged from

n=23 (study 1) to 4352 (study 3), with 7423 participants included in total across

the 16 studies. As a result of inclusion criteria, all participants were over the age

of 60 years. Average age ranged between 65.0 years (study 12) and 73.4 years

(study 5).

The studies employed a range of depression and WM measures. Many

studies used more than one measure of each construct. The most commonly

used measure of mood was the Hamilton Rating Scale for Depression (HRSD),

which was used in seven studies (studies 2, 6-9, 11 and 13). Five studies used

the Structured Clinical Interview for DSM-IV (SCID; studies 1, 2, 6, 9 and 16),

and four used the Geriatric Depression Scale (GDS; studies 3, 10, 12 and 15).

---

[4] Study numbers throughout this section refer to those given within Table 3.

Other studies used a range of generally well-validated measures of depression

and anxiety. Digit span (in both forward and backward form) was the most

frequently used measure of WM, used in six studies (studies 7, 10, 11 and 14-

16). The *n*-back task (studies 1, 6 and 9) and the Digit Symbol Substitution Test

(studies 2, 6 and 9) were used in three studies each. A number of studies did

not explicitly refer to these tasks as measures of WM, instead describing them

as measures of 'attention', 'processing speed' or 'executive function'. The

variability in tests used may be a source of bias, given that no

neuropsychological measure can claim to be a pure assessment tool (e.g. digits

forwards as a measure of phonological loop, digits backwards as central

executive function; Redick & Lindsey, 2013).

Table 3

Summary of Studies Included in the Review

Study	Participants	Design	Measure of WM	Measure of mood	Main findings	**Comments on study quality**	Quality Rating (EPHPP QATQS)
1 Dumas & Newhouse (2015)	n=23 total; comparison of current MDD (n=11) and healthy controls (n=12) Mean age 73.3yrs	Cross-sectional case-control	Verbal N-back	SCID-IV	Depressed participants demonstrated decreased sensitivity on N-back task with increased WM load compared to healthy controls. Depressed participants were more conservative on 0-back trials than healthy controls. *Effect size:* insufficient raw data reported to enable calculation	*Strengths:* Clearly describes drop outs, high rate of study completion *Weaknesses:* Failed to report differences between groups, or attempts to control for confounds. Raw data not reported pre-analysis. Small sample size.	A – Moderate B – Moderate C – Weak D – Moderate E – Strong F – Strong Global rating – **Moderate**
2 Koenig et al. (2015)	n=438 total; comparison of currently depressed (n=120), previously depressed but currently euthymic (n=190), and never-depressed (n=128). Mean age 72.5yrs	Cross-sectional case-control	DSST*	SCID-IV HRSD-17	Significant difference in DSST performance across the 3 groups, such that previously and currently depressed participants performed worse than never-depressed participants. *Effect size:* d=0.6 for the difference between currently and never depressed groups on DSST	*Strengths:* Clear reporting of and controlling for confounds. Data collection tools had strong psychometric properties. *Weaknesses:* Assessors not blinded to categorisation of participants.	A – Moderate B – Moderate C – Strong D – Moderate E – Strong F – n/a Global rating - **Strong**
3 Shimada et al. (2014)	n=4352 total; comparison of depression (n=87), depressive symptoms (n=570), and no depression (n=3695) Mean age 71.1yrs	Cross-sectional case-control	SDST +	GDS-15	Significant group differences in SDST. Post-hoc analyses indicate poorer performance in 'depressed participants' as compared to 'no depressive symptoms' participants. SDST performance was a significant predictor of having depression (compared to no depressive symptoms) in logistic regression analyses. *Effect size*: d=0.3 for the difference between currently depressed and non-depressed groups on SDST	*Strengths:* Withdrawals from study and confounding variables both accurately reported. Full reporting of data in appendices. *Weaknesses:* No reporting of how many individuals were approached to participate.	A – Moderate B – Moderate C – Strong D – Moderate E – Strong F – Moderate Global rating - **Strong**

Study	Participants	Design	Measure of WM	Measure of mood	Main findings	Comments on study quality	Quality Rating (EPHPP QATQS)
4 Baune et al. (2006)	N=364 total; comparison of depressed (n=36) and non-depressed (n=328) Mean age 72.4yrs	Cross-sectional case-control	LDST*	CES-D	Non-significant difference between groups in LDST performance. LDST performance not significantly affected by increase in CES-D scores. Complex relationship between severity of depression and cognitive performance. *Effect size:* d=0.3 for the difference between depressed and non-depressed groups on LDST	*Strengths:* Some attempts to ensure participants representative of target population. Confounding variables identified and controlled. *Weaknesses:* Assessors not blinded to categorisation of participants.	A – Moderate B – Moderate C – Strong D – Moderate E – Strong F – Strong Global rating - **Strong**
5 Sair et al. (2006)	N-258 total; comparison of depressed (n=129) and non-depressed (n=129) Mean age 73.4yrs	Cross-sectional case-control	ADT	NIMH Diagnostic Interview Schedule MADRS	Depressed participants had significantly lower score on ADT. MADRS score was not significantly associated with ADT performance in depressed participants i.e. no association between depression severity and task performance. *Effect size:* insufficient raw data reported to enable calculation	*Strengths:* Assessment tools psychometrically strong. Confounders identified and controlled for. *Weaknesses:* No reporting of number of individuals approached to participate.	A – Moderate B – Moderate C – Strong D – Moderate E – Strong F – n/a Global rating - **Strong**
6 Nebes et al. (2003)	N=94 total; comparison of depressed (n=73) and non-depressed (n=21) Mean age 70.3yrs	Cross-sectional case-control	*N*-back test DSST +	SCID-IV HRSD-17	Depressed participants scored significantly lower on DSST and *N*-back tests. Secondary finding following treatment suggests that remission of depression does not relate to improvement in WM or processing speed performance. *Effect size:* d=0.9 for both the difference between depressed and non-depressed participants on the DSST and the *N*-back	*Strengths:* Data checked for confounding variables. *Weaknesses:* Unclear if data reported elsewhere. Poor reporting of study withdrawals. Not all participants completed relevant tasks.	A – Moderate B – Moderate C – Strong D – Weak E – Moderate F – Weak Global rating - **Weak**

Study	Participants	Design	Measure of WM	Measure of mood	Main findings	Comments on study quality	Quality Rating (EPHPP QATQS)
7 Elderkin-Thompson et al. (2003)	N=94 total; comparison of minor depressed (n=28); major depressed (n=26) and healthy controls (n=38) Mean age 69.6yrs	Cross-sectional case-control	Digit span (forward and backward)	HRSD-17 DSM-IV	WM did not significantly separate participants into relevant diagnostic groups, although there was a trend for separating healthy elderly from MDD. WM did not correlate with HRSD-17 scores for any of the three groups. WM did correlate significantly with depression in females. *Effect size:* d=0.7 for the difference between controls and major depression on DS	*Strengths:* Identified confounders controlled for during analysis. Multiple data collection used to triangulate identification. *Weaknesses:* Assessors not blinded to categorisation of participants.	A – Moderate B – Moderate C – Strong D – Moderate E – Strong F – n/a Global rating - **Strong**
8 Brodaty et al. (2001)	N=73 total; comparison of participants with early onset (n=35) and late onset (n=38) depression Mean age 68.4yrs	Cross-sectional case-control	SDMT *	HRSD-21 ZDS CORE NES	SDMT performance was not significantly different between participants with early and late onset depression. There were no apparent differences in the proportion of participants with significantly impaired performance (>2SD worse than normative data) in either group. *Effect size:* d=0.1 for the difference between late- and early-onset depression on SDMT.	*Strengths:* Description of population targeted and number of individuals consenting to participate. *Weaknesses:* Design unable to causally explain results.	A – Strong B – Moderate C – Strong D – Moderate E – Strong F – n/a Global rating - **Strong**
9 Nebes et al. (2000)	N=58 total; comparison of depressed (n=39) and control (n=19) participants Mean age 70.8yrs	Cross-sectional case-control	*N*-back test DSST +	SCID HRSD-17	Performance on both tasks was significantly worse for depressed participants than controls. HRSD scores did not correlate with either task performance. Performance on tests did not differ when participants with early- and late-onset depression were compared. Secondary finding that remission in depression with treatment was not reflected by notable improvement in performance on tasks of WM/PS (i.e. performance not improved more than is explainable by practise effects). Additional finding that WM explains great deal of variance in performance on other cognitive tests, which would otherwise be attributable to depression. *Effect size:* d=0.9 for difference between depressed participants and controls on the *N*-back test.	*Strengths:* Measurement tools have strong psychometric properties. Intervention (medication) adherence described. *Weaknesses:* Unclear if data reported elsewhere.	A – Moderate B – Moderate C – Strong D – Moderate E – Strong F – Strong Global rating - **Strong**

Study	Participants	Design	Measure of WM	Measure of mood	Main findings	Comments on study quality	Quality Rating (EPHPP QATQS)
10 Salazar-Villanea et al. (2015)	N=184 community participants total Mean age 68.4yrs	Cross-sectional	Digit Span (fwd and bwd) CAMCOG Attention/ Calculation subscale	GDS PANAS	Increased GDS score correlated with fewer digits recalled on digits backwards. Increased depression significantly predicted decreased WM capacity. *Effect size: $r=.06$* for relationship between GDS scores and latent WM factor.	*Strengths:* Wide range of well-validated tools used. *Weaknesses:* Cross-sectional design.	A – Moderate B – Weak C – Strong D – Moderate E – Strong F – n/a Global rating - **Moderate**
11 Potter, McQuoid, & Steffens (2015)	N=322 total Mean age 69.0yrs	Cross-sectional	Digit Span (fwd and bwd)	HRSD-17 MADRS	Greater depressive appetite loss correlated with lower digit span scores, however this relationship did not hold in multivariate regression models. Correlations between WM and other depression factors were non-significant. Depression scores did not correlate significantly with WM performance. *Effect size: $r=0.5$* for relationship between anhedonia and WM	*Strengths:* Confounders identified and controlled for within analysis. *Weaknesses:* Cross-sectional design.	A – Moderate B – Weak C – Strong D – Moderate E – Strong F – n/a Global rating - **Moderate**
12 Delaloye et al. (2010)	N=71 total; comparison between early onset depression (n=30), late onset depression (n=11) and healthy controls (n=30) Mean age 65.0yrs	Cross-sectional case control	LNST Reading span test	MINI GDS	No significant difference across groups in performance in either WM task. *Effect size: $d=0.6$* for difference in reading span and 0.9 for difference in LNST between late-onset and healthy controls	*Strengths:* Differentiation in sample between late- and early-onset. *Weaknesses:* Confounders identified but not controlled for within analysis. Assessment tools poorly validated. Small sample size.	A – Weak B – Moderate C – Weak D – Moderate E – Weak F – n/a Global rating - **Weak**

Study	Participants	Design	Measure of WM	Measure of mood	Main findings	Comments on study quality	Quality Rating (EPHPP QATQS)
13 Caudle et al. (2007)	N=208 adults with GAD total. Further analysis focused on participants who completed CBT treatment (n=65) Mean age 66.7yrs	Cohort	WM subscale of MMSE (serial 7s or 'world' backwards)	ADIS PSWQ HAMA BDI HDRS	In a comparison of participants who made WM errors and those who did not, there was significant difference in GAD severity, HDRS and BDI scores. Additional analysis showed no relationship between WM performance and response to treatment. *Effect size:* d=0.4 for GAD severity and 0.6 for HDRS scores between participants making and not-making WM errors.	*Strengths:* Withdrawals from study reported. *Weaknesses:* Poorly validated measure of WM. Consistency of CBT treatment described as 'similar' but not reported.	A – Moderate B – Moderate C – Strong D – Moderate E – Strong F – Strong Global rating - **Strong**
14 Korten et al. (2014)	N=510 total; comparison of depressed (n=378) and non-depressed (n=132) participants Mean age 70.1yrs	Cross-sectional case control	Digit Span (forward and backward)	CIDI IDS BAI Worry Scale	Depressed participants did not score significantly lower than non-depressed participants on digit span. Across the full sample, there were statistically significant correlations between WM performance and both anxiety and depression severity. Relationship between WM and worrying was non-significant. Higher mood and somatic symptom scores were related to worse WM. *Effect size:* d-.01 for WM Z-score when compared depressed and non-depressed participants.	*Strengths:* Study representative of target group. Study controls for identified and theoretical confounds within analysis. Study withdrawals reported in full. Effect sizes reported. *Weaknesses:* It is not reported what proportion of approached individuals declined participation.	A – Moderate B – Moderate C – Strong D – Moderate E – Strong F – Strong Global rating - **Strong**
15 Turner et al. (2015)	N=298 community African American participants total Mean age 73.9yrs	Prospective	Digit Span (fwd and bwd, digit ordering) Digit ordering	CES-D GDS	No relationship between CES-D and WM decline over 5-year follow up of study. However, higher GDS scores related to faster decline in WM over follow-up. When individuals with MCI excluded this relationship was no longer significant. Increased GDS negative affect associated with faster decline in WM. *Effect size:* Insufficient raw data reported to enable calculation.	*Strengths:* Proportion of approached individuals declining participation is reported. Wide range of relevant confounders identified and controlled for. *Weaknesses:* Assessors not blinded to categorisation of participants.	A – Moderate B – Moderate C – Strong D – Moderate E – Strong F – Weak Global rating - **Moderate**

Study	Participants	Design	Measure of WM	Measure of mood	Main findings	Comments on study quality	Quality Rating (EPHPP QATQS)
16 Yuan et al. (2009)	N=76 total; comparison of first-episode remitted geriatric depression (n=40) and healthy controls (n=36) Mean age 70.3yrs	Cross-sectional case control	Digit Span (unspecified which forms) SDMT	SCID	No significant difference between groups in digit span test performance. Significantly worse performance at SDMT by depressed participants compared to controls. *Effect size:* d=0.6 for SDMT and 0.3 for digit span when comparing depressed participants and healthy controls.	*Strengths:* Provides both imaging and behavioural measures. *Weaknesses:* Unclear whether selected participants are representative of the target population.	A – Moderate B – Moderate C – Strong D – Moderate E – Strong F – n/a Global rating - **Strong**

Note.

ADIS – Anxiety Disorders Interview Schedule
ADT - Ascending Digits Task
BAI – Beck Anxiety Inventory
BDI – Beck Depression Inventory
CES-D - Centre for Epidemiologic Studies Depression Scale
CIDI – Composite International Diagnostic Interview
DSST - Digit Symbol Substitution Test
EPHPP QATQS – Effective Public Health Practice Project Quality Assessment Tool for Quantitative Studies
GAD – Generalized Anxiety Disorder
GAS – Goldberg Anxiety Scale
GDS-15 – Geriatric Depression Scale
HAMA – Hamilton Rating Scale for Anxiety
HRSD – Hamilton Rating Scale for Depression
IDS – Inventory of Depressive Symptomatology
LDST – Letter Digit Substitution test
LNST – Letter Number Sequence Task
MADRS – Montgomery Asberg Depression Rating Scale

MDD – Major Depressive Disorder
MINI – Mini International Neuropsychiatric Inventory Interview
MMSE – Mini Mental State Examination
NES - Newcastle Endogeneity Scale
OA – Older Adult
PS – Processing Speed
PSWQ- Penn State Worry Questionnaire
QATQS ratings: A - Selection Bias, B - Study Design, C - Confounders, D - Blinding, E - Data Collection Method, F - Withdrawals and Dropouts
SCID - Structured Clinical Interview for DSM-IV
SDMT – Symbol Digit Modalities Test
SDST - Symbol Digit Substitution Test
WM – Working Memory
ZDS – Zung Depression Scale

* Described as a measure of attention/processing speed
+ Described as a test of processing speed

Risk of Bias within Studies

Studies were evaluated using the QATQS (EPHPP, 2009; Appendix B), identifying 10 studies of 'strong' quality (studies 2-5, 7-9 and 13-16), four of 'moderate' quality (studies 1, 10, 11 and 15) and two of 'weak' quality (studies 6 and 12).

All studies scored as moderate quality with regard to study design, due to all being either case-control or cohort studies without randomisation, introducing a possible source of bias within the studies. Ten of the studies used statistical methods to control for relevant confounders, which included age, gender and years of education. Four studies (studies 6, 9, 13 and 16) did not identify significant differences on these variables and therefore did not take steps to control for them within analysis. Study 1 did not describe information relating to confounders, and was therefore scored as 'weak' in quality for this area. Study 12 found differences between groups in gender, age and education but only controlled for age within analysis, also resulting in a 'weak' rating for confounds.

All studies appeared to use a single non-blinded assessor to collect ratings of mood and WM for all participant groups. This is a clear source of potential bias. Fourteen studies (all apart from studies 4 and 8) neglected to include information regarding specific recruitment numbers, both in terms of the proportion of individuals approached who agreed to participate in the study – introducing potential selection bias - and of individuals who withdrew during data collection or whose data was lost/not useable, i.e. attrition bias.

The majority of studies (14/16) were rated as strong quality for data collection methods, as studies tended to use reliable and valid tools to measure

both mood and WM. Inconsistencies were seen across studies in the specific

constructs assessed by the different measures of WM (i.e. SDMT was referred

to as a test of attention and processing speed in study 8, and as a measure of

WM by study 16).

It is relevant to note that studies did not tend to take participants'

medication status in to account, and as a result participants often varied in their

prescribed medication, both within and between study groups. It is therefore

impossible to ascertain the possible impact of medication on task performance.

Given that it is possible to hypothesise that the impact of psychotropic

medication for mood may operate by increasing WM capacity (Biringer,

Rongve, & Lund, 2009), studies which control for medication status are vital to

understanding the complex relationship between mood and WM.

Results of Individual Studies

See Appendix C for prose descriptions of the findings of individual

studies, as recommended within PRISMA guidelines (Moher, et al., 2009).

Risk of Bias across Studies

Alongside risk of bias within studies, it is important to consider what bias

may have been introduced across studies (Sterne, Egger & Moher, 2011). It is

increasingly clear that publication bias – the tendency for studies to be

published only if results are significant - impacts upon the literature available to

reviews. Many of the studies included report non-significant results. This may

provide some reassurance regarding the impact of publication bias, although in

the absence of analysis of a register of protocols it is not possible to define the

extent to which publication bias may have influenced the studies available to

this review.

All identified studies have been published since 2000, i.e. within the 16

years prior to the search. There is therefore a strong possibility that time-lag

bias, i.e. the tendency to publish significant or unusual results more quickly,

may have had an impact on this relatively young area of research. Efforts have

been taken to combat this in contacting relevant researchers in the field to

enquire about as-yet unpublished studies, however without response.

Outcome reporting bias has had a significant impact on the studies

included within the review. As can be seen in Figure 2, 34% (57/166) of studies

excluded at the point of full text review were excluded on the basis that the

relationship between mood and WM was not directly explored. These were

studies in which these variables were measured, but in which the specific

analyses around the relationship were not reported. It is possible that these

studies did not report these analysis due to lack of significance, therefore

leading to a bias in which only significant relationships between variables are

evidenced within the literature. Also of note here is that many of the studies

explored WM as a facet of cognition alongside a number of other cognitive

facets (e.g. processing speed, executive function). The number of analyses

undertaken both within and across studies is likely to lead to an increase in the

Type I error rate.

The majority of studies included primarily female participants (up to 70%

sample in study 15), suggesting possible participant bias. Given that study 7

demonstrated a differential relationship between mood and WM for female and

male participants, it is possible that the predominantly female participants

across studies have biased the results towards a non-generalisable finding. A

positive result in this review is evidence of studies in multiple cultural settings

(e.g. Latin America – study 10, China – study 16) and with participants of

different ethnicities (e.g. African Americans – study 15).

Discussion

Summary of evidence

The findings of the studies can be grouped according to their specific

designs and comparative groups of interest, in accordance with the guidance of

Popay and colleagues (2006). Table 4 attempts to summarise the results of the

studies with reference to their methodological quality.

Table 4

Synthesis of Findings from Studies Included in the Review

	Case-control analysis					Correlational analysis	Prospective analysis
	Relationship with WM				*Relationship with anxiety*		
	Depression vs. controls	**Early- vs. late-onset**	**Current depression vs. prev depression vs. controls**	**Depression vs. depression symptoms vs. controls**	**WM errors vs. no WM errors**	**Mood symptoms and WM**	**Relationship between treatment of mood and WM**
Evidence of relationship (p<.05)	1 3 5 6 9 16		2	3	13	10 11 14	
No evidence of relationship (p>.05)	4 14	8 9 12		7		5 7 9 15	6 9 13

Note. QATQS Quality rating: Strong, moderate, weak; WM – working memory

Studies may appear more than once in the table where investigations covered more than one area.

Half of the included studies compared ability on WM tasks between individuals who met diagnostic criteria for depression and healthy controls. Six of these eight studies (1, 3, 5, 6, 9 and 16) found statistically significant differences between the two groups, with worse performance on WM tasks in individuals with depression. Two studies (4 and 14) did not reject the null hypothesis. The studies that found a significant difference tended to employ measures of WM that are reliant on the manipulation of information within the WM store (e.g. DSST, ADT), whereas tests using tasks focusing more selectively on WM capacity (e.g. digit span; study 14) were non-significant. It may be that there is a specific relationship between depression and the manipulation aspects of WM. This has been echoed in studies comparing manipulation and straight-capacity tasks of WM in adults with depression (Joorman, Levens, & Gotlib, 2011). Evidence such as this has been used to support the 'impaired disengagement' hypothesis (Koster, De Lissnyder, Derakshan, & De Raedt, 2011) of the relationship between WM and depression. The hypothesis suggests that an inability to manipulate or remove information within the WM store predisposes individuals to ruminative thinking styles and, as a result, to depression. This theory suggests that interventions to improve impaired updating ability may help reduce depression.

All three studies that compared individuals with early-onset depression to those with late-onset depression found no significant difference in WM performance. Given the publication bias common in the literature, this is highly suggestive that there is no association between age at onset of depression and WM impairment. This is interesting, given that much of the theoretical literature is predictive of a difference; alluding to possible neurological or vascular origins

to late-onset depression (Baldwin & O'Brien, 2002). It may be that the relationship between depression and WM in OAs is reflective of the relationship seen in adults of working age (Christopher & MacDonald, 2005).

The only study (study 2) to compare currently depressed participants with those who had previously been depressed and controls, found support for a relationship between depression and WM. The study indicated that both previously and currently depressed participants had poorer performance on a WM task than never-depressed participants. This gives some support to the hypothesis that WM impairment is a trait, not state, variable in OAs who experience depression. However, further research is needed to replicate the finding.

Studies 3 and 7 compared individuals with diagnosable depression to individuals with some level of depressive symptomatology and healthy controls. Study 3 found a significant difference in WM performance across groups, attributable to the poorer performance in depressed participants as compared to controls. Although study 7 did not reach significance in this regard, there was evidence of a trend in the same direction. Interestingly, study 7 employed the *N*-back task, while study 3 used a substitution task. It is possible that the difference in tasks holds some bearing on strength of relationships displayed in the studies.

Study 13 compared individuals making errors on WM tasks in the MMSE to those who did not, finding higher levels of depression and anxiety in those with errors. It is important to note that this is within a population of individuals who already met diagnostic criteria for an anxiety disorder, and without the use of control group. It is therefore difficult to ascertain to what extent WM

impairment is unique to anxiety or depression as compared to the general

population. This was the only study to look at anxiety, and it is clear that further

research is needed before firm conclusions can be reached.

Of the seven studies that undertook correlational analysis between

measures of mood and WM, only 3 studies reported a significant relationship.

Three quarters of the studies finding no evidence of a relationship had a 'strong'

quality rating, while only 1/3 of those supporting a relationship did, lending

some weight to the former category. It therefore seems likely that severity of

depression or anxiety symptoms does not relate to worsening WM ability.

Although not a specific focus of this review, three of the studies (6, 9 and

13) went on to measure WM performance following successful treatment (both

pharmacological and psychotherapeutic). All three studies found no evidence

for an improvement in WM capacity following amelioration of depression, once

again indicating the trait nature of WM impairment in individuals who have

experienced depression.

Limitations

Study Level. The studies are, in general (10/16), of 'strong' quality,

meaning there is a low risk of bias inherent within the studies. However, six

studies do not meet this quality standard, including two that were graded as

'weak'. Several studies were limited in their reporting of blinding and of drop-

outs. None of the studies are randomised control trials, the 'gold standard' study

design, which limits the ability of included studies to control for confounds.

Importantly, the cross-sectional nature of the majority of the studies means that

it is not possible to reach conclusions as to the direction of the relationship

between mood and WM. Future studies may benefit from experimental study designs, e.g. the elicitation of low mood/anxiety, or inclusion of tasks to limit WM capacity, which would enable further investigation of the nature of the relationship.

Review Level. This review is the first to explore the unique relationship between WM and mood in OAs. As such, it offers an opportunity to bring together the existing literature and identify important directions for future work. The review is, however, limited in the number of studies contained (16) and in the relative infancy of the study area (all studies published since 2000). It is possible that relevant articles were missed during the screening process, although every effort was made to minimise the possibility of this occurring (e.g. searches of reference lists, approaching key authors).

A number of studies were excluded during screening as they did not sufficiently explore/report the relationship between WM and mood, meaning that the review may not contain all the available data on the subject and may be vulnerable to reporting bias.

Only one study explored the relationship between anxiety and WM, meaning that there is little evidence on which to draw firm conclusions. The significant result of this study suggests that further investigation is warranted.

Conclusions

The review has summarised and critiqued available research investigating a relationship between WM and mood in adults over the age of 60 years. The included studies comprised a range of non-randomised designs, providing evidence of a relationship between low mood and decreased WM

ability, particularly with regard to the ability to manipulate information within the WM store. It appears that impaired WM is a trait variable; with no evidence for a difference in WM ability in early- compared to late-onset depression, no improvement in WM with effective treatment and no relationship between severity of mood symptoms and degree of impairment. In other words, OAs with depression are likely to have WM impairment regardless of age at onset, degree of symptoms and treatment response. However, further research using experimental study designs is needed.

Funding

Clinical Implications

The conclusions of this review have clear clinical implications. There is robust evidence of the presence of trait impairment in at least some aspects of WM within OAs with low mood, regardless of age of depression onset. This impairment appears unlikely to be significantly improved with effective treatment of depression. It is therefore important that clinical services catering for the needs of OAs with depression are aware of the need to provide effective support for WM impairment, including the provision of well-paced therapeutic input, inclusive of tangible reminders or aide-mémoire where relevant.

Clinicians may also wish to explore with clients how they make sense of any substantive WM deficits, and provide psycho-education and normalisation based on the current literature. This is likely to be reassuring for OAs who may

be at risk of misinterpreting WM deficits as indicators of declining health. At the same time, the trait nature of WM impairment within depression suggests that any significant and recent decline in cognitive ability is unlikely to be causally related to low mood. Such presentations should receive onward referral to memory assessment services if needed.

While there is less evidence to support the role of WM impairment in anxiety in OAs, clinicians and clients are likely to benefit from an awareness of the possible role of WM impairment in the client's difficulties, and responding to and working on WM when this is clinically useful.

Take home points for clinicians are as follows[5]:

- OAs presenting to services with depression are likely to have some degree of deficit in WM. Services should be designed to engage with and support these individuals, e.g. by providing written summaries of sessions and appointment times to ease pressure on WM.
- WM deficits are likely to be persistent in OAs with depression. Therefore rehabilitation and compensatory strategies to ameliorate the impact of long-term WM deficits are likely to be helpful, alongside treatments targeting mood.

---

[5] This section is necessary for publication in Clinical Gerontology: *'The clinical implications section must include 2-3 bulleted "take home" points for clinicians which will be set out in a text box.'*

References

Armijo-Olivo, S., Stiles, C. R., Hagen, N. A., Biondo, P. D., & Cummings, G. G. (2012).

Assessment of study quality for systematic reviews: a comparison of the

Cochrane Collaboration Risk of Bias Tool and the Effective Public Health

Practice Project Quality Assessment tool: methodological research. *Journal of*

Evaluation in Clinical Practice, 18, 12-18.

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The*

psychology of learning and motivation: Advances in research and theory (Vol.

8, pp. 47–89). New York: Academic Press.

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory?

Trends in Cognitive Sciences, 4, 417-423.

Baldwin. R. C., & O'Brien, J. (2002). Vascular basis of late-onset depressive disorder.

British Journal of Psychiatry, 180, 157-160. doi: 10.1192/bjp.180.2.157

Baune, B. T., Suslow, T., Engelien, A., Arolt, V., & Berger, K. (2006). The Association

between depressive mood and cognitive performance in an elderly general

population – The MEMO study. *Dementia and Geriatric Cognitive Disorders, 22,*

142-149. doi: 10.1159/000093745

Biringer, E., Rongve, A., & Lund, A. (2009). A review of modern antidepressants'

effects on neurocognitive function. *Current Psychiatry Reviews, 5,* 164-176.

Bopp, K. L., & Verhaeghen, P. (2005). Aging and verbal memory span: a meta-

analysis. *Journal of Gerontology: Psychological Sciences & Social Sciences,*

60, 223-233.

Braver, T. S., & West. R (2007). Working memory, executive control and aging. In

    Craik, F.I.M. & Salthouse, T.A. (Eds.) *The handbook of aging and cognition:*

    Third edition (311-373). Sussex: Psychology Press.

Brodarty, H., Luscombe, G., Parker, G., Wilhelm, K., Hickie, I., Austin, M., & Mitchell, P.

    (2001). Early and late onset depression in old age: different aetiologies, same

    phenomenology. *Journal of Affective Disorders, 66,* 225-236.

Butters, M. A., Bhalla, R. K., Andreescu, C., Loebach Wetherall, J., Mantella, E.,

    Begley, A. E., & Lenze, E.J. (2011). Changes in neuropsychological functioning

    following treatment for late-life generalised anxiety disorder. *The British Journal*

    of Psychiatry, 199, 211-218 doi: 10.1192/bjp.bp.110.090217

Caudle, D. D., Senior, A. C., Wetherell, J. L., Rhoades, H. M., Beck, J. G., Kunik, M .E.,

    … Stanley, M. A. (2007). Cognitive errors, symptom severity, and response to

    cognitive behavior therapy in older adults with generalized anxiety disorder.

    American Journal of Geriatric Psychiatry, 15, 680-689

Chan, R., Shum, S., Toulopoulou, T., & Chen, E. (2008). Assessment of executive

    functions: Review of instruments and identification of critical issues. *Archives of*

    Clinical Neuropsychiatry, 23, 201-216. doi: 10.1016/j.acn.2007.08.010

Channon S., Baker J. E., & Robertson, M. M. (1993). Working memory in clinical

    depression: an experimental study. *Psychological Medicine,* 23, 87-91.

Christopher, G., & MacDonald, J. (2005). The impact of clinical depression on working

    memory. *Cognitive Neuropsychiatry, 10,* 379-399.

Daselaar, S., & Cabeza, R. (2013). Age-related decline in working memory and

    episodic memory. In Ochsner, K. N., & Kosslyn, S. (Eds) *The Oxford handbook*

    of cognitive neuroscience, volume 1: Core topics. UK: Oxford University Press.

    doi: 10.1093/oxfordhb/9780199988693.013.0022

Delaloye, C., Moy, G., de Bilbao, F., Baudois, S., Weber, K., Hofer, F., … Gold, G.

(2010). Neuroanatomical and neuropsychological features of elderly euthymic

depressed patients with early- and late-onset. *Journal of the Neurological*

Sciences, 299, 19-23.

Dobbs, A. R., & Rule, B. G.(1989). Adult age differences in working memory.

Psychology and Aging, 4, 500-503.

Dumas, J. A., & Newhouse, P. A. (2015). Impaired working memory in geriatric

depression: An fMRI study. *American Journal of Geriatric Psychiatry, 23,* 433-

436.

Effective Public Health Practice Project. (2009). *Quality Assessment Tool for*

Quantitative Studies. Retrieved from http://www.ephpp.ca/tools.html

Elderkin-Thompson, V., Kumar, A., Bilker, W. B., Dunking, J. J., Mintz, J., Moberg, P.

J., … Gur, R. E. (2003). Neuropsychological deficits among patients with late-

onset minor and major depression. *Achieves of Clinical Neuropsychology, 18,*

529-549.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current*

Directions in Psychological Science, 11, 19-23.

Gohier, B., Ferracci, L., Surguladze, S.A., Lawrence, E., El Hage, W., Zied Kefi, M., …

& Le Gall, D. (2009). Cognitive inhibition and working memory in unipolar

depression. *Journal of Affective Disorders, 116,* 100-105. doi:

10.1016/j.jad.2008.10.028

Harvey, P. O., Le Bastard, G., Pochon, J. B., Levy, R., Allilaire, J. F., Dubois, B., &

Fossati, P. (2004). Executive functions and updating of the contents of working

memory in unipolar depression. *Journal of Psychiatric Research, 38,* 567-576.

doi: 10.1016/j.jpsychires.2004.03.003

Joorman, J., Levens, S. M., & Gotlib, I. H. (2011). Sticky thoughts: depression and

rumination are associated with difficulties manipulating emotional material in

working memory. *Psychological Science, 22,* 979-983. doi:

10.1177/0956797611415539

Koenig, A. M., DeLozier, I. J., Zmuda, M. D., Marron, M. M., Begley, A. E., Anderson,

S. J., … Butters, M. A. (2015). Neuropsychological functioning in the acute and

remitted states of late-life depression. *Journal of Alzheimer's Disease, 45*, 175-

185. doi: 10.3233/JAD-148006.

Korten, N., Penninx, B., Kok, R. M., Stek, M. L., Voshaar, R., Deeg, D., & Comijs, H. C.

(2014). Heterogeneity of late-life depression: relationship with cognitive

functioning. *International Psychogeriatrics, 26,* 953-963.

Koster, E. H., De Lissnyder, E., Derakshan, N., & De Raedt, R. (2011). Understanding

depressive rumination from a cognitive science perspective: the impaired

disengagement hypothesis. *Clinical Psychology Review, 31,* 138-45. doi:

10.1016/j.cpr.2010.08.005.

Mental Health Foundation (2015). Mental health statistics: Older people. Retrieved

from: http://www.mentalhealth.org.uk/help-information/mental-health-

statistics/older-people/

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009).

Preferred reporting items for systematic reviews and meta-analyses: The

PRISMA Statement. *PLoS Med, 6*, e1000097.

doi:10.1371/journal.pmed.1000097

Nebes, R. D., Butters, M. A., Mulsant, B. H., Pollock. B. G., Zmuda, M. D., Houck, P.

R., & Reynolds, C. F. (2000). Decreased working memory and processing

speed mediate cognitive impairment in geriatric depression. *Psychological

Medicine, 30*, 679-691.

Nebes, R. D., Pollock, B. G., Houck, P. R., Butters, M. A., Mulsant, B. H., Zmuda, M. D., & Reynolds, C. F. (2003). Persistence of cognitive impairment in geriatric patients following antidepressant treatment: a randomized, double-blind clinical trial with nortriptyline and paroxetine. *Journal of Psychiatric Research, 37,* 99-108.

Owens, M., Koster, E. H., & Derakshan, N. (2012) Impaired filtering of irrelevant information in dysphoria: An ERP study. *Social Cognition and Affective Neuroscience, 7,* 752-763. doi: 10.1093/scan/nsr050

Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M…. Duffy, S. (2006). *Guidance on the conduct of narrative synthesis in systematic reviews: A product from the ESRC methods programme.* UK: Lancaster University.

Potter, G. G., McQuoid, D. R., & Steffens, D. C. (2015). Appetite loss and neurocognitive deficits in late-life depression. *International Journal of Geriatric Psychiatry, 30,* 647-654. doi: 10.1002/gps.4196

Redick, T. S., & Lindsey, D. (2013). Complex span and n-back measures of working memory: A meta-analysis. *Psychonomic Bulletin Review, 20*, 1102-1113. doi: 10.3758/s13423-013-0453-9

Rose, E. J., & Ebermeier, K. P. (2006). Pattern of impaired working memory during major depression. *Journal of Affective Disorders, 90*, 149-161. doi: 10.1016/j.jad.2005.11.003

Sair, H. I., Welsh-Bohmer, K. A., Wagner, H. R., & Steffens, D. C. (2006) Ascending digits task as a measure of executive function in geriatric depression. *The Journal of Neuropsychiatry and Clinical Neurosciences, 18*, 117-120.

Salazar-Villana, M., Liebmann, E., Garneir-Villarreal, M., Montenegro-Montenegro, E., & Johnson. D. K. (2015). Depressive symptoms affects working memory in

healthy older adult Hispanics. *Journal of Depression and Anxiety, 4,* doi:

10.4172/2167-1044.1000204.

Salthouse, T. A. (1994). The aging of working memory. *Neuropsychology, 8,* 535-543.

Shimada, H., Park, H., Makizako, H., Doi, T., Lee, S., & Suzuki, T. (2014). Depressive

symptoms and cognitive performance in older adults. *Journal of Psychiatric*

Research, 57, 149-156.

Snyder, H. (2013). Major depressive disorder is associated with broad impairments on

neuropsychological measures of executive function: A meta-analysis and

review. *Psychological Bulletin, 139*, 81–132. doi: 10.1037/a0028727

Sterne, J., Egger, M., & Moher, D. (2011). Addressing reporting biases. In Higgins, J.,

& Green, S. (Eds.) *Cochrane handbook for systematic reviews of interventions.*

Version 5.1.0 [updated March 2011]. The Cochrane Collaboration, 2011.

Available from www.handbook.cochrane.org.

Turner, A. D., Capuano, A. W., Wilson, R. S., & Barnes, L. L. (2015). Depressive

symptoms and cognitive decline in older African Americans: Two scales and

their factors. *American Journal of Geriatric Psychiatry, 23,* 568-578. doi:

10.1016/j.jagp.2014.08.003.

von Hippel, W., Vasey, M., Gonda, T., & Stern, T. (2008). Executive function deficits,

rumination and late-onset depressive symptoms in older adults. *Cognitive*

Therapy and Research, 32, 474-487. doi: 10.1007/s10608-006-9034-9

Walters, J. (2016). *Effects of trait anxiety and depression on working memory updating*

(Unpublished dissertation). University of Colorado, Boulder: USA.

World Health Organization. (1992). *The ICD 10 classification of mental and behavioural*

disorders: Clinical descriptions and diagnostic guidelines. Geneva: World

Health Organization. Retrieved from

http://www.who.int/classifications/icd/en/bluebook.pdf on 13th March 2017.

Yuan, Y., Zhang, Z., Bai, F., Yu, H., You, J., Shi, Y., … Jiang, T. (2009). Larger

regional white matter volume is associated with executive function deficit in

remitted geriatric depression: An optimized voxel-based morphometry study.

Journal of Affective Disorders, 115, 225-229.

Appendices

Appendix A: QATQS

QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES

EPHPP
Effective Public Health Practice Project

COMPONENT RATINGS

A)   SELECTION BIAS

(Q1)   Are the individuals selected to participate in the study likely to be representative of the target population?
1   Very likely
2   Somewhat likely
3   Not likely
4   Can't tell

(Q2)   What percentage of selected individuals agreed to participate?
1   80 - 100% agreement
2   60 – 79% agreement
3   less than 60% agreement
4   Not applicable
5   Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

B)   STUDY DESIGN

Indicate the study design
1   Randomized controlled trial
2   Controlled clinical trial
3   Cohort analytic (two group pre + post)
4   Case-control
5   Cohort (one group pre + post (before and after))
6   Interrupted time series
7   Other specify _____
8   Can't tell

Was the study described as randomized? If NO, go to Component C.
        No   Yes

If Yes, was the method of randomization described? (See dictionary)
        No   Yes

If Yes, was the method appropriate? (See dictionary)
        No   Yes

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

Appendix B: QATQS Ratings of Included Studies

2: Nevo 4 Ta

QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES

EPHPP — Effective Public Health Practice Project

COMPONENT RATINGS

A) SELECTION BIAS

B) STUDY DESIGN

C) CONFOUNDERS

D) BLINDING

E) DATA COLLECTION METHODS

F) WITHDRAWALS AND DROP-OUTS

G) INTERVENTION INTEGRITY

H) ANALYSES

GLOBAL RATING

3: Dep S + cog perform.

QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES

EPHPP
Effective Public Health Practice Project

COMPONENT RATINGS

A) SELECTION BIAS

B) STUDY DESIGN

C) CONFOUNDERS

D) BLINDING

E) DATA COLLECTION METHODS

F) WITHDRAWALS AND DROP-OUTS

752 / 5104 = 14%

N/a.

G) INTERVENTION INTEGRITY

H) ANALYSES

N/A.

GLOBAL RATING

COMPONENT RATINGS

		STRONG	MODERATE	WEAK
A	SELECTION BIAS	1	2	3
B	STUDY DESIGN	1	2	3
C	CONFOUNDERS	1	2	3
D	BLINDING	1	2	3
E	DATA COLLECTION METHOD	1	2	3
F	WITHDRAWALS AND DROPOUTS	1	2	3	Not Applicable

GLOBAL RATING FOR THIS PAPER (circle one)

1 STRONG
2 MODERATE
3 WEAK

Final decision of both reviewers (circle one)   1   STRONG
2   MODERATE
3   WEAK

4: Assoc between dep need, cog performance.

QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES

EPHPP
Effective Public Health Practice Project

COMPONENT RATINGS

A) SELECTION BIAS

(Q1) Are the individuals selected to participate in the study likely to be representative of the target population?
1 Very likely
2 Somewhat likely
3 Not likely
4 Can't tell

(Q2) What percentage of selected individuals agreed to participate?
1 80 - 100% agreement
2 60 - 79% agreement
3 less than 60% agreement
4 Not applicable
5 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

B) STUDY DESIGN

Indicate the study design
1 Randomized controlled trial
2 Controlled clinical trial
3 Cohort analytic (two group pre + post)
4 Case-control
5 Cohort (one group pre + post (before and after))
6 Interrupted time series
7 Other specify _____
8 Can't tell

Was the study described as randomized? If NO, go to Component C.
No   Yes

If Yes, was the method of randomization described? (See dictionary)
No   Yes

If Yes, was the method appropriate? (See dictionary)
No   Yes

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

C) CONFOUNDERS

(Q1) Were there important differences between groups prior to the intervention?
1 Yes
2 No
3 Can't tell

The following are examples of confounders:
1 Race
2 Sex
3 Marital status/family
4 Age
5 SES (income or class)
6 Education
7 Health status
8 Pre-intervention score on outcome measure

(Q2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design (e.g. stratification, matching) or analysis)?
1 80 - 100% (most)
2 60 - 79% (some)
3 Less than 60% (few or none)
4 Can't Tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

D) BLINDING

(Q1) Was (were) the outcome assessor(s) aware of the intervention or exposure status of participants?
1 Yes
2 No
3 Can't tell

(Q2) Were the study participants aware of the research question?
1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

E) DATA COLLECTION METHODS

(Q1) Were data collection tools shown to be valid?
1 Yes
2 No
3 Can't tell

(Q2) Were data collection tools shown to be reliable?
1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

F) WITHDRAWALS AND DROP-OUTS

(Q1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?
1 Yes
2 No
3 Can't tell
4 Not Applicable (i.e. one time surveys or interviews)

(Q2) Indicate the percentage of participants completing the study. (If the percentage differs by groups, record the lowest).
1 80 - 100%
2 60 - 79%
3 less than 60%
4 Can't tell
5 Not Applicable (i.e. Retrospective case-control)

314 / 335 = 94.5%

RATE THIS SECTION	STRONG	MODERATE	WEAK	
See dictionary	1	2	3	Not Applicable

G) INTERVENTION INTEGRITY

(Q1) What percentage of participants received the allocated intervention or exposure of interest?
1 80 - 100%
2 60 - 79%
3 less than 60%
4 Can't tell

n/a.

(Q2) Was the consistency of the intervention measured?
1 Yes
2 No
3 Can't tell

(Q3) Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results?
4 Yes
5 No
6 Can't tell

H) ANALYSES

(Q1) Indicate the unit of allocation (circle one)
community organization/institution practice/office individual

(Q2) Indicate the unit of analysis (circle one)
community organization/institution practice/office individual

(Q3) Are the statistical methods appropriate for the study design?
1 Yes
2 No
3 Can't tell

(Q4) Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?
1 Yes
2 No
3 Can't tell

N/a.

GLOBAL RATING

COMPONENT RATINGS
Please transcribe the information from the grey boxes on pages 1-4 onto this page. See dictionary on how to rate this section.

		STRONG	MODERATE	WEAK	
A	SELECTION BIAS	1	2	3	
B	STUDY DESIGN	1	2	3	
C	CONFOUNDERS	1	2	3	
D	BLINDING	1	2	3	
E	DATA COLLECTION METHOD	1	2	3	
F	WITHDRAWALS AND DROPOUTS	1	2	3	Not Applicable

GLOBAL RATING FOR THIS PAPER (circle one):

1 STRONG (no WEAK ratings)
2 MODERATE (one WEAK rating)
3 WEAK (two or more WEAK ratings)

With both reviewers discussing the ratings:

Is there a discrepancy between the two reviewers with respect to the component (A-F) ratings?
No   Yes

If yes, indicate the reason for the discrepancy
1 Oversight
2 Differences in interpretation of criteria
3 Differences in interpretation of study

Final decision of both reviewers (circle one):
1 STRONG
2 MODERATE
3 WEAK

5: Ascending digits task as measure of EF.

QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES

EPHPP
Effective Public Health Practice Project

COMPONENT RATINGS

A) SELECTION BIAS

(Q1) Are the individuals selected to participate in the study likely to be representative of the target population?

(Q2) What percentage of selected individuals agreed to participate?

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

B) STUDY DESIGN

Indicate the study design

Was the study described as randomized? If NO, go to Component C.

If Yes, was the method of randomization described? (See dictionary)

If Yes, was the method appropriate? (See dictionary)

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

C) CONFOUNDERS

(Q1) Were there important differences between groups prior to the intervention?

The following are examples of confounders:

(Q2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design (e.g. stratification, matching) or analysis)?

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

D) BLINDING

(Q1) Was the outcome assessor aware of the intervention or exposure status of participants?

(Q2) Were the study participants aware of the research question?

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

E) DATA COLLECTION METHODS

(Q1) Were data collection tools shown to be valid?

(Q2) Were data collection tools shown to be reliable?

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

F) WITHDRAWALS AND DROP-OUTS

(Q1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?

N/a.

(Q2) Indicate the percentage of participants completing the study. (If the percentage differs by groups, record the lowest).

RATE THIS SECTION	STRONG	MODERATE	WEAK	
See dictionary	1	2	3	Not Applicable

G) INTERVENTION INTEGRITY

(Q1) What percentage of participants received the allocated intervention or exposure of interest?

(Q2) Was the consistency of the intervention measured?

N/a.

(Q3) Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results?

H) ANALYSES

(Q1) Indicate the unit of allocation (circle one)

(Q2) Indicate the unit of analysis (circle one)

(Q3) Are the statistical methods appropriate for the study design?

(Q4) Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?

N/a.

GLOBAL RATING

COMPONENT RATINGS

Please transcribe the information from the grey boxes on pages 1-4 onto this page. See dictionary on how to rate this section.

		STRONG	MODERATE	WEAK
A	SELECTION BIAS	1	2	3
B	STUDY DESIGN	1	2	3
C	CONFOUNDERS	1	2	3
D	BLINDING	1	2	3
E	DATA COLLECTION METHOD	1	2	3
F	WITHDRAWALS AND DROPOUTS	1	2	3	Not Applicable

GLOBAL RATING FOR THIS PAPER (circle one):

1 STRONG (no WEAK ratings)
2 MODERATE (one WEAK rating)
3 WEAK (two or more WEAK ratings)

With both reviewers discussing the ratings:

Is there a discrepancy between the two reviewers with respect to the component (A-F) ratings?

No   Yes

If yes, indicate the reason for the discrepancy

1 Oversight
2 Differences in interpretation of criteria
3 Differences in interpretation of study

Final decision of both reviewers (circle one):

1 STRONG
2 MODERATE
3 WEAK

6: Resistance of coy equipment

Quality Assessment Tool for Quantitative Studies (EPHPP) — Effective Public Health Practice Project

7: Neurol deficits

QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES — EPHPP (Effective Public Health Practice Project)

8: Early- & late-onset dep.

QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES

EPHPP — Effective Public Health Practice Project

COMPONENT RATINGS

A) SELECTION BIAS

(Q1) Are the individuals selected to participate in the study likely to be representative of the target population?
1 Very likely
2 Somewhat likely
3 Not likely
4 Can't tell

(Q2) What percentage of selected individuals agreed to participate?
1 80 - 100% agreement
2 60 - 79% agreement
3 less than 60% agreement
4 Not applicable
5 Can't tell

73/407 =

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

B) STUDY DESIGN

Indicate the study design
1 Randomized controlled trial
2 Controlled clinical trial
3 Cohort analytic (two group pre + post)
4 Case-control
5 Cohort (one group pre + post (before and after))
6 Interrupted time series
7 Other specify _____
8 Can't tell

Was the study described as randomized? If NO, go to Component C.
No Yes

If Yes, was the method of randomization described? (See dictionary)
No Yes

If Yes, was the method appropriate? (See dictionary)
No Yes

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

C) CONFOUNDERS

(Q1) Were there important differences between groups prior to the intervention?
1 Yes
2 No
3 Can't tell

The following are examples of confounders:
1 Race
2 Sex
3 Marital status/family
4 Age
5 SES (income or class)
6 Education
7 Health status
8 Pre-intervention score on outcome measure

(Q2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design (e.g. stratification, matching) or analysis)?
1 80 - 100% (most)
2 60 - 79% (some)
3 Less than 60% (few or none)
4 Can't Tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

D) BLINDING

(Q1) Was (were) the outcome assessor(s) aware of the intervention or exposure status of participants?
1 Yes
2 No
3 Can't tell

(Q2) Were the study participants aware of the research question?
1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

E) DATA COLLECTION METHODS

(Q1) Were data collection tools shown to be valid?
1 Yes
2 No
3 Can't tell

(Q2) Were data collection tools shown to be reliable?
1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

F) WITHDRAWALS AND DROP-OUTS

(Q1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?
1 Yes
2 No
3 Can't tell
4 Not Applicable (i.e. one time survey or interview)

(Q2) Indicate the percentage of participants completing the study. (If the percentage differs by groups, record the lowest).
1 80 -100%
2 60 - 79%
3 less than 60%
4 Can't tell
5 Not Applicable (i.e. Retrospective case-control)

RATE THIS SECTION	STRONG	MODERATE	WEAK	
See dictionary	1	2	3	Not Applicable

G) INTERVENTION INTEGRITY

(Q1) What percentage of participants received the allocated intervention or exposure of interest?
1 80 -100%
2 60 - 79%
3 less than 60%
4 Can't tell

(Q2) Was the consistency of the intervention measured?
1 Yes
2 No
3 Can't tell

N/R.

(Q3) Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results?
1 Yes
2 No
3 Can't tell

H) ANALYSES

(Q1) Indicate the unit of allocation (circle one)
community organization/institution practice/office individual

(Q2) Indicate the unit of analysis (circle one)
community organization/institution practice/office individual

(Q3) Are the statistical methods appropriate for the study design?
1 Yes
2 No
3 Can't tell

(Q4) Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?
1 Yes
2 No
3 Can't tell

GLOBAL RATING

COMPONENT RATINGS
Please transcribe the information from the gray boxes on pages 1-4 onto this page. See dictionary on how to rate this section.

		STRONG	MODERATE	WEAK
A	SELECTION BIAS	STRONG	MODERATE	WEAK
		1	2	3
B	STUDY DESIGN	STRONG	MODERATE	WEAK
		1	2	3
C	CONFOUNDERS	STRONG	MODERATE	WEAK
		1	2	3
D	BLINDING	STRONG	MODERATE	WEAK
		1	2	3
E	DATA COLLECTION METHOD	STRONG	MODERATE	WEAK
		1	2	3
F	WITHDRAWALS AND DROPOUTS	STRONG	MODERATE	WEAK	
		1	2	3	Not Applicable

GLOBAL RATING FOR THIS PAPER (circle one):
1 STRONG (no WEAK ratings)
2 MODERATE (one WEAK rating)
3 WEAK (two or more WEAK ratings)

With both reviewers discussing the ratings:

Is there a discrepancy between the two reviewers with respect to the component (A-F) ratings?
No Yes

If yes, indicate the reason for the discrepancy
1 Oversight
2 Differences in interpretation of criteria
3 Differences in interpretation of study

Final decision of both reviewers (circle one)
1 STRONG
2 MODERATE
3 WEAK

9: Dec WM, B (handwritten)

QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES

EPHPP
Effective Public Health Practice Project

COMPONENT RATINGS

A) SELECTION BIAS

(Q1) Are the individuals selected to participate in the study likely to be representative of the target population?
1. Very likely
2. Somewhat likely
3. Not likely
4. Can't tell

(Q2) What percentage of selected individuals agreed to participate?
1. 80 - 100% agreement
2. 60 - 79% agreement
3. less than 60% agreement
4. Not applicable
5. Can't tell

⟹ *same sample as b.* (handwritten)

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

B) STUDY DESIGN

Indicate the study design
1. Randomized controlled trial
2. Controlled clinical trial
3. Cohort analytic (two group pre + post)
4. Case-control
5. Cohort (one group pre + post (before and after))
6. Interrupted time series
7. Other specify _____
8. Can't tell

Was the study described as randomized? If NO, go to Component C.
No   Yes

If Yes, was the method of randomization described? (See dictionary)
No   Yes

If Yes, was the method appropriate? (See dictionary)
No   Yes

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

C) CONFOUNDERS

(Q1) Were there important differences between groups prior to the intervention?
1. Yes
2. No
3. Can't tell

The following are examples of confounders:
1. Race
2. Sex
3. Marital status/family
4. Age
5. SES (income or class)
6. Education
7. Health status
8. Pre-intervention score on outcome measure

(Q2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design (e.g. stratification, matching) or analysis)?
1. 80 - 100% (most)
2. 60 - 79% (some)
3. Less than 60% (few or none)
4. Can't Tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

D) BLINDING

(Q1) Was (were) the outcome assessor(s) aware of the intervention or exposure status of participants?
1. Yes
2. No
3. Can't tell

(Q2) Were the study participants aware of the research question?
1. Yes
2. No
3. Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

E) DATA COLLECTION METHODS

(Q1) Were data collection tools shown to be valid?
1. Yes
2. No
3. Can't tell

(Q2) Were data collection tools shown to be reliable?
1. Yes
2. No
3. Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

F) WITHDRAWALS AND DROP-OUTS

(Q1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?
1. Yes
2. No
3. Can't tell
4. Not Applicable (i.e. no drop-outs or interviews)

(Q2) Indicate the percentage of participants completing the study. (If the percentage differs by group, record the lowest)
1. 80 -100%
2. 60 - 79%
3. less than 60%
4. Can't tell
5. Not Applicable (i.e. Retrospective case-control)

RATE THIS SECTION	STRONG	MODERATE	WEAK	
See dictionary	1	2	3	Not Applicable

E) INTERVENTION INTEGRITY

(Q1) What percentage of participants received the allocated intervention or exposure of interest?
1. 80 -100%
2. 60 - 79%
3. less than 60%
4. Can't tell

(Q2) Was the consistency of the intervention measured?
1. Yes
2. No
3. Can't tell

(Q3) Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results?
1. Yes
2. No
3. Can't tell

H) ANALYSES

(Q1) Indicate the unit of allocation (circle one)
community organization/institution practice/office individual

(Q2) Indicate the unit of analysis (circle one)
community organization/institution practice/office individual

(Q3) Are the statistical methods appropriate for the study design?
1. Yes
2. No
3. Can't tell

(Q4) Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?
1. Yes
2. No
3. Can't tell

GLOBAL RATING

COMPONENT RATINGS
Please transcribe the information from the grey boxes on pages 1-4 onto this page. See dictionary on how to rate this section.

		STRONG	MODERATE	WEAK	
A	SELECTION BIAS	1	2	3	
B	STUDY DESIGN	1	2	3	
C	CONFOUNDERS	1	2	3	
D	BLINDING	1	2	3	
E	DATA COLLECTION METHODS	1	2	3	
F	WITHDRAWALS AND DROPOUTS	1	2	3	Not Applicable

GLOBAL RATING FOR THIS PAPER (circle one):
1. STRONG (no WEAK ratings)
2. MODERATE (one WEAK rating)
3. WEAK (two or more WEAK ratings)

With both reviewers discussing the ratings:

Is there a discrepancy between the two reviewers with respect to the component (A-F) ratings?
No   Yes

If yes, indicate the reason for the discrepancy
1. Oversight
2. Differences in interpretation of criteria
3. Differences in interpretation of study

Final decision of both reviewers (circle one)
1. STRONG
2. MODERATE
3. WEAK

10: Dp Se affeb.WM.

QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES

EPHPP
Effective Public Health Practice Project

COMPONENT RATINGS

A) SELECTION BIAS

(Q1) Are the individuals selected to participate in the study likely to be representative of the target population?
1 Very likely
2 Somewhat likely
3 Not likely
4 Can't tell

(Q2) What percentage of selected individuals agreed to participate?
1 80 - 100% agreement
2 60 - 79% agreement
3 less than 60% agreement
4 Not applicable
5 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

B) STUDY DESIGN

Indicate the study design
1 Randomized controlled trial
2 Controlled clinical trial
3 Cohort analytic (two group pre + post)
4 Case-control
5 Cohort (one group pre + post (before and after))
6 Interrupted time series
7 Other specify _____ cross-sectional
8 Can't tell

Was the study described as randomized? If NO, go to Component C.
No   Yes

If Yes, was the method of randomization described? (See dictionary)
No   Yes

If Yes, was the method appropriate? (See dictionary)
No   Yes

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

C) CONFOUNDERS

(Q1) Were there important differences between groups prior to the intervention?   N/A
1 Yes
2 No
3 Can't tell

The following are examples of confounders:
- Race
- Sex
- Marital status/family
- Age
- SES (income or class)
- Education
- Health status
- Pre-intervention score on outcome measure

(Q2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design (e.g. stratification, matching) or analysis)?
1 80 - 100% (most)
2 60 - 79% (some)
3 Less than 60% (few or none)
4 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

D) BLINDING

(Q1) Was (were) the outcome assessor(s) aware of the intervention or exposure status of participants?
1 Yes
2 No
3 Can't tell

(Q2) Were the study participants aware of the research question?
1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

E) DATA COLLECTION METHODS

(Q1) Were data collection tools shown to be valid?
1 Yes
2 No
3 Can't tell

(Q2) Were data collection tools shown to be reliable?
1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

F) WITHDRAWALS AND DROP-OUTS

(Q1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?
1 Yes
2 No
3 Can't tell
4 Not Applicable (i.e. one time surveys or interviews)

(Q2) Indicate the percentage of participants completing the study. (If the percentage differs by groups, record the lowest)
1 80 -100%
2 60 - 79%
3 less than 60%
4 Can't tell
5 Not Applicable (i.e. Retrospective case-control)

RATE THIS SECTION	STRONG	MODERATE	WEAK	Not Applicable
See dictionary	1	2	3	

G) INTERVENTION INTEGRITY

(Q1) What percentage of participants received the allocated intervention or exposure of interest?   N/A
1 80 -100%
2 60 - 79%
3 less than 60%
4 Can't tell

(Q2) Was the consistency of the intervention measured?
1 Yes
2 No
3 Can't tell

(Q3) Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results?
4 Yes
5 No
6 Can't tell

H) ANALYSES

(Q1) Indicate the unit of allocation (circle one)
community   organization/institution   practice/office   individual

(Q2) Indicate the unit of analysis (circle one)
community   organization/institution   practice/office   individual

(Q3) Are the statistical methods appropriate for the study design?
1 Yes
2 No
3 Can't tell

(Q4) Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?
1 Yes
2 No
3 Can't tell

GLOBAL RATING

COMPONENT RATINGS
Please transcribe the information from the grey boxes on pages 1-4 onto this page. See dictionary on how to rate this section.

		STRONG	MODERATE	WEAK
A	SELECTION BIAS	1	2	3
B	STUDY DESIGN	1	2	3
C	CONFOUNDERS	1	2	3
D	BLINDING	1	2	3
E	DATA COLLECTION METHOD	1	2	3
F	WITHDRAWALS AND DROPOUTS	1	2	3

GLOBAL RATING FOR THIS PAPER (circle one):

1 STRONG (no WEAK ratings)
2 MODERATE (one WEAK rating)
3 WEAK (two or more WEAK ratings)

With both reviewers discussing the ratings:

Is there a disagreement between the two reviewers with respect to the component (A-F) ratings?
No   Yes

If yes, indicate the reason for the discrepancies
1 Oversight
2 Differences in interpretation of criteria
3 Differences in interpretation of study

Final decision of both reviewers (circle one)
1 STRONG
2 MODERATE
3 WEAK

11: Appetite loss, renal deficits

QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES

EPHPP — Effective Public Health Practice Project

COMPONENT RATINGS

A) SELECTION BIAS

(B1) Are the individuals selected to participate in the study likely to be representative of the target population?

(B2) What percentage of selected individuals agreed to participate?

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

B) STUDY DESIGN

Indicate the study design

cross-sectional

Was the study described as randomized? If NO, go to Component C.

If Yes, was the method of randomization described? (See dictionary)

If Yes, was the method appropriate? (See dictionary)

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

C) CONFOUNDERS

(C1) Were there important differences between groups prior to the intervention?

N/A

The following are examples of confounders:

(C2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design e.g. stratification, matching or analysis)?

Ag, 4E, sex, race, hello, illness

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

D) BLINDING

(D1) Was the outcome assessor aware of the intervention or exposure status of participants?

(D2) Were the study participants aware of the research question?

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

E) DATA COLLECTION METHODS

(E1) Were data collection tools shown to be valid?

(E2) Were data collection tools shown to be reliable?

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

F) WITHDRAWALS AND DROP-OUTS

(F1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?

(F2) Indicate the percentage of participants completing the study. (If the percentage differs by groups, record the lowest).

RATE THIS SECTION	STRONG	MODERATE	WEAK	
See dictionary	1	2	3	Not Applicable

G) INTERVENTION INTEGRITY

(G1) What percentage of participants received the allocated intervention or exposure of interest?

N/A.

(G2) Was the consistency of the intervention measured?

(G3) Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results?

H) ANALYSES

(H1) Indicate the unit of allocation (circle one)

(H2) Indicate the unit of analysis (circle one)

(H3) Are the statistical methods appropriate for the study design?

(H4) Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?

GLOBAL RATING

COMPONENT RATINGS

		STRONG	MODERATE	WEAK
A.	SELECTION BIAS	1	2	3
B.	STUDY DESIGN	1	2	3
C.	CONFOUNDERS	1	2	3
D.	BLINDING	1	2	3
E.	DATA COLLECTION METHOD	1	2	3
F.	WITHDRAWALS AND DROPOUTS	1	2	3

GLOBAL RATING FOR THIS PAPER (circle one):

1. STRONG (no WEAK ratings)
2. MODERATE (one WEAK rating)
3. WEAK (two or more WEAK ratings)

With both reviewers discussing the ratings:

Is there a discrepancy between the two reviewers with respect to the component (A-F) ratings?

If yes, indicate the reason for the discrepancy:

1. Oversight
2. Differences in interpretation of criteria
3. Differences in interpretation of study

Final decision of both reviewers (circle one): 1 STRONG 2 MODERATE 3 WEAK

12: Nonanatomical, neuro & features

QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES

EPHPP — Effective Public Health Practice Project

COMPONENT RATINGS

A) SELECTION BIAS

(Q1) Are the individuals selected to participate in the study likely to be representative of the target population?
1 Very likely
2 Somewhat likely
3 Not likely
4 Can't tell

(Q2) What percentage of selected individuals agreed to participate?
1 80 - 100% agreement
2 60 - 79% agreement
3 less than 60% agreement
4 Not applicable
5 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

B) STUDY DESIGN

Indicate the study design
1 Randomized controlled trial
2 Controlled clinical trial
3 Cohort analytic (two group pre + post)
4 Case control
5 Cohort (one group pre + post (before and after))
6 Interrupted time series
7 Other specify
8 Can't tell

Was the study described as randomized? If NO, go to Component C.
No Yes

If Yes, was the method of randomization described? (See dictionary)
No Yes

If Yes, was the method appropriate? (See dictionary)
No Yes

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

C) CONFOUNDERS

(Q1) Were there important differences between groups prior to the intervention?
1 Yes
2 No
3 Can't tell

The following are examples of confounders:
1 Race
2 Sex
3 Marital status/family
4 Age
5 SES (income or class)
6 Education
7 Health status
8 Pre-intervention score on outcome measure

(Q2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design (e.g. stratification, matching) or analysis)?
1 80 - 100% (most)
2 60 - 79% (some)
3 Less than 60% (few or none)
4 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

D) BLINDING

(Q1) Was (were) the outcome assessor(s) aware of the intervention or exposure status of participants?
1 Yes
2 No
3 Can't tell

(Q2) Were the study participants aware of the research question?
1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

E) DATA COLLECTION METHODS

(Q1) Were data collection tools shown to be valid?
1 Yes
2 No
3 Can't tell

(Q2) Were data collection tools shown to be reliable?
1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

F) WITHDRAWALS AND DROP-OUTS

(Q1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?
1 Yes
2 No
3 Can't tell
4 Not Applicable (i.e. one time survey or interview)

(Q2) Indicate the percentage of participants completing the study. (If the percentage differs by groups, record the lowest).
1 80 -100%
2 60 - 79%
3 less than 60%
4 Can't tell
5 Not Applicable (i.e. Retrospective case-control)

RATE THIS SECTION	STRONG	MODERATE	WEAK	
See dictionary	1	2	3	Not Applicable

G) INTERVENTION INTEGRITY

(Q1) What percentage of participants received the allocated intervention or exposure of interest?
1 80-100%
2 60-79%
3 less than 60%
4 Can't tell

(Q2) Was the consistency of the intervention measured?   N/A
1 Yes
2 No
3 Can't tell

(Q3) Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results?
4 Yes
5 No
6 Can't tell

H) ANALYSES

(Q1) Indicate the unit of allocation (circle one)
community organization/institution practice/office individual

(Q2) Indicate the unit of analysis (circle one)
community organization/institution practice/office individual

(Q3) Are the statistical methods appropriate for the study design?
1 Yes
2 No
3 Can't tell

(Q4) Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?
1 Yes
2 No
3 Can't tell

GLOBAL RATING

COMPONENT RATINGS
Please transcribe the information from the gray boxes on pages 1-4 onto this page. See dictionary on how to rate this section.

		STRONG	MODERATE	WEAK	
A	SELECTION BIAS	1	2	3	
B	STUDY DESIGN	STRONG	MODERATE	WEAK	
		1	2	3	
C	CONFOUNDERS	STRONG	MODERATE	WEAK	
		1	2	3	
D	BLINDING	STRONG	MODERATE	WEAK	
		1	2	3	
E	DATA COLLECTION METHOD	STRONG	MODERATE	WEAK	
		1	2	3	
F	WITHDRAWALS AND DROPOUTS	STRONG	MODERATE	WEAK	
		1	2	3	Not Applicable

GLOBAL RATING FOR THIS PAPER (circle one):
1 STRONG (no WEAK ratings)
2 MODERATE (one WEAK rating)
3 WEAK (two or more WEAK ratings)

With both reviewers discussing the ratings:

Is there a discrepancy between the two reviewers with respect to the component (A-F) ratings?
No Yes

If yes, indicate the reason for the discrepancy
1 Oversight
2 Differences in interpretation of criteria
3 Differences in interpretation of study

Final decision of both reviewers (circle one):
1 STRONG
2 MODERATE
3 WEAK

B: Cog errors, Sa anxiety, resp to CBT.

QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES

EPHPP — Effective Public Health Practice Project

COMPONENT RATINGS

A) SELECTION BIAS

(Q1) Are the individuals selected to participate in the study likely to be representative of the target population?
- 1 Very likely
- 2 Somewhat likely
- 3 Not likely
- 4 Can't tell

(Q2) What percentage of selected individuals agreed to participate?
- 1 80 - 100% agreement
- 2 60 - 79% agreement
- 3 less than 60% agreement
- 4 Not applicable
- 5 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

B) STUDY DESIGN

Indicate the study design
- 1 Randomized controlled trial
- 2 Controlled clinical trial
- 3 Cohort analytic (two group pre + post)
- 4 Case-control
- 5 Cohort (one group pre + post (before and after))
- 6 Interrupted time series
- 7 Other specify _____
- 8 Can't tell

Was the study described as randomized? If NO, go to Component C.
- No Yes

If Yes, was the method of randomization described? (See dictionary)
- No Yes

If Yes, was the method appropriate? (See dictionary)
- No Yes

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

C) CONFOUNDERS

(Q1) Were there important differences between groups prior to the intervention?
- 1 Yes
- 2 No
- 3 Can't tell

The following are examples of confounders:
- 1 Race
- 2 Sex
- 3 Marital status/family
- 4 Age
- 5 SES (income or class)
- 6 Education
- 7 Health status
- 8 Pre-intervention score on outcome measure

(Q2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design (e.g. stratification, matching) or analysis)?
- 1 80 - 100% (most)
- 2 60 - 79% (some)
- 3 Less than 60% (few or none)
- 4 Can't Tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

D) BLINDING

(Q1) Was (were) the outcome assessor(s) aware of the intervention or exposure status of participants?
- 1 Yes
- 2 No
- 3 Can't tell

(Q2) Were the study participants aware of the research question?
- 1 Yes
- 2 No
- 3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

E) DATA COLLECTION METHODS

(Q1) Were data collection tools shown to be valid?
- 1 Yes
- 2 No
- 3 Can't tell

(Q2) Were data collection tools shown to be reliable?
- 1 Yes
- 2 No
- 3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

F) WITHDRAWALS AND DROP-OUTS

(Q1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?
- 1 Yes
- 2 No
- 3 Can't tell
- 4 Not Applicable (i.e. one time survey or interview)

1/65

(Q2) Indicate the percentage of participants completing the study. (If the percentage differs by groups, record the lowest.)
- 1 80 - 100%
- 2 60 - 79%
- 3 less than 60%
- 4 Can't tell
- 5 Not Applicable (i.e. Retrospective case-control)

RATE THIS SECTION	STRONG	MODERATE	WEAK	
See dictionary	1	2	3	Not Applicable

G) INTERVENTION INTEGRITY

(Q1) What percentage of participants received the allocated intervention or exposure of interest?
- 1 80 - 100%
- 2 60 - 79%
- 3 less than 60%
- 4 Can't tell

65/208 = 31%

(Q2) Was the consistency of the intervention measured?
- 1 Yes
- 2 No
- 3 Can't tell

'similar'

(Q3) Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results?
- 1 Yes
- 2 No
- 3 Can't tell

H) ANALYSES

(Q1) Indicate the unit of allocation (circle one)
community organization/institution practice/office individual

(Q2) Indicate the unit of analysis (circle one)
community organization/institution practice/office individual

(Q3) Are the statistical methods appropriate for the study design?
- 1 Yes
- 2 No
- 3 Can't tell

(Q4) Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?
- 1 Yes
- 2 No
- 3 Can't tell

GLOBAL RATING

COMPONENT RATINGS
Please transcribe the information from the gray boxes on pages 1-4 onto this page. See dictionary on how to rate this section.

		STRONG	MODERATE	WEAK	
A	SELECTION BIAS	1	2	3	
B	STUDY DESIGN	1	2	3	
C	CONFOUNDERS	1	2	3	
D	BLINDING	1	2	3	
E	DATA COLLECTION METHOD	1	2	3	
F	WITHDRAWALS AND DROPOUTS	1	2	3	Not Applicable

GLOBAL RATING FOR THIS PAPER (circle one):
- 1 STRONG (no WEAK ratings)
- 2 MODERATE (one WEAK rating)
- 3 WEAK (two or more WEAK ratings)

With both reviewers discussing the ratings:

Is there a discrepancy between the two reviewers with respect to the component (A-F) ratings?
- No Yes

If yes, indicate the reason for the discrepancy
- 1 Oversight
- 2 Differences in interpretation of criteria
- 3 Differences in interpretation of study

Final decision of both reviewers (circle one)
- 1 STRONG
- 2 MODERATE
- 3 WEAK

14: Heterogeneity of late-life dep

QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES

EPHPP
Effective Public Health Practice Project

COMPONENT RATINGS

A) SELECTION BIAS

B) STUDY DESIGN

C) CONFOUNDERS

D) BLINDING

E) DATA COLLECTION METHODS

F) WITHDRAWALS AND DROP-OUTS

G) INTERVENTION INTEGRITY

H) ANALYSES

GLOBAL RATING

		STRONG	MODERATE	WEAK	
A	SELECTION BIAS		⃝	3	
B	STUDY DESIGN	STRONG	⃝	WEAK	
C	CONFOUNDERS	STRONG	2	WEAK	
D	BLINDING	⃝	MODERATE	WEAK	
E	DATA COLLECTION METHOD	1	⃝	3	
F	WITHDRAWALS AND DROPOUTS	⃝	2	3	Not Applicable

GLOBAL RATING FOR THIS PAPER (circle one)

⃝ STRONG
MODERATE
WEAK

15. Dep Sx, cog decline

QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES

EPHPP — Effective Public Health Practice Project

COMPONENT RATINGS

A) SELECTION BIAS

(Q1) Are the individuals selected to participate in the study likely to be representative of the target population?
1 Very likely
2 Somewhat likely
3 Not likely
4 Can't tell

(Q2) What percentage of selected individuals agreed to participate?
1 80 - 100% agreement
2 60 - 79% agreement
3 less than 60% agreement
4 Not applicable
5 Can't tell

298/402 = 74%

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

B) STUDY DESIGN

Indicate the study design
1 Randomized controlled trial
2 Controlled clinical trial
3 Cohort analytic (two group pre + post)
4 Case-control
5 Cohort (one group pre + post (before and after))
6 Interrupted time series
7 Other specify _____
8 Can't tell

Was the study described as randomized? If NO, go to Component C.
No   Yes

If Yes, was the method of randomization described? (See dictionary)
No   Yes

If Yes, was the method appropriate? (See dictionary)
No   Yes

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

C) CONFOUNDERS

(Q1) Were there important differences between groups prior to the intervention?
1 Yes
2 No
3 Can't tell

The following are examples of confounders:
1 Race
2 Sex
3 Marital status/family
4 Age
5 SES (income or class)
6 Education
7 Health status   *vascular.*
8 Pre-intervention score on outcome measure

(Q2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design (e.g. stratification, matching) or analysis)?
1 80 - 100% (most)
2 60 - 79% (some)
3 Less than 60% (few or none)
4 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

D) BLINDING

(Q1) Was (were) the outcome assessor(s) aware of the intervention or exposure status of participants?
1 Yes
2 No
3 Can't tell

(Q2) Were the study participants aware of the research question?
1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

E) DATA COLLECTION METHODS

(Q1) Were data collection tools shown to be valid?
1 Yes
2 No
3 Can't tell

(Q2) Were data collection tools shown to be reliable?
1 Yes
2 No
3 Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

F) WITHDRAWALS AND DROP-OUTS

(Q1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?
1 Yes
2 No
3 Can't tell
4 Not Applicable (i.e. one time survey or interview)

(Q2) Indicate the percentage of participants completing the study. (If the percentage differs by groups, record the lowest)
1 80 -100%
2 60 - 79%
3 less than 60%
4 Can't tell
5 Not Applicable (i.e. Retrospective case-control)

RATE THIS SECTION	STRONG	MODERATE	WEAK	
See dictionary	1	2	3	Not Applicable

G) INTERVENTION INTEGRITY

(Q1) What percentage of participants received the allocated intervention or exposure of interest?
1 80 -100%
2 60 - 79%
3 less than 60%
4 Can't tell

(Q2) Was the consistency of the intervention measured?
1 Yes
2 No
3 Can't tell

N/A.

(Q3) Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results?
4 Yes
5 No
6 Can't tell

H) ANALYSES

(Q1) Indicate the unit of allocation (circle one)
community   organization/institution   practice/office   individual

(Q2) Indicate the unit of analysis (circle one)
community   organization/institution   practice/office   individual

(Q3) Are the statistical methods appropriate for the study design?
1 Yes
2 No
3 Can't tell

(Q4) Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?
1 Yes
2 No
3 Can't tell

GLOBAL RATING

COMPONENT RATINGS
Please transcribe the information from the grey boxes on pages 1-4 onto this page. See dictionary on how to rate this section.

		STRONG	MODERATE	WEAK
A	SELECTION BIAS	1	2	3
B	STUDY DESIGN	1	2	3
C	CONFOUNDERS	1	2	3
D	BLINDING	1	2	3
E	DATA COLLECTION METHOD	1	2	3
F	WITHDRAWALS AND DROPOUTS	1	2	3 Not Applicable

GLOBAL RATING FOR THIS PAPER (circle one)
1 STRONG (no WEAK ratings)
2 MODERATE (one WEAK rating)
3 WEAK (two or more WEAK ratings)

With both reviewers discussing the ratings:

Is there a discrepancy between the two reviewers with respect to the component (A-F) ratings?
No   Yes

If yes, indicate the reason for the discrepancy
1 Oversight
2 Differences in interpretation of criteria
3 Differences in interpretation of study

Final decision of both reviewers (circle one)
1 STRONG
2 MODERATE
3 WEAK

16: Large regional white matter

QUALITY ASSESSMENT TOOL FOR QUANTITATIVE STUDIES

EPHPP
Effective Public Health Practice Project

COMPONENT RATINGS

A) SELECTION BIAS

(Q1) Are the individuals selected to participate in the study likely to be representative of the target population?
1. Very likely
2. Somewhat likely
3. Not likely
4. Can't tell

(Q2) What percentage of selected individuals agreed to participate?
1. 80 - 100% agreement
2. 60 - 79% agreement
3. less than 60% agreement
4. Not applicable
5. Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

B) STUDY DESIGN

Indicate the study design
1. Randomized controlled trial
2. Controlled clinical trial
3. Cohort analytic (two group pre + post)
4. Case-control
5. Cohort (one group pre + post (before and after))
6. Interrupted time series
7. Other specify _____
8. Can't tell

Was the study described as randomized? If NO, go to Component C.
No   Yes

If Yes, was the method of randomization described? (See dictionary)
No   Yes

If Yes, was the method appropriate? (See dictionary)
No   Yes

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

C) CONFOUNDERS

(Q1) Were there important differences between groups prior to the intervention?
1. Yes
2. No
3. Can't tell

The following are examples of confounders:
1. Race
2. Sex
3. Marital status/family
4. Age
5. SES (income or class)
6. Education
7. Health status
8. Pre-intervention score on outcome measure

(Q2) If yes, indicate the percentage of relevant confounders that were controlled (either in the design (e.g. stratification, matching) or analysis)?
1. 80 - 100% (most)
2. 60 - 79% (some)
3. Less than 60% (few or none)
4. Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

D) BLINDING

(Q1) Was (were) the outcome assessor(s) aware of the intervention or exposure status of participants?
1. Yes
2. No
3. Can't tell

(Q2) Were the study participants aware of the research question?
1. Yes
2. No
3. Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

E) DATA COLLECTION METHODS

(Q1) Were data collection tools shown to be valid?
1. Yes
2. No
3. Can't tell

(Q2) Were data collection tools shown to be reliable?
1. Yes
2. No
3. Can't tell

RATE THIS SECTION	STRONG	MODERATE	WEAK
See dictionary	1	2	3

F) WITHDRAWALS AND DROP-OUTS

(Q1) Were withdrawals and drop-outs reported in terms of numbers and/or reasons per group?
1. Yes
2. No
3. Can't tell
4. Not Applicable (i.e. one time surveys or interviews)

(Q2) Indicate the percentage of participants completing the study. (If the percentage differs by groups, record the lowest)
1. 80 -100%
2. 60 - 79%
3. less than 60%
4. Can't tell
5. Not Applicable (i.e. Retrospective case-control)

RATE THIS SECTION	STRONG	MODERATE	WEAK	
See dictionary	1	2	3	Not Applicable

G) INTERVENTION INTEGRITY

(Q1) What percentage of participants received the allocated intervention or exposure of interest?
1. 80 -100%
2. 60 - 79%
3. less than 60%
4. Can't tell

(Q2) Was the consistency of the intervention measured?
1. Yes
2. No
3. Can't tell

(Q3) Is it likely that subjects received an unintended intervention (contamination or co-intervention) that may influence the results?
1. Yes
2. No
3. Can't tell

H) ANALYSES

(Q1) Indicate the unit of allocation (circle one)
community   organization   group   individual   practice/office

(Q2) Indicate the unit of analysis (circle one)
community   organization   group   individual   practice/office

(Q3) Are the statistical methods appropriate for the study design?
1. Yes
2. No
3. Can't tell

(Q4) Is the analysis performed by intervention allocation status (i.e. intention to treat) rather than the actual intervention received?
1. Yes
2. No
3. Can't tell

GLOBAL RATING

COMPONENT RATINGS
Please transcribe the information from the gray boxes on pages 1-4 onto this page. See dictionary on how to rate this section.

		STRONG	MODERATE	WEAK
A	SELECTION BIAS	1	2	3
B	STUDY DESIGN	1	2	3
C	CONFOUNDERS	1	2	3
D	BLINDING	1	2	3
E	DATA COLLECTION METHOD	1	2	3
F	WITHDRAWALS AND DROPOUTS	1	2	3	Not Applicable

GLOBAL RATING FOR THIS PAPER (circle one)
1. STRONG (no WEAK ratings)
2. MODERATE (one WEAK rating)
3. WEAK (two or more WEAK ratings)

With both reviewers discussing the ratings:

Is there a discrepancy between the two reviewers with respect to the component (A-F) ratings?
No   Yes

If yes, indicate the reason for the discrepancy
1. Oversight
2. Difference in interpretation of criteria
3. Difference in interpretation of study

Final decision of both reviewers (circle one)
1. STRONG
2. MODERATE
3. WEAK

Appendix C: Written Descriptions of Results of Individual Studies

1: Dumas and Newhouse (2015) compared *n*-back task performance between 11 older adults with a diagnosis of current depression to 12 healthy controls who had never had depression. Participants were scanned using functional magnetic resonance imaging (fMRI) while undertaking the task. Four depressed participants were taking anti-depressant medication, while none of the controls were, therefore the possible impact of medication on task performance cannot be predicted. The study found that depressed participants performance decreased more compared to healthy controls as WM load increased, and that depressed participants were more conservative on 0-back trials. fMRI data related this to decreased activation of the bilateral frontal lobe and precuneus, which the authors suggest indicates a cognitive element to depression in OAs.

2: Koenig and colleagues (2015) have extended findings such as these by delineating between currently and previously depressed participants, alongside healthy controls. Their study, using the digit symbol substitution test (DSST) within a large battery of neuropsychological tests, found that individuals currently or previously presenting with depression performed significantly worse on the DSST than never-depressed participants. The lack of differentiation between previously and currently depressed participants on the measures leads the authors to conclude that cognitive impairments related to depression are trait variables.

3: Shimada and colleagues (2014) also used a three-group strategy to explore the relationship between mood and cognition in OAs. They distinguished individuals who showed some depressive symptoms (6+ on the

GDS-15) from those who received a diagnosis of depression (from a doctor).

These two groups were compared to individuals with no depressive symptoms

(<6 on the GDS-15) on a neuropsychological battery (NCGG-FAT). Analysis

using ANOVA indicated a significant difference in SDST performance, which

post-hoc analysis indicated related to a difference between individuals with

depression and those without any symptoms. Although the study used a well-

validated measure to distinguish between the groups, it is not clear that the

points of cleavage allowed for sufficient distinction between group participants.

4: Baune et al. (2006) used participants from a larger scale study of

morbidity in OAs to explore the relationship between depression and cognition.

The researchers used a CES-D cut-off of 16 to compare individuals with and

without clinically significant levels of depression on a number of

neuropsychological tests, including the Letter Digit Substitution Test (LDST).

There was no significant difference. Again, this may relate to insufficient

differentiation between 'depressed' and 'non-depressed' participants, as

selection was only based on a self-report scale.

5: Sair and colleagues (2006) also compared depressed and non-

depressed participants on a measure of working memory. They divided

participants according to the NIMH Diagnostic Interview Schedule as well as

self-rating scale scores. They found that depressed participants performed

more poorly on the Ascending Digits Task (ADT) than non-depressed

counterparts, even when age, gender, race and education were controlled for.

However, they found no relationship between severity of depression and ADT

performance. The authors suggest that depression may be particularly related

to tasks in which information in WM must be manipulation, requiring executive

input.

   6 and 9: In baseline measures for a longitudinal treatment study, ***Nebes***

and colleagues (2003 & 2006) compared depressed participants with healthy

controls, matched for age and education, on the *n*-back and digit symbol

substitution (DSST) tests. Performance on both tasks was significantly lower in

depressed compared to control participants across both studies. As treatment

studies, the authors went on to track changes in task performance with

commencement of anti-depressant medication (paroxetine or nortriptyline).

There was no change in WM performance over and above that expected by

practice effects, indicating that cognitive impairment associated with geriatric

depression may not be influenced by depression treatment and may require

additional intervention. Nebes et al. (2000) used regression analyses to

demonstrate that a high proportion of variance on neuropsychological tasks

accorded to depression could in fact be best explained by cognitive processing

abilities.

   7: Elderkin-Thompson and colleagues (2003) used a three group

design to compare working memory and other cognitive abilities in OAs – minor

depressed, major depressed and healthy controls. Minor depressed participants

differed from major depressed participants on the number of DSM-IV symptoms

they experienced (depression/adhedonia with between 1-4 additional

symptoms), and their HRSD scores (between 8-16). Minor depressed

participants were and controls were recruited from the community, while

depressed participants were from clinics. The study combined digit span

forward and backward to a single 'WM' factor, which was then compared across

the three groups. There was no significant difference across groups, although post-hoc pair-wise comparisons indicated a trend towards significance when controls and major depressed participants were compared (p=.08). There was no correlation between severity of depression and WM scores in the total sample, although when women were analysed separately it was significant (-.312, p=.04).

8: Brodaty and colleagues (2001) focused on a comparison between early- (before 60years age) and late-onset depression on a battery of neuropsychological measures, including the symbol-digit modalities test (SDMT). No differences were found in test performance between the two groups, and both contained a similar proportion of impaired performers (2 SDs below mean; around 20-40% across tests). They suggest that, while poor cognition may have a causative role in late-onset depression, it may be a *result* of long-term depression in individuals with early-onset depression. So while both groups demonstrate impaired performance, the pathways to this are distinct.

10: Salazar-Villanea and colleagues (2015) sampled participants from Costa Rica, constructing structural equation models (SEM) of the relationships between affect, depression and cognitive performance. They demonstrated a small (-0.21) but significant correlation between Geriatric Depression Scale scores and digit span backwards. Although depression symptomatology correlated with WM performance, negative affect – as measure by the PANAS – did not. The full SEM demonstrated that increased depression significantly predicted a decreased WM capacity, which the authors suggest related to the size of WM capacity, not the speed.

11: Potter and colleagues (2015) investigated cognitive performance in adults with a diagnosis of late-onset MDD. Ninety one percent of participants were taking medication to manage their mood. A factor analysis of the HRSD-17 and MADRS resulted in five factors to describe groups of symptoms relevant to depression. Only one of these factors (appetite loss) was significantly correlated with WM performance (digit span) (r= -0.12, p<.04). A global depression score did not significantly associate with WM.

12: Delaloye and colleagues (2010) compared participants with early- and late-onset depression to healthy controls, using reading span and letter-number sequencing tasks as measures of WM. Neither task effectively differentiated between the three groups, despite a prediction of impaired cognitive performance in adults with late-onset depression. MRI scanning indicated an increase in periventricular white matter hyperintensities in this group, lending support to the vascular origin of late-onset depression.

13: Caudle and collegues (2007) provide the only study identified within the review to focus specifically on anxiety, taking as their sample participants with generalised anxiety disorder (GAD). Participants who made errors within the WM domain of the Mini-Mental State Exam (MMSE) were shown to have higher levels of GAD severity, HRSD and Beck Depression Inventory scores, i.e. has increased levels of both depression and anxiety. WM errors had no impact on response to CBT. As a small subscale of a larger screening test, the WM subscale may not be the most reliable or valid measure of the construct of WM.

14: Korten and colleagues (2014) compared depressed participants with non-depressed controls on a range of cognitive measures, including digit

span task. Here was no significant difference between the groups in WM performance, and only a very modest relationship between the digit span forward scores and depression severity (t(351)=-2.26, p,.03). Post-hoc analyses also indicated a relationship between WM performance and anxiety severity, but not worrying. When symptom dimensions of depression were analysed separately, WM was associated with mood and somatic aspects of depression. The authors also highlight that participants taking tricyclic anti-depressants demonstrated worse performance on cognitive testing than their counter-parts taking selective serotonin reuptake inhibitors, suggesting that SSRIs be the treatment of choice for OAs.

15: Turner and colleagues (2015) identified that a number of studies have utilised primarily white participants. They explored the relationship between depression (using the GDS and CES-D) and WM (digit span forward and backward, and digit ordering) at baseline, and over an average of 5years follow up. There were no significant relationships between WM and measures of mood at baseline. However, higher GDS scores were associated with faster decline in WM over follow up. In exploration of association between GDS factors and cognitive abilities, negative affect was significantly associated with fast decline in WM. The longitudinal design of this study lends support to the idea of depression precipitating WM impairment.

16: Yuan and colleagues (2009) compared performance on a neuropsychological battery, including digit span and Symbol Digit Modalities (SDMT) tests, between participants with remitted depression and healthy controls. Performance on SDMT was significantly poorer for depressed participants than healthy controls, while there was no significant difference in

digit span performance. An imaging aspect to the study also found that

depressed participants had increased white-matter volumes of left inferior

parietal lobule and right inferior frontal gyrus.

SCHOOL OF PSYCHOLOGY

DOCTORATE IN CLINICAL PSYCHOLOGY

EMPIRICAL PAPER

The Role of a Working Memory Training Program in Reducing Repetitive

Negative Thinking in Older Adults.

Trainee Name: **Jodie Rawlings**

Primary Research Supervisor: **Dr Jenny Limond**

Senior Lecturer and Co-Director of Research

Secondary Research Supervisor: **Dr Philip Yates**

Research Tutor

Target Journal: Cognition and Emotion

Word Count: 7973

Submitted in partial fulfilment of requirements for the Doctorate Degree in

Clinical Psychology, University of Exeter

Acknowledgements

Abstract

Objective: Recent research has supported the construction of a model of depression and anxiety in older adults (>60years), in which the decline in working memory with age is implicated in increasing propensity to engage in repetitive negative thinking, thereby predisposing older adults to anxiety and depression. The study detailed here intended to extend these research findings by exploring the efficacy and acceptability of a working memory intervention in older adults.

Methods: The study employed a randomised multiple-baseline single case research design, in which six older adults participated in baseline assessment of mood, repetitive negative thinking and working memory. Participants then tracked their repetitive negative thinking daily throughout a baseline phase, the length of which was defined by the randomisation procedure. Following phase change, participants continued to rate their repetitive thinking daily, as well as engaging in a daily working memory training intervention. On completion of the intervention phase, outcome measures of mood, repetitive negative thinking and working memory were repeated. Full visual and statistical analysis of all data was undertaken to support exploration of the findings.

Results: Results of the study indicated that that the working memory training program was unlikely to have elicited a significant impact upon participants' working memory. However, two participants demonstrated reliable improvement in both repetitive thinking and mood. Analysis of relevant variables to predict the selective impact of the intervention was not fruitful, but may indicate that improvements in working memory underscore the improvements in

thinking and mood, lending support to the proposed model. There appeared to be a small, non-significant decrease in daily repetitive negative thinking across five of the six participants.

Conclusion: Further research is needed to identify factors that may predict response to working memory training within older adult populations. The research supports the on-going investigation of innovative working memory interventions within an older adult population, although results are not sufficiently robust to indicate wider adoption of these models within health services or as routine treatments for this population.

Keywords: older adults, working memory, repetitive negative thinking, depression, anxiety.

Introduction

Working Memory

Working memory (WM) can be defined as the process of "actively maintaining… or manipulating information across a short delay" (Snyder, 2013, p.82). WM is considered to be an essential factor in a range of complex cognitive tasks (Baddeley, 1997). Much attention has therefore been paid to methods of measuring WM within individuals. Many of the measures most commonly used represent WM capacity tasks, such as the digit span forwards, in which individuals are asked to recall a set of information across a short delay (Conway et al., 2005). WM capacity tasks have been shown to have good reliability (Oberauer, Süß, Schulze, Wilhelm, & Wittmann, 2000) and validity (Kane & Engel, 2000). However, there is increasing evidence that effective WM requires not only a capacity to store a volume of information, but also to actively manipulate this information according to need, ('updating'; Miyake, Friedman, Emerson, Witzki, & Howerter, 2000). Tasks such as the *n*-back[6] are considered to be effective measures of this concept (Wilhelm, Hildebrandt, & Oberauer, 2013).

Working memory and mood.

Channon and colleagues (1993) found poorer performance in individuals with depression compared to healthy controls on two tasks of WM that required input of the central executive. More recent research indicates that depressed individuals are likely to have impairment across all elements of WM (Christopher & MacDonald, 2005). A thorough meta-analysis from Snyder

---

[6] In the *n*-back task, the participant is requested to remember the stimulus *n* steps back in a task.

(2013) demonstrated impairments for depressed patients on both visuospatial and verbal WM tasks, with reported composite small-medium effect sizes for both at d = .45. In a sample of individuals with chronic depression, Harvey and colleagues (2004) reported positive correlations between performance on the n-back task and both the number of hospitalisations and longitudinal course of depression. Imaging studies, using event-related potential, have demonstrated that biases within WM can be identified in individuals with a diagnosis of depression, but not healthy controls (Deldin, Deveney, Kim, & Webb, 2001). The relationship between low mood and WM is maintained even at non-clinical levels, with both Owens, Koster and Derakshan (2012) and Hubbard et al. (2016) finding low WM capacity in dysphoric compared to non-dysphoric participants.

Evidence also indicates both a decreased WM capacity and decreased ability to update information within WM for highly anxious participants (Darke, 1988). There appears to be a specific relationship between anxiety and visuospatial WM, with studies suggesting an association between increasing anxiety and worsening spatial WM performance (Shackman et al., 2006).

Working memory in older adults.

Research indicates that WM declines with increasing age (Swanson, 1999). Dobbs and Rule (1989) administered a battery of WM tasks to participants across the ages of 30-99 years, and found a significant decline in WM updating ability between the ages 60-69 years, and 70+ years. While this study demonstrated stable WM capacity throughout the lifetime, other studies

have noted a decline in capacity with aging (Babcock & Salthouse, 1990).

Studies have indicated that active manipulation of information in WM appears

particularly impaired in older as compared to younger samples (Vecchi &

Cornoldi, 1999; Hester, Kinsella, & Ong, 2004). These results can be seen as

evidence for Hasher and Zacks' (1988) inhibition-deficit hypothesis, which

states that older adults (OAs) are impaired in the inhibition of irrelevant

information, such that their WM becomes full of information they are unable to

remove.

Repetitive Negative Thinking

Repetitive thinking has been defined as the "process of thinking

attentively, repetitively or frequently about one's self and one's world,"

(Segerstrom, Roach, Evans, Schipper, & Darville, 2003, p. 909). When this

thinking style is negative in content or orientation, it is described as repetitive

negative thinking (RNT). Increasingly, RNT is seen as a transdiagnostic

process, related to both the onset and maintenance of a number of Axis I

disorders (Ehring & Watkins, 2008). A recent investigation of the RNT construct

using structural equation modelling (Topper, Molenaar, Emmelkamp, & Ehring,

2014), has identified RNT as a unitary factor, with sub-factors of worry and

rumination. These two factors may be best differentiated by their temporal

focus, with rumination considered as past-oriented, and worry as future-oriented

(Nolen-Hoeksema, Wisco, & Lyubomirsky, 2008). Although much of the

literature has distinguished RNT into these separate constructs, Topper and

colleagues (2014) suggest that, given its unitary structure, RNT is best

assessed using generic measures.

Repetitive negative thinking and mood.

Evidence suggests that RNT is particularly relevant to both depression and anxiety. Nolen-Hoeksema's (1991) response styles theory (RST) of depression suggests that rumination contributes to the maintenance and exacerbation of depression by increasing negative thinking, limiting ability to problem solve or engage in instrumental behaviour and eroding social support. High scores on the Ruminative Responses Scale (RRS; Treynor, Gonzales, & Nolen-Hoeksema, 2003) are predictive of future depression, chronicity of depression, anxiety (Nolen-Hoeksema, 2000), and length of depressive episodes (Nolen-Hoeksema, Morrow, & Fredrickson, 1993). As with rumination, worry is predictive of both depression and anxiety (Hoyer, Gloster, & Herzberg, 2009). Worry is also a key tenet of generalised anxiety disorder (GAD; Wells, 2006).

Repetitive negative thinking in older adults.

While older adults (OAs) have been found to ruminate less than their younger counterparts (Sütterlin, Paap, Babic, Kübler, & Vögele, 2012), the relationship between rumination and depression has been replicated in this population (Ricarte, Ros, Serrano, Martinez-Lorca, & Latorre, 2016). Segerstrom and colleagues (2010) found that the valence of repetitive thought in OAs influenced its relationship with outcomes; RNT was associated with poorer psychological, physical and cognitive health, while positive repetitive thought was related to improved psychological wellbeing. There is suggestion by the authors that poorer cognitive function predisposes OAs to more negatively-valenced repetitive thought. However, the sample for this study comprised only married individuals, with few health problems, and is therefore

not representative of the wider OA population.

The Association between Working Memory, Repetitive Negative Thinking and Mood

Working memory, RNT and depression.

There is increasing evidence that the relationship between WM and depression may be mediated by RNT (Onraedt & Koster, 2014). Joorman and Gotlib (2008) explored performance on a WM updating task across three groups; participants with a diagnosis of depression, healthy controls following a sad mood induction, and healthy controls following a neutral exercise. The authors found that depressed participants had significantly higher intrusion effects than either control group, indicating poorer WM updating, and that these intrusion effects correlated significantly with self-report measures of rumination (r = .49, p<.05). Unfortunately, the cross-sectional nature of this study limits the ability to draw conclusions as to the causal nature of these relationships.

Gotlib and Joorman (2010) suggest that deficits in WM may lead to mood disturbance by "setting the stage for ruminative responses to negative events" (p.13). Poor WM may cause individuals to become stuck with RNTs they are unable to take 'offline'. Koster and colleagues (2011) have termed this the 'impaired disengagement' model of rumination; highlighting difficulties updating negative information in the WM store. Joorman and Tran (2009) have found that individuals with depression have difficultly supressing negative words in directed forgetting tasks. Joorman and colleagues (2011) extend findings such as these by linking difficulties in information-sorting WM tasks with self-reported

rumination. There is even evidence to suggest that filling WM with an alternative task prevents a negative mood elicitation exercise from having its intended effect, with authors of this study suggesting that this is because RNT is not able to occur (Van Dillen, & Koole, 2007). Kircanski and colleagues (2012) have suggested that alongside difficulties removing information from WM, individuals may be more prone to negative information entering WM. It is therefore possible to theorise a pathway in which ineffective WM management leads to RNT, which may in turn result in low mood or depression. It is relevant to note that both rumination and negative emotion are likely to feedback to reduce WM capacity (Curci, Lanciano, Soleti, & Rimé, 2013).

Working memory, RNT and depression in older adults.

Von Hippel and colleagues (2008) compared older adults (OAs) with late (>60yrs at first episode) and early (<60yrs) onset experiences of depression, finding that OA with late-onset depression were more impaired on tasks of executive function (EF) – including a simple task of WM – than their early-onset peers. Furthermore, the performance of OA with late-onset depression demonstrated negative correlations between EF measures and rumination scores, which were absent for early-onset OA. The authors take these results to support an extension of the WM-RNT-depression model described above. They suggest that age-related cognitive decline (Hasher & Zacks, 1988), in tandem with increasing likelihood of negative life events such as bereavement, leads to increased rumination, therefore increasing susceptibility to depression. While this study is limited by a small sample size and self-report of depression onset, this provides some evidence for the relationships between WM, RNT and depression.

Further research has supported elements of this model in OAs, with replication of findings of increased neuropsychological EF impairment in late- compared to early-onset depression in OA (Thomas et al., 2009; Herrmann, Goodwin, & Ebmeier, 2007). Ricarte and colleagues (2016) have provided support for this reasoning, with their experiment indicating that OA, but not younger adults, demonstrate a relationship between WM performance and depression, mediated by 'brooding' rumination. Philippot & Agrigoroaei (2016) indicate that impairment in executive functions such as WM may particularly instigate OAs to engage in abstract (rather than concrete) RNT, which is in turn more closely correlated with depression.

It appears that the relationship between WM and RNT in OAs holds even where depression is mild (Pantzar et al., 2014) or does not reach clinical levels (von Hippel, Vasey, Gonda, & Stern, 2008). Sexton and colleagues (2012) have used magnetic resonance imaging (MRI) to identify neural changes relevant to EF deficits in late-life depression. However, a key limitation of studies in this area is their cross-sectional design, and therefore inability to elucidate causal relationships.

Working memory, RNT and anxiety.

The research literature directly exploring the inter-relationships between WM, RNT and anxiety is somewhat scarcer. However, a recent meta-analysis found that anxiety was related to poorer WM performance across a wide range of WM tasks, with the author suggesting a mediational role for worry/RNT in this

relationship (Moran, 2016).

Working memory, RNT and anxiety in older adults.

There is emerging evidence of the validity of this model within an older adult (OA) population. Butters and colleagues (2011) compared OA with generalised anxiety disorder (GAD) to healthy OA controls using the digit span task, finding poorer WM in the GAD group. Interestingly, following pharmacological intervention (escitalopram), participants demonstrated small but significant improvements in WM alongside clinical improvements. The lack of a depression control group prevents analysis of the similarities and differences in WM impairment between these two clinical groups.

Computerised Working Memory Intervention

The above literature has led to investigation of WM as an intervention target to decrease both RNT and depression. Using the *n*-back task as a training paradigm over 8 days, Owens and colleagues (2013) were able to detect improvements in WM capacity using both behavioural and neuroimaging measures in dysphoric participants. Schweizer and colleagues (2013) have evidenced that improvement in WM capacity following 20 days WM training relates to increases in affective cognitive control. Similarly, Iacoviello and colleagues (2014) reported 6 of 11 participants with depression showed a greater than 50% decrease in symptoms following a cognitive-emotional WM training program. These findings lend support to the causal role of WM in affective disturbance. However, it should be noted that studies in this field have

been limited in a number of ways, not least by small sample sizes (e.g. n=21 split across experimental and control groups in Iacoviello et al., 2014). At the same time, the validity of control groups used within these studies has been questioned. Many employ a 'non-adaptive' (Owens et al., 2013) or 'non-affective' (Iacoviello et al., 2014) control training task, which has not been independently investigated and evaluated for effects. This impedes researchers' ability to identify the root cause of any gains seen.

There is evidence to suggest that the positive impact of WM training may be via a reduction in RNT. Quinn and colleagues (2014) demonstrated that trait rumination moderated the impact of executive control training, such that only 'high' ruminators decreased in stress response following WM intervention, with 'low' ruminators showing no intervention effect. Hoorelbeke and colleagues (2015) found that, in students with high levels of trait rumination, 10 sessions of cognitive control training were effective in decreasing rumination in response to a naturalistic stressor at a four-week follow-up.

A number of studies have, however, failed to find positive effects of WM-based training interventions (Wanmaker, Geraerts, & Franken, 2015). Onraedt and Koster (2014) found no evidence of improvement in WM, rumination or depression following a six-day WM intervention. They suggest that, to best demonstrate efficacy, training programs should occur for at least 17 days. A meta-analysis on the subject concluded that gains in WM were often not maintained at follow-up, and that transfer effects to alternative tasks were rarely seen (Melby-Lervåg & Hulme, 2013). The evidence for a wider psychological impact is also not clear, with Wanmaker and colleagues (2014) demonstrating increases in WM capacity with training, but no impact on psychopathology,

including rumination. It is clear that further work is needed to better understand the most effective form of WM training, and the possible validity of this form of intervention.

Computerised cognitive training for older adults.

There is evidence that common interventions for depression have reduced efficacy for older adults (OAs; Frazer, Christensen, & Griffiths, 2005). Impaired EF may be the cause of this, acting as a barrier to benefitting from typical treatments, (CBT: Pimontel, Culand-Reinlieb, Morimoto, & Sneed, 2011; pharmacological intervention: Sneed et al., 2007). Bogner and colleagues (2007) described low remission rates for OA with executive dysfunction receiving normal treatment plans. As such, training programs to increase EF and WM abilities may represent a valid treatment route or adjunct for OA with depression (von Hippel et al., 2008) and/or anxiety (Mohlman & Gorman, 2005).

Lampit and colleagues' (2014) meta-analysis of the literature on computerised cognitive training in OAs suggests small effect sizes for such intervention of around g=0.22 (95% CIs 0.09-0.35). Reijnders and colleagues (2013) conclude their SR of the field by suggesting there is evidence that cognitive training programs can be effective for OA, but that it is difficult to identify transfer effects. Another SR from Kueider and colleagues (2012), concludes that computerised cognitive training is an effective method of delivering interventions as well as being less labour intensive for researchers/clinicians. Recent evidence suggests that these programs have an equal efficacy to pharmacological treatment (escitalopram; Morimoto et al., 2014).

Research has begun to focus on the specific factors that might enhance efficacy of WM training programs for OAs. McAvinue and colleagues (2013) found improvements in WM only for OAs undertaking an adaptive, compared to non-adaptive, training protocol. Toril and colleagues (2016) have also demonstrated promising results using a 'video game' interface, highlighting a possible factor of engagement in facilitating positive gains. Further evidence suggests that training can be targeted to specific individuals likely to benefit, with Zinke and colleagues (2014) able to predict outcome of a WM training program on the basis of baseline performance and age. Researchers also suggest that sessions should be over 30 minutes duration and should take place outside of the participant's home (Lampit et al., 2014).

Rationale and Aim of Current Study

As the literature review above has demonstrated, there is increasing theoretical and experimental evidence of a relationship between WM, RNT and mood. A robust theoretical model has been proposed that indicates that the tendency toward decline in WM with increasing age may predispose older adults (OAs) to RNT, and consequentially, depression and anxiety. Given low response rates to classic interventions for mood in this population, it is reasonable to suggest that interventions that target improvements in WM may have a positive impact on low mood within an OA population.

This study therefore aims to explore the impact and acceptability of a WM training intervention in OAs, tracking the effect of the program on WM capacity in both training and transfer tasks along with bearing on RNT and

mood. As such, the study explores a potential new area of research in the support of OA with mood difficulties.

Research Questions

Primary research questions.

– Is a working memory training program (WMTP) effective in training WM in OAs?

– Does a WMTP reduce repetitive negative thinking in OAs?

Secondary research questions.

– Is the WMTP successful in reducing depression and anxiety in OAs?

– Is the WMTP acceptable and feasible for this population?

Research Hypotheses

H1: Completion of an affective WMTP will result in transfer effects to non-training measure of affective WM (modified Sternberg).

H2: Completion of an affective WMTP will elicit a reduction in state RNT as measured by a self-report questionnaire (Perseverative Thinking Questionnaire).

H3: Depression and anxiety (Hospital Anxiety and Depression Scale) will decrease following completion of an affective WMTP.

Method

Design

The study employed a multiple-baseline single-case experimental design (SCED) to evaluate the impact of an affective working memory training program (WMTP) on RNT in older adults (OAs). Single-case research is an effective and efficient method for inferring causality when sample size is small (Nock, Michel, & Photos, 2007), while multiple baselines allow changes in outcome measure to be causally related to the introduction of the intervention (Kazdin & Kopel, 1975; Watson & Workman, 1981). For these reasons, SCED is frequently used in neuropsychological intervention research (Crawford & Garthwaite, 2012). The Oxford Centre for Evidence-Based Medicine (2016) has included randomised n-of-1 trials as 'Level 1' evidence in making treatment decisions. The quality of the study design of these trials can be assessed using standardised scales, such as the Risk of Bias in N-of-1 Trials (RoBiN-T; Tate, et al, 2013). Applying the RoBiN-T scale to this study design suggests a design of moderate quality (Appendix A).

The study also included pre-/post-intervention measures of WM, depression and anxiety. This enables basic analysis of changes in these variables related to intervention.

Participants

Sample characteristics.

Participants for the study were recruited through a partnership between the University of Exeter and the University of the Third Age (U3A). The U3A is

an organisation for (semi-)retired individuals to co-operate in shared learning

opportunities (University of the Third Age, 2015). U3A members represent a

particular subset of the OA population, as they are heavily involved in academic

and social activities. However, the use of this non-clinical sample allowed

access to a large group of individuals who are able to participate in the study.

Parallels can be drawn with the use of student populations as an adjunct for the

general adult population; while students represent a comparatively high-

functioning group, they have been shown to demonstrate sufficiently high levels

of rumination, depression and anxiety (e.g. Roberts, Watkins, & Wills, 2013). It

is therefore anticipated that OA from the U3A constitute an appropriate

population for this proof of principle study. Participants were reimbursed £10 for

their engagement.

Six participants (aged 65 to 77 years, mean 69.83 years, SD 4.75; four

females) completed the study, including baseline and intervention phases. As

an experimental design for use with small sample sizes, and with analysis

strategies in service of this, SCED is exempt from classic power analyses to

determine sample size. The inclusion of randomisation within the design

increases the power of the study (Onghena, & Edginton, 2005), such that four

or greater subjects is entirely appropriate.

Study advert shared on both Exeter and Bath U3A's websites and newsletters.

Interested parties invited to email researcher.

Email contact from interested parties, who are then sent detailed information about the study.

(n = 25)

Interested parties responded

(n = 17)

Declined participation

(n = 6)

Reasons:
None provided (n = 3)
Too big an ask (n = 1)
Participant too young (n = 1)
No Internet access (n = 1)

Interested parties completed screening tool (Perseverative Thinking Questionnaire; PTQ)

(n = 11)

Scored below cut-off on PTQ

(n = 1)

Scored above cut-off on PTQ

(n = 10)

Unable to identify suitable period to undertake study

(n = 2)

Commenced participation in study

(n = 8)

Withdrew from study

(n = 2)

Reasons:
Study too time consuming (n = 2)

Completed study

(n = 6)

Figure 1. Participant flow chart through study.

Inclusion and exclusion criteria.

To be eligible for inclusion in the study, participants needed to be over

the age of 60 years and have access to their own home computer/laptop with

Internet access (to complete the online WMTP). Inclusion also required high

scores on a screening measure of repetitive negative thinking, the

Perseverative Thinking Questionnaire (PTQ; Ehring et al., 2011) in order to

diminish potential floor effects. The mean average score on the PTQ in a non-

clinical sample is 28.14 (SD 13.23; T. Ehring, personal communication,

November 22, 2015), suggesting that participants with scores >28 are likely to

engage in relatively high levels of repetitive negative thinking. These

participants were therefore selected for inclusion within the study.

Participants were excluded from participation if they had a diagnosis of

dementia, intellectual disability or moderate/severe head injury; or were

currently receiving treatment for a mental health condition. Participants with

high levels of suicidal ideation on the PHQ-9 were also excluded from

participation for ethical reasons. In this scenario, the centre's risk management

protocol was followed.

Ethical Considerations

Ethical approval for the study was obtained from the University of Exeter

Ethics Committee (Appendix B). The research abided by the BPS Code of

Human Research Ethics (2010). Further analysis of risks and ethical

considerations for the study is found in Appendix C.

Apparatus and Materials

Screening Measures.

Before inclusion in the study, the researcher discussed with participants their participation and administered screening measures. Participants were given the information sheet and consent form for the study (Appendix B), and had opportunity to discuss any questions. Demographic information (age, gender, occupation) and information about medical diagnoses relevant to study inclusion was collected. Standardised measures undertaken at screening included:

- *Perseverative Thinking Questionnaire (PTQ; Ehring et al., 2011).* This measure of RNT was developed in response to recent research suggesting that RNT is a transdiagnostic phenomenon, requiring a content-independent measure. It comprises 15 items, which participants rate on a 5-point Likert scale ('never' to 'almost always'). Psychometric properties of the scale indicate excellent internal consistency ($\alpha = 0.95$) and strong convergent validity with other disparate measures of RNT (Ehring et al., 2011). A score of >28 was sufficient for inclusion.

- The *Test Your Memory measure* (TYM; Brown, Pengas, Dawson, Brown, & Clatworthy, 2009) was administered as a screen for significant cognitive difficulties that may preclude study inclusion. Scores below 42 are considered to represent likely memory decline (Brown et al., 2009) and therefore participants scoring below this were excluded from participation.

- High risk of self-injury was screened using the *Patient Health Questionnaire-9* (PHQ-9; Kroenke, Spitzer & Williams, 2001). Participants indicating high levels of depressed mood were consulted regarding the potential for participation to worsen their mood and given

clear opportunities to withdraw from the study at any time. The

researcher emailed participants weekly to monitor any adverse impact of

the WMTP on participants' mood.

Baseline Measures.

Participants and researcher met at the beginning of the agreed study period

to undertake baseline measures. These were:

- *Hospital Anxiety and Depression Scale (HADS; Zigmond & Snaith,
  1983).* This measure asks participants whether they have

  experienced any of seven anxiety and seven depression symptoms,

  each scored on a four point Likert scale (0-3). Scores range from 0-21

  for each subscale. A review of the relevant literature found a mean

  Cronbach alpha of .83 for HADS-anxiety and .82 for HADS-

  depression (Bjelland, Dahl, Haug, & Neckleman, 2002). The HADS

  has been shown to be applicable for use in older adults (OAs;

  Roberts, Fletcher, & Merrick, 2014).

- *Modified Sternberg (MS; Oberauer, 2001).* This WM measure

  requires participants to update WM using negatively-valenced stimuli.

  Participants are asked to memorise two lists of words, then cued with

  an indication to which of the lists the probe will be evaluated against

  (asked "was the probe from this cued word list?"). The MS is scored

  using the difference in reaction time between new words and words

  from the cued list. Oberauer (2001) included the use of the MS with

  an OA population.

- The PTQ was repeated at baseline meeting to ensure up-to-date measurement of RNT.

- A 'short form' version of the PTQ (SF-PTQ) was used to track daily fluctuations in RNT throughout the study (baseline and intervention phases). This was abridged by the researchers who selected five items identified by factor analysis of the scale (Ehring et al., 2011) as best representing the disengagement difficulties most relevant to the form of RNT targeted by the intervention (e.g. 'I get stuck on certain issues and I can't move on.').

All measures can be found detailed in Appendix D.

Intervention.

The intervention stage of the study required participants to engage in an affective WMTP for between 17-26 days. Seventeen days is the minimum length suggested to show effects from similar WMTPs (Onraedt & Koster, 2014). The WMTP was accessible to participants via the Internet on a home computer/laptop.

The intervention itself comprises a training paradigm adapted from that of Ecker and colleagues (2010), which focuses on updating information within the WM store. At the beginning of the day's training, participants engaged in a four-minute training session to learn that day's seven randomly assigned words. All words were negatively affective (e.g. 'fear'), in keeping with Schweizer and colleagues' (2013) findings that to produce gains in affective areas, training needs to be affective in content. Participants then completed 10 blocks of the

training task. In each block they were presented with three of the day's words.

They were then required to perform one of three steps to update these words

within WM storage: recall (recall which of the words was in the indicated

position); transform (+1 or +2 words working through the seven item word set)

or substitute (replace previous word with a different word). The number of steps

(n) within each block was adaptive. Each day's block started with the highest

number of steps from the previous day, less one (n-1). The number of steps

was then increased by one (n+1) when participants were achieving >50%

correct responses across a set of 5 blocks, and reduced by one step (n-1) if

>50% incorrect responses. At the end of the required number of steps,

participants were asked if a word displayed in a target box was the correct word

(Y/N response). Daily sessions lasted between 30 minutes and one hour.

The program automatically stored training data from participants for later

analysis. The highest number of steps successfully completed within a training

session was used as a measure of WM capacity.

Outcome Measures

Once participants had completed the intervention phase, they met again

with the researcher to undertake outcome measures. These were repeat

measurement of the full PTQ, HADS and modified Sternberg. Participants were

also asked to complete a short evaluation questionnaire to detail their

experience of the program.

Procedures

In keeping with randomised study design, participants were

independently randomised to intervention phase lengths of between 17-26

days, using random.org. To ensure all participants were allocated the same

study duration, baseline length varied with respect to identified study length,

such that baseline was between 10 and 20 days. Smith (2012) recommends at

least three data points for baseline, although length is generally dependent on

stability of the measured variable. This baseline length was considered likely to

allow the relation of change in RNT to the introduction of the WMTP (Kazdin,

2011). Throughout baseline participants completed the SF-PTQ daily.

Following defined baseline length, participants transitioned to

intervention phase. During this phase they continued to complete the SF-PTQ

daily, as well as undertaking daily sessions of the WMTP. At the end of this

phase, an exit meeting was undertaken to complete outcome measures of

mood (HADS), WM (MS), repetitive thinking (PTQ), along with study evaluation.

```
┌─────────────────────────────────────────┐
│              Screening                   │
│ PTQ                                      │
│ Medical History questionnaire            │
│ PHQ-9                                    │
│ TYM                                      │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│         Start of baseline meeting        │
│ HADS                                     │
│ Modified Sternberg task                  │
│ PTQ                                      │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│       Baseline phase: 10-20 days         │
│ Daily completion of SF-PTQ               │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│      Intervention phase: 17-26 days      │
│ Daily completion of WMTP                 │
│ Daily completion of SF-PTQ               │
└─────────────────────────────────────────┘
                    │
                    ▼
┌─────────────────────────────────────────┐
│              Exit meeting                │
│ HADS                                     │
│ Modified Sternberg task                  │
│ PTQ                                      │
│ WMTP evaluation form                     │
└─────────────────────────────────────────┘
```
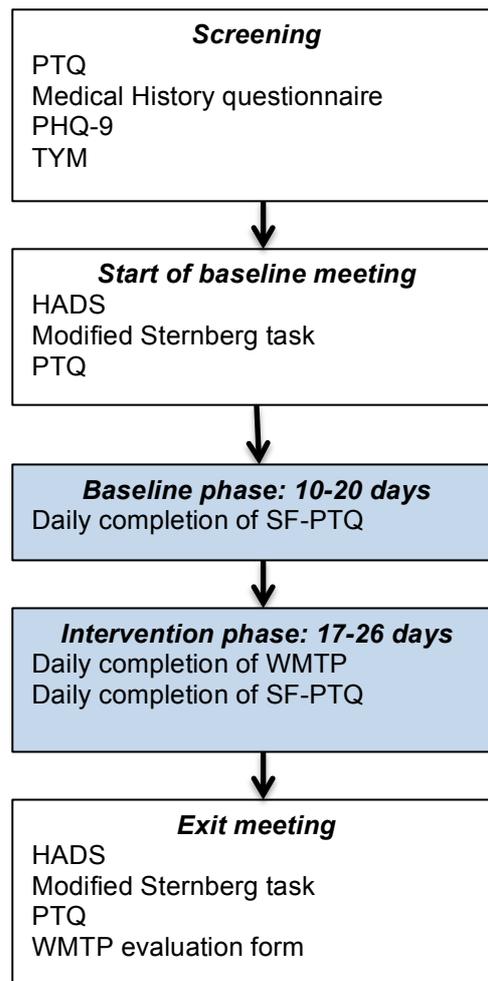
Figure 2. Study procedure.

Data Analysis Plan

Visual analysis of data is the most common form of analysis in SCED (Smith, 2012). Following the recommendations of Lane and Gast (2014), each participant's daily SF-PTQ scores will be presented as a graph and systematically compared. Trend lines will be added as these have been shown to aid visual analysis of data (Johnson & Ottenbacher, 1991). In accordance with the recommendations of Morley and Adams (1991), trend lines will be set

as the broadened median. This allows for appropriate visual analysis of variability and trend, without undue influence from outliers (Franklin, 2014).

Increasingly, statistical analysis of SCED data is recommended as an adjunct to visual analysis (Kazdin, 2011; Park, Marasculio, & Gaylord-Ross, 1990). Randomisation tests (Todman, 2002) will be used to investigate the impact of the WMTP on RNT for the older adult (OA) participants. Randomisation tests have been chosen for analysis in this design due to their robustness against Type I error (Levin, Ferron, & Gafurov, 2014) and lack of reliance on parametric assumptions, including the absence of serial dependence (Bulte & Onghena, 2009). In order to meet the assumptions of randomization tests, participants were randomly assigned to transition from baseline to experimental phase between 10-20 days after study commencement, giving ten possible moments of phase change (Kratochwill et al., 2010). Analysis will be undertaken in R (R Core Team, 2013), using the single case analysis package developed by Bulte and Onghena (SCDA; 2008). This analysis will indicate whether RNT has changed due to the introduction of the WMTP (H2).

In line with recommendations that SCED employ a measure of effect size alongside both visual and statistical analysis (Bulte & Onghena, 2008), percentage of all non-overlapping data (PAND; Parker, Hagan-Burke, & Vannest, 2007) will be calculated using SPSS to provide an indication of size of the effect. PAND was chosen due to its applicability to multiple baseline designs, in terms of required power and interpretability, and robustness to variability and outliers.

In order to indicate whether any training gains in WM have transferred to alternative WM measure (H1), pre-post scores on the Modified Sternberg will be analysed using a repeated measures t-test.

In order to ascertain the impact of the WMTP on RNT (H2) and depression/anxiety (H3), indicators of both reliable and clinically significant change will be calculated. Reliable change gives an indication of whether degree of change in a measured variable is greater than would be predicted by measurement error (Jacobson, Follette, & Revenstorf, 1984). Clinically significant change represents an analysis of each participants score to assess whether they have moved from 'problem' score to 'normal population' score (Jacobsen & Traux, 1991).

Evaluation of the WMTP will be exploratory and used to feed back in to the development of the training paradigm.

Results

Sample Characteristics

Characteristics of the sample can be seen summarised in Table 1 below. The sample consisted of six individuals, ranging in age from 65 to 77 years. All were retired and residing independently within the community. All met the inclusion and exclusion criteria for the study.

Table 1

Participant Characteristics

Participant No	Age	Gender
301	68	Female
302	65	Male
303	73	Female
304	65	Female
305	71	Male
306	77	Female

Screening, characterisation and pre-post data from the study is found in Appendix E.

Working Memory

Data regarding the participants' WM performance was stored by the WMTP (Figure 3).

Average

301
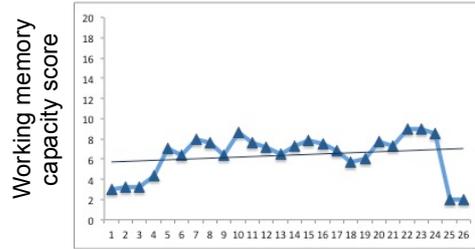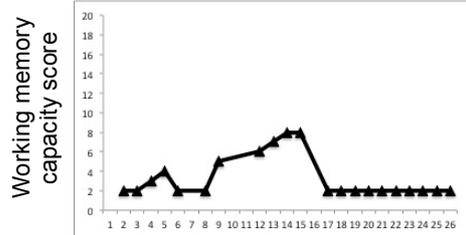
302
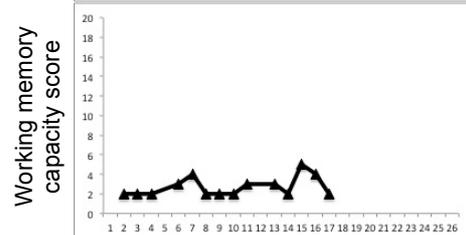
303
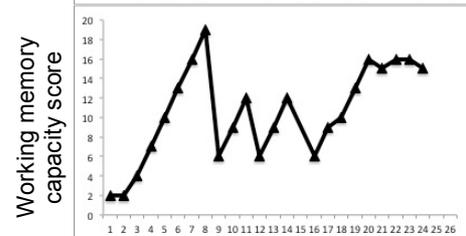
304
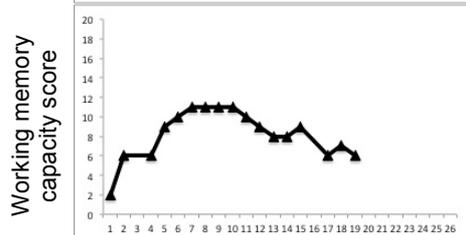
305
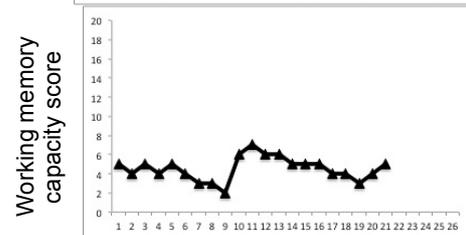
306
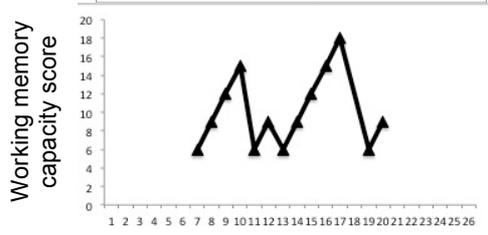
Days of training

Figure 3: Participants' WMTP performance plotted against days of training.

The trend line suggests a slight increase in WM capacity across the training, although any effect is small and within participant variability is clear. This may reflect errors in recording data within the WMTP, which required participants to complete and fully exit the WMTP in order to store data; hence a number of points of missing data. It is therefore not possible to conclude that the participants have improved WM functioning during training.

A manipulation check was undertaken using participants pre-post performance on the Modified Sternberg task. For analysis, participants' reaction times (RT) to interference cues (words they had been directed to forget) was compared against novel cues, with the expectation that improved WM updating would be demonstrated in an alignment of these RTs i.e. the difference in RTs to be closer to zero. A repeated measure t-test of the difference in RTs between interference and novel cues at baseline and outcome suggested that there was no significant difference ($t(5) = -2.498$, $p = 0.055$). Interestingly, the trend within the result is suggestive of an increased RT to interference compared to novel cues at outcome. Results here do not support a specific nor generalizable improvement in WM updating ability in participants following the WMTP. Hypothesis 1 must therefore be rejected.

Table 2

Reaction Times to Intrusion Compared to Novel Cues at Baseline and Outcome

on the Modified Sternberg

	Baseline			Outcome		
	Ave intrusion cue RT (ms)	*Ave novel cue RT (ms)*	*Difference RTs (ms)*	*Ave intrusion cue RT(ms)*	*Ave novel cue RT (ms)*	*Difference RTs (ms)*
301	2763	2850.625	-87.625	3056.857	1857.125	1199.732
302	3843.143	2674	1169.143	3437.625	1717.75	1719.875
303	1963.286	1595.875	367.411	1648.143	972.5	675.643
304	3098	1997.125	1100.875	3597.75	1852	1745.75
305	2111.333	2052.125	59.208	2376.857	1564.143	812.714
306	3476.667	2175.125	1301.542	2701.167	1712.857	989.31

Repetitive Negative Thinking

Participants tracked their RNT with daily completion of the SF-PTQ throughout baseline and intervention phases. Graphs to support visual analysis of this data are displayed below, in accordance with the guidelines of Barton and Reichow (2012). Larger format versions can be found in Appendix F.
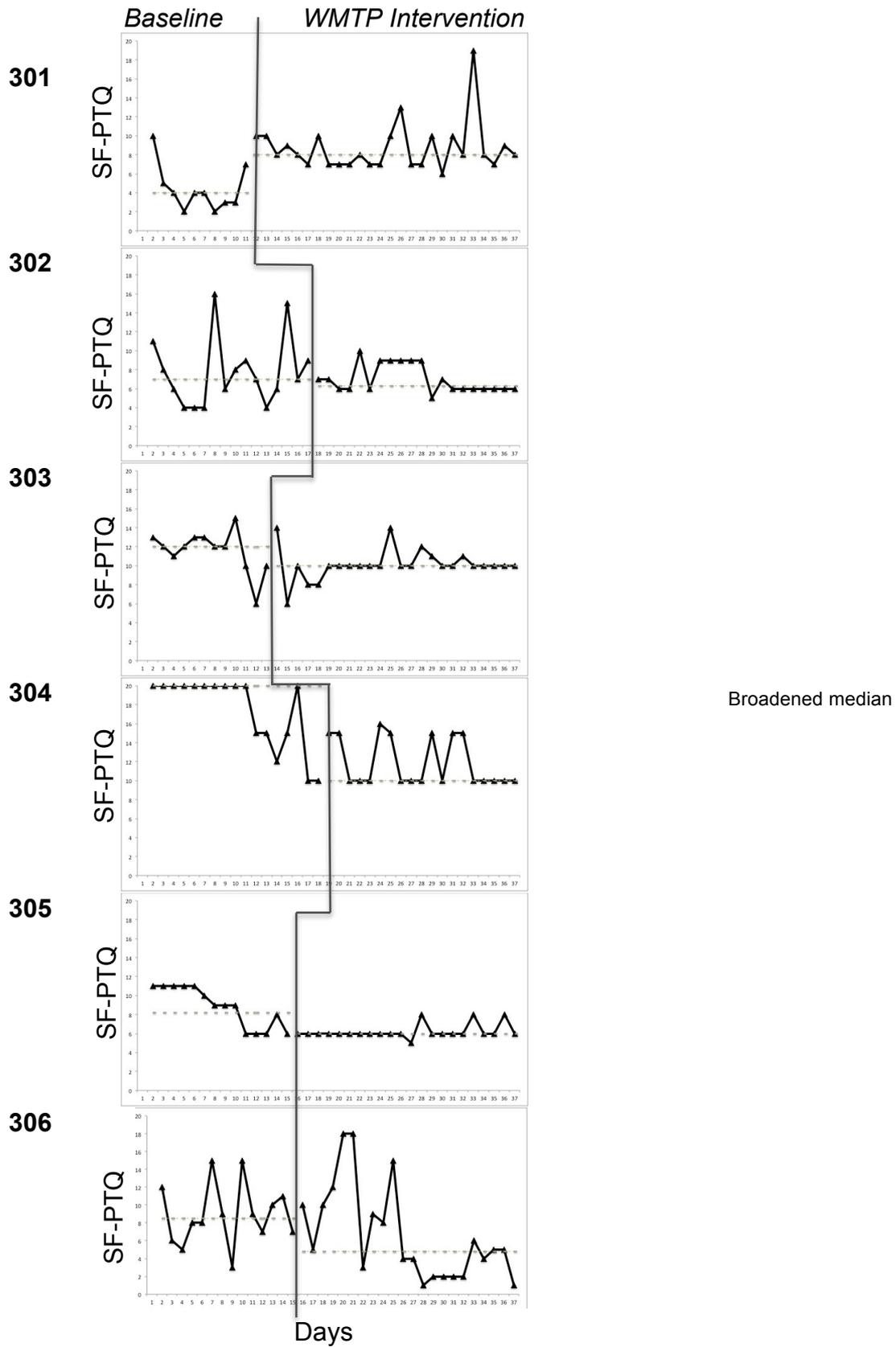
Figure 4: Multiple baseline graphs charting daily ratings of the SF-PTQ

throughout baseline and intervention phases.

For five of the six participants (302-306), broadened medians indicate a drop in SF-PTQ scores with commencement of the WMTP. However, such falls appear small and are diminished in visual analysis by the large degree of variability in scores across the study. As such, further statistical analysis is warranted. For four of the participants, visual trend analysis (Appendix G) indicated decreasing SF-PTQ scores during baseline, followed by stability during intervention. This may provide some guidance as to the length of time (up to 20days) needed to achieve stability in scores on this measure. The remaining two participants demonstrated increasing SF-PTQ scores during baseline, followed by a decline during intervention. For participant 306 this decline during intervention is steep, potentially indicating the impact of the intervention.

Statistical analysis was undertaken using the R package developed by Bulte and Onghena (SCDA; 2008). Script and output for the analysis can be found in Appendix H. With Monte Carlo randomisation (n=1000), a p value of 0.88 was established, suggesting that the WMTP did not significantly impact upon participants' WM. Percentage of all no overlapping data (PAND) resulted in a small-to-medium effect size (*Phi* = 0.28), indicating that only a small (*Phi*^{2} = 0.08) proportion of the variance in participants' SF-PTQ scores could be explained by their participation in the WMTP.

Reliable and clinically significant change in scores on the PTQ was assessed using norms provided by Ehring (2015).

Table 3

Participant Scores, RCI and CSC for the PTQ

Participant	Baseline PTQ	Outcome PTQ	Change	RCI	CSC
301	30	24	6	None	None
302	30	33	-3	None	None
303	33	31	2	None	None
304	54	42	12	Improve	None
305	29	37	-8	Deteriorate	None
306	36	22	14	Improve	Yes

RCI – reliable change indicator; CSC – clinically significant change

These results indicated that three participants showed no change on the PTQ, one participant (305) demonstrated reliable deterioration, and two (304 and 306) showed reliable improvement. Participant 306 also demonstrated a level of change which would be considered clinically significant i.e. moved out of the 'problem' range and in to the normal population range. It is interesting to note that participants 304 and 306 had the highest baseline PTQ scores, perhaps suggesting a possible floor effect in other participants, whose scores are closer to the general population average (28; Ehring, 2015). There is therefore some tentative support for hypothesis two, as a reduction occurred for one third of the participants, and there is a small-to-medium effect size when looking at the entire sample. However, the relationship does not meet statistical significance.

Mood

Participants completed the HADS (Zigmond & Snaith, 1983) at baseline and outcome to track change in depression and anxiety. Reliable and clinically significant change indicators were calculated for this data, demonstrating reliable change in two participants (304 and 306), with participant 306 also demonstrating clinically significant change.

Table 4

Participant Scores, RCI and CSCI for the HADS

Participant	Baseline HADS	Outcome HADS	Change	RCI	CSC
301	17	17	0	No change	None
302	21	19	2	No change	None
303	17	18	-1	No change	None
304	27	21	8	Improve	None
305	7	8	-1	No change	None
306	20	10	10	Improve	Yes

RCI – reliable change indicator; CSC – clinically significant change

Both participants 304 and 306 had also demonstrated reliable change on the PTQ, perhaps adding support for the relationship between these two concepts. It appears that, for a subset of participants, participation in the study related to a decrease in mood-related symptomatology. However, for the majority of participants, there was no change in mood symptoms following participation. The study can only therefore lend tentative support for hypothesis three.

WMTP Evaluation

As a novel intervention within this population, it was important to assess the acceptability of the intervention to participants. Results can be seen in Table 5 below.

Table 5

Participant Responses to Evaluation Questionnaire

Participant	Ease of use	Usefulness	Enjoyability
301	1	5	3
302	1	2	1
303	1	2	1
304	4	3	1
305	4	4	4
306	3	5	4

Scored on Likert scale from 1 (difficult/not useful/not enjoyable) to 5 (easy/useful/enjoyable)

Participants tended to report that the intervention was 'quite difficult' to use (modal response = 1), although two participants indicated they found it 'easy' to use. There was a spread of responses regarding the usefulness of the program, from 'not very useful' (two participants) to 'very useful' (two participants). The majority of participants (three) indicated that the program was 'not at all enjoyable', although two participants reported finding participation 'quite enjoyable'.

Qualitative feedback regarding the program is summarised in Figure 5 below.
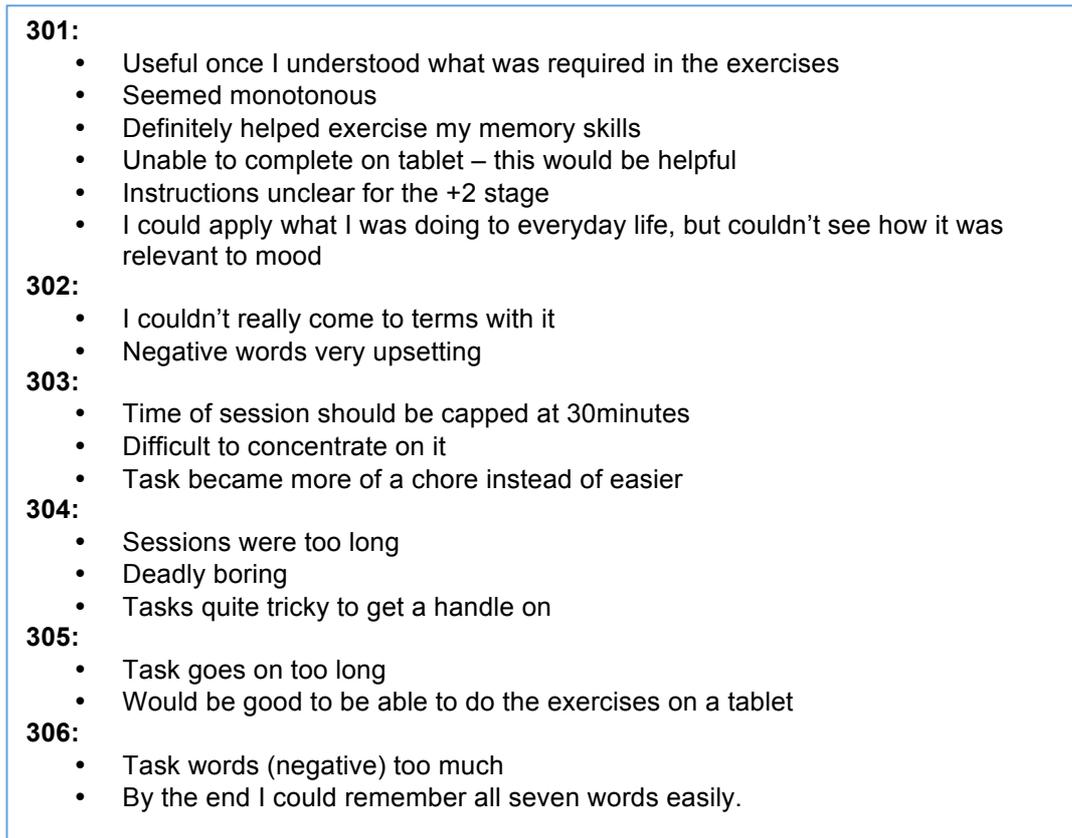
301:
- Useful once I understood what was required in the exercises
- Seemed monotonous
- Definitely helped exercise my memory skills
- Unable to complete on tablet – this would be helpful
- Instructions unclear for the +2 stage
- I could apply what I was doing to everyday life, but couldn't see how it was relevant to mood

302:
- I couldn't really come to terms with it
- Negative words very upsetting

303:
- Time of session should be capped at 30minutes
- Difficult to concentrate on it
- Task became more of a chore instead of easier

304:
- Sessions were too long
- Deadly boring
- Tasks quite tricky to get a handle on

305:
- Task goes on too long
- Would be good to be able to do the exercises on a tablet

306:
- Task words (negative) too much
- By the end I could remember all seven words easily.

Figure 5: Qualitative feedback from participants regarding the WMTP.

It appears that participants tended to find the WMTP fairly 'boring' to complete, particularly given that sessions could be long (due to the adaptive nature of the program). Some participants also reported that the negative affective valence of the target words was upsetting and limited their engagement with the intervention. Although the decision to use negatively valenced words was based upon relevant research literature (Schweizer, Grahn, Hampshire, Mobbs, & Dalgleish, 2013), it may be that this is not ethically defensible with a population of this kind. The qualitative feedback indicates four clear routes to improve the acceptability of the program:

- Improved clarity of instructions

- Limit to a predictable timescale (i.e. max 30 minutes)

- Mix of negative affective and neutral words

• Enable accessibility of WMTP on tablets as well as desktops/laptops

Discussion

The results of the study present a complex picture regarding possible usefulness of WMTP in an older adult (OA) population. The primary research question related to the possible impact of the WMTP in improving WM within this population. Evidence from data collected by the WMTP indicate a small trend towards an improvement in WM updating ability, with an increase in the number of steps from an average of just under 6 at commencement, moving to around 7 at the end of the intervention. However, graphical analysis of participants' data indicated variability in performance, which may be reflective of errors within the recording of capacity by the WMTP, such that participants were required to complete the training for their scores to be stored. Variability in scores may also be indicative of poor efficacy of the program, possible ceiling effects or possible mood-influences outside the control of the study (e.g. life events).

Pre- and post-intervention scores on the Modified Sternberg test were used as a manipulation check. It was anticipated that improvement in WM would be reflected in decreased difference in reaction time (RT) between intrusion and novel cues on this measure. Improvement in this direction would indicate that participants' ability to delete irrelevant information from their WM had improved. Results from analysis within this study were not in line with this prediction, with no significant difference in RTs at baseline compared to outcome. There is therefore no evidence of a transfer effect from the WMTP to

other assessments of WM. As Dahlin and colleagues (2008) have noted, there is little evidence of transfer effects in OA samples, indicating particular difficulty in transfer within this population. Together with inconclusive evidence of an improvement in WM from the WMTP data, it is most reasonable to conclude that the WMTP has not been universally effective in improving the WM of participants. As such, it is difficult for the results of this study to be used in support of the model proposed within the literature, as decreases in RNT and mood cannot be directly related to gains in WM ability.

Although there is not clear support for an improvement in WM, there is a suggestion within the data that participation has led to a degree of improvement in RNT in a subset of participants. Visual analysis of daily SF-PTQ data suggests a small decrease in SF-PTQ score during intervention phase, although further statistical analysis indicated that this improvement did not meet statistical significance. Two participants (one third of the sample) demonstrated reliable change in their PTQ scores at outcome compared to baseline. These participants had the highest baseline scores within the sample, reflecting Onraedt and Koster's (2014) finding that levels of rumination must be sufficiently high to demonstrate improvements from WM intervention. Both participants also demonstrated gains on the WMTP, with participant 304 showing a maximum increase of 9 during the program, and participant 306 demonstrating an increase of 12 during the intervention. These improvements may indicate that these two participants specifically benefitted from the program and experienced gains in their WM, and therefore could support improvements in WM as prompting improvements in RNT.

Unfortunately, one participant (305) demonstrated a reliable increase in RNT over the course of the study. Interestingly, this does not reflect their daily SF-PTQ, on which their scores showed a slight improvement during intervention. Questions included in the SF-PTQ were selected for their applicability to WM updating, and it may be that the participant improved in this aspect while declining in the wider aspects of RNT. This represents a weakness of the study in assessing the wider variability in repetitive thinking, with recent studies demonstrating differential predictive validity between abstract and concrete (Philippot, & Agrigoroaei, 2016) and positively- and negatively-valenced (Segerstrom et al., 2010) repetitive thinking.

A secondary research question for the study related to possible improvements in mood in participants following participation. A similar pattern was seen in this variable as in PTQ scores, with two participants demonstrating reliable improvement in mood scores. The same participants (304 and 306) improved in both RNT and mood, providing further support for the relationship between these two variables (Ricarte et al., 2016). Participant 306 demonstrated a clinically significant improvement in both RNT and mood, suggesting the possibility that – for some participants – a short intervention targeting WM updating can be effective in improving wellbeing in an OA population. Unfortunately, the information acquired regarding each participant is not sufficient to identify the individualised variables that may predict or explain the unique efficacy of the WMTP for this individual. Participant 306 had similar baseline WM and mood scores to other participants, and was not randomised to longer within the WMTP than other subjects, indicating that these would not be explanatory factors. Zinke and colleagues (2016) have indicated that

improvements in WM following a training intervention can be best predicted by

baseline performance, suggesting that this may be the most fruitful aspect to

focus on within future studies.

The final research question related to the acceptability of the WMTP

within an OA population. Participants varied in their experience of the

usefulness, enjoyability and ease of use of the program. Qualitative comments

made clear that participants struggled with the unpredictable length of sessions,

and tended to experience the program as monotonous to complete. Recent

research from Toril and colleagues (2016) suggests that more engaging

computerised interventions for OAs are more likely to be effective. As such, the

lack of reported engagement with the WMTP in this study may go some way to

explain its apparent lack of efficacy. Recent research has suggested that

effective cognitive training programs for OAs should last for over 30 minutes

and take place outside the home setting (Lampit et al., 2014). This does not

reflect the experience of participants within this study, who requested that the

daily program length be limited to half an hour, and that the WMTP should be

available on different platforms to aid completion at home. This represents a

common tension within psychological intervention between maximising potential

efficacy and encouraging engagement. Ethically, the needs of participants must

be respected, and participants' comments will be implemented ahead of future

use of the WMTP.

Methodological Critique and Directions for Future Research

The study demonstrates the usefulness of SCED in understanding the impact of interventions within a small sample. The use of visual and statistical analysis, alongside reporting measures of effect size, is in line with best practice interpretation in the area of SCED (Bulte, & Onghena, 2008). Analysis of the design using the RoBiN-T scale (Tate et al., 2013) indicates that the study is of adequate quality, and meets the majority of criteria required of high quality SCED. However, the study was limited in the ability to include blinding of researchers and participants within the design. Inclusion of a non-adaptive training program to participants unaware of this distinction, and support from a second researcher to undertake assessments, would enable these features to be built in to future studies.

The analysis of the daily SF-PTQ scores is limited by the variability seen in this measure. Stability of baseline data is an important tenet of SCED (Byiers, Reichle, & Symons, 2012), however the necessity of introducing randomisation of phase length (in order to undertake randomisation analysis) meant that it was not possible to extend baselines individually for participants until stability had been achieved. However, additional piloting with the measure to ascertain likely length of stabilisation may have allowed for adaptation of the protocol to account for this process (i.e. by extending minimum baseline length). Furthermore, improved recording for external factors influencing participants' mood (e.g. health or social difficulties) would allow greater control for these factors.

The design of the study also restricts its ability to ascertain delayed effects of training. Participants were engaged in outcome testing on the final

day of their completion of the program. As such, slower-acting impacts of the

WMTP may not have been accounted for in post-intervention data. In future

studies, a second 'intervention' phase, which continues to track daily SF-PTQ

scores following completion of the program, would allow for this information to

be collected. Similarly, a longer-term follow-up of outcome measures would

allow for analysis of the maintenance of any effects shown.

Theoretical and Clinical Implications and Future Directions

The study aimed to assess the impact and clinical utility of a WMTP

intervention for older adults (OAs). Results suggest that for one third of

participants, engagement in the study was related to a reliable decrease in RNT

and mood symptoms. For these participants it is possible that gains in these

areas relate to improvements in WM found through participation in the WMTP,

and therefore tentatively adds to the impaired disengagement hypothesis of

Koster and colleagues (2011), indicating that decline in WM ability may be a

causal factor in low mood and anxiety, due to the decreased ability to update

negative information within the WM store. The use of this model, and an

updating task-based WMTP, within an OA population also adds to the literature

surrounding Hasher and Zacks' (1988) inhibition deficit hypothesis, suggesting

that OAs may be particularly susceptible to difficulties with WM updating.

However, for the remaining two thirds of participants, this effect was not

seen. It therefore is not possible to conclude a universal effect. The efficacy of

the WMTP does not appear to be based on pre-intervention mood, RNT or WM

scores, nor to relate to length of engagement in the WMTP. It is therefore likely

that an alternative factor is needed to explain the individualised impact for two

participants. Research has indicated that mechanisms for low mood may be

different depending on age of depression onset, with WM impairment more likely to be a causal factor in late-onset depression (von Hippel et al., 2008). Information regarding the onset and course of RNT and/or mood difficulties was not obtained for this study, and may represent a possible confounding factor in predicting which participants benefitted from engagement. Future studies should take steps to identify predictive variables, such that, within clinical practice, relevant interventions can be targeted towards individuals most likely to benefit.

In evaluating the impact of this WMTP, it is important to also to consider the wider literature regarding the efficacy of WM training. In their meta-analysis on the subject, Melby-Lervåg & Hulme (2013) have concluded that WM improvement does not transfer beyond small, short-term and training-specific gains. Given the demonstrated relationship between WM and fluid intelligence (Salthouse & Pink, 2008), and the evidence that WM ability is a stable, trait factor (Koenig et al., 2015), it is appropriate to question to what degree WM is amenable to intervention. The failure of the WMTP used within this study to significantly improve ability at the training or transfer tasks may therefore simply indicate that WM is not responsive to this type of intervention.

At the same time, research supports the assertion that WM is a complex, multi-faceted construct, with executive elements relating to updating most likely to be of relevance to the WM impairments seen in depression (Channon, Baker, & Robertson, 1993). It may be that the development of nuanced computerised interventions targeted closely to this specific aspect are most likely to prove efficacious. Similarly, downstream interventions, including the use of

rumination-focused CBT (Watkins et al., 2007) may also be helpful in providing

clients with tools to manage stable WM updating deficits that precipitate RNT.

The study also indicated factors related directly to the WMTP that may

have influenced its efficacy, namely, that participants described the intervention

as 'boring', and the use of negatively-valenced stimuli as upsetting. In the

service of both ethics and efficacy, future investigations using this intervention

ought to increase its accessibility. This could be achieving by enforcing

predictable time limits to sessions, using a mix of positively- and negatively-

valenced stimuli and developing a more visually engaging interface or scenario.

Conclusion

In conclusion, the results of the study contribute to a growing body of

literature exploring the relationship between WM, RNT and mood in older adults

(OAs). OAs represent a vital population for research in this area, given the

natural decline in WM seen with age (Swanson, 1999), and the public health

concern over the lack of effective interventions for mood disorders within OA

populations (Frazer, Christensen, & Griffiths, 2005). Interventions targeting WM

may therefore represent an appropriate avenue for investigation in the

development of effective, appropriate interventions for OAs. This study provides

clear evidence of the usefulness of such interventions for a proportion of the OA

population, although further research is warranted to investigate the factors that

may predict a positive impact of WM training. The study highlights the

importance of ensuring that computerised cognitive interventions for OA

populations are designed in an engaging manner.

References

Babcock, R. L., & Salthouse, T. A. (1990). Effects of increased processing

demands on age differences in working memory. *Psychology and Aging, 5,*

421-428.

Baddeley, A. D. (1997). *Human memory: Theory and practice.* Hove, UK:

Psychology Press Ltd.

Barton, E. E., & Reichow, B. (2012). Guidelines for graphing data with Microsoft®

Office 2007™, Office 2010™, and Office for Mac™ 2008 and 2011. *Journal*

of Early Intervention, 39, 124-150. doi: 10.1177/1053815112456601

Bjelland, I., Dahl, A. A., Haug, T. T., & Neckelmann, D. (2002). The validity of the

Hospital Anxiety and Depression Scale: An updated literature review.

Journal of Psychosomatic Research, 52, 69-77. doi:10.1016/S0022-

3999(01)00296-3

Bogner, H. R., Bruce, M. L., Reynolds, C. F., Mulsant, B. H., Cary, M. S., Morales,

K., & Alexopolous, G. S. (2007). The effects of memory, attention, and

executive dysfunction on outcomes of depression in a primary care

intervention trial: the PROSPECT study. *International Journal of Geriatric*

Psychiatry, 22, 922-9.

British Psychological Society. (2010). *Code of Human Research Ethics.* Retrieved

from

http://www.bps.org.uk/sites/default/files/documents/code_of_human_researc

h_ethics.pdf

Brown, J., Pengas, G., Dawson, K., Brown, L., & Clatworthy, P. (2009). Self-

administered cognitive screening test (TYM) for detection of Alzheimer's

disease: cross sectional study. *British Medical Journal, 338,* 2030-39. doi:

10.1136/bmj.b2030

Bulte, I., & Onghena, P. (2008). An R package for single-case randomization tests.

Behaviour Research Methods, 40, 467-478.

Bulte, I., & Onghena, P. (2009). Randomization tests for multiple-baseline designs:

an extension of the SCRT-R package. *Behavior Research Methods, 41,*

477-485. doi: 10.3758/BRM.41.2.477.

Butters, M. A., Bhalla, R. K., Andreescu, C., Loebach Wetherall, J., Mantella, E.,

Begley, A. E., & Lenze, E. J. (2011). Changes in neuropsychological

functioning following treatment for late-life generalised anxiety disorder. *The

British Journal of Psychiatry, 199,* 211-218 doi: 10.1192/bjp.bp.110.090217

Byiers, B .J., Reichle, J., & Symons, F. J. (2012). Single-subject experimental

design for evidence-based practice. *American Journal of Speech and

Language Pathology, 21,* 397-414. doi: 10.1044/1058-0360(2012/11-0036)

Channon, S., Baker, J., & Robertson, M. M. (1993). Working memory in clinical

depression: An experimental study. *Psychological Medicine, 23,* 87-91.

Christopher, G., & MacDonald, J. (2005). The impact of clinical depression on

working memory. *Cognitive Neuropsychiatry, 10,* 379-399. doi:

10.1080/13546800444000128

Conway, A., Kane, M., Bunting, M., Hambrick, D., Wilhem, O., & Engle, R. (2005).

Working memory span tasks: A methodological review and user's guide.

Psychonomic Bulletin & Review, 12, 769-786. doi: 10.3758/BF03196772

Crawford, J. R., & Garthwaite, P. H. (2012). Single-case research in

neuropsychology: A comparison of five forms of t-test for comparing a case

to controls. *Cortex, 48,* 1009-101.

Curci, A., Lanciano, T., Soleti, E., & Rimé, B. (2013). Negative emotional

experiences arouse rumination and affect working memory capacity.

Emotion, 13, 867-80. doi: 10.1037/a0032492

Dahlin, E., Neely, A. S., Larsson, A. Backman, L., & Nyberg, L. (2008). Transfer of

learning after updating training mediated by the striatum. *Science, 13,* 1510-

1512. doi: 10.1126/science.1155466.

Darke, S. (1988). Anxiety and working memory capacity. *Cognition and Emotion, 2,*

145-154. doi: 10.1080/02699938808408071

Deldin, P., Deveney, C., Kim, A., & Webb, J. (2001). A slow wave investigation of

working memory bias in mood disorders. *Journal of Abnormal Psychology,*

110, 267-281. doi: 10.1037//0021-843X.110.2.267

Dobbs, A. R., & Rule, B. G. (1989). Adult age differences in working memory.

Psychology and Aging, 4, 500-503.

Ehring, T., & Watkins, E. R. (2008). Repetitive negative thinking as a

transdiagnostic process. *International Journal of Cognitive Therapy, 1,* 192-

205.

Ehring, T, Zetsche, U., Weidacker, K., Wahl, K., Schonfeld, S., & Ehlers, A. (2011).

The Perseverative Thinking Questionnaire (PTQ): Validation of a content-

independent measure of repetitive negative thinking. *Journal of Behaviour*

Therapy and Experimental Psychiatry, 42, 225-232. doi:

10.1016/j.jbtep.2010.12.003

Franklin, R.D., Allison, D.B., & Gorman, B.S. (2014). *Design and analysis of single-case research*. Hove: Psychology Press.

Frazer, C. J., Christensen, & Griffiths, K.M. (2005) Effectiveness of treatments for depression in older people. *Medical Journal of Australia, 182*, 627-632.

Gotlib, I., & Joorman, J. (2010) Cognition and depression: Current status and future directions. *Annual Review of Clinical Psychology, 6,* 1-28. doi: 10.1146/annurev.clinpsy.121208.131305

Harvey, P. O., Le Bastard, G., Pochon, J.B., Levy, R., Allilaire, J.F., Dubois, B., & Fossati, P.J., (2004). Executive functions and updating of the contents of working memory in unipolar depression. *Journal of Psychiatric Research, 38*, 567-76.

Hasher, L., & Zacks, R. T. (1988). Working memory, comprehension, and aging: A review and a new view. In G.H. Bower (Ed), *The psychology of learning and motivation* (Vol. 22, pp. 193-225). New York: Academic Press.

Herrmann, L. L., Goodwin, G. M, & Ebmeier, K. P. (2007). The cognitive neuropsychology of depression in the elderly. *Psychological Medicine, 37,* 1693-1702. doi: http://dx.doi.org/10.1017/S0033291707001134

Hester, R., Kinsella, G., & Ong, B. (2004). Effect of age on forward and backward span tasks. *Journal of the International Neuropsychological Society, 10,* 475-481. doi: https://doi.org/10.1017/S1355617704104037

Hoorelbeke, K., Koster, E. H., Vanderhasselt, M. A., Callewaert, S., & Demeyer, I. (2015). The influence of cognitive control training on stress reactivity and rumination in response to a lab stressor and naturalistic stress. *Behaviour Research and Therapy, 69,* 1-10. doi: 10.1016/j.brat.2015.03.010.

Hoyer, J., Gloster, A. T., & Herzberg, P. Y. (2009). Is worry different from

rumination? Yes, it is more predictive of psychopathology! *Psychosocial*

Medicine, 6, 1-9. doi: 10.3205/psm000062

Hubbard, N., Hutchinson, J., Turner, M., Montroy, J., Bowles, R., Rympa, B.

(2016). Depressive thoughts limit working memory capacity in dysphoria.

Cognition and Emotion, 30, 193-209. doi: 10.1080/02699931.2014.991694

Iacoviello, B. M., Wu, G., Alvarez, E., Huryk, K., Collins, K. A., Murrough, J.W.,

Iosifescu, D. V., & Charney, D. S. (2014). Cognitive-emotional training as an

intervention for major depressive disorder. *Depression and Anxiety: Focus*

on Treatment, 31, 699–706. doi: 10.1002/da.22266

Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome

research: methods for reporting variability and evaluating clinical

significance. *Behavior Therapy, 15,* 336-352.

Jacobson, N. S., & Truax, P. (1991). Clinical significance: a statistical approach to

defining meaningful change in psychotherapy research. *Journal of*

Consulting and Clinical Psychology, 59, 12-19.

Johnson, M. B., & Ottenbacjer, K. J. (1991). Trend line influence on visual analysis

of single-subject data in rehabilitation research. *International Disability*

Studies, 13, 55-9.

Joorman, J., & Gotlib, I. H. (2008). Updating the contents of working memory in

depression: Interference from irrelevant negative material. *Journal of*

Abnormal Psychology, 117, 182-192. doi: 10.1037/0021-843X.117.1.182

Joormann, J., Levens, S., & Gotlib, I. H. (2011). Sticky thoughts: Depression and

rumination are associated with difficulties manipulating emotional material in

working memory. *Psychological Science, 22,* 979-983*.*

Joorman, J., & Tran, T. B. (2009). Rumination and intentional forgetting of

emotional material. *Cognition and Emotion, 23,* 1233-1246.

Kane, M. J., & Engle, R. W. (2000). Working memory capacity, proactive

interference, and divided attention: Limits on long-term memory retrieval.

Journal of Experimental Psychology: Learning, Memory, & Cognition, 26,

333-358.

Kazdin, A. E., & Kopel, S. A. (1975). On resolving ambiguities of the multiple-

baseline design: Problems and recommendations. *Behavior Therapy, 6,*

601-08.

Kazdin, A. E. (2011). *Single-case research designs: Methods for clinical and

applied settings* (2nd ed.). New York, USA: Oxford University Press.

Kircanski, K., Joormann, J., & Gotlib, I. H. (2012), Cognitive aspects of depression.

WIREs Cognitive Science, 3, 301–313. doi: 10.1002/wcs.1177

Koenig, A. M., DeLozier, I. J., Zmuda, M. D., Marron, M. M., Begley, A. E.,

Anderson, S. J., … Butters, M. A. (2015). Neuropsychological functioning in

the acute and remitted states of late-life depression. *Journal of Alzheimer's

Disease, 45*, 175-185. doi: 10.3233/JAD-148006

Koster, E. H., De Lissnyder, E., Derakshan, N., & De Raedt, R. (2011).

Understanding depressive rumination from a cognitive science perspective:

the impaired disengagement hypothesis. *Clinical Psychology Review, 31,*

138-45. doi: 10.1016/j.cpr.2010.08.005.

Kratochwill, T. R., Hitchcock, J., Horner, R. H., Levin, J. R., Odom, S. L., Rindskopf, D. M., & Shadish, W. R. (2010). Single-case designs technical documentation. In *What Works Clearinghouse: Procedures and standards handbook* (Version 2.0). Retrieved from What Works Clearinghouse website: http://ies.ed.gov/ncee/wwc/pdf/wwc_scd.pdf

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The PHQ-9: Validity of a brief depression severity measure. *Journal of General Internal Medicine, 16,* 606-13.

Kueider, A. M., Parisi, J. M., Gross, A. L., & Rebok, G. W. (2012). Computerized cognitive training with older adults: A systematic review. *PloS ONE, 7*, e40588. doi:10.1371/journal.pone.0040588

Lampit, A., Hallock, H., & Valenzuela, M. (2014). Computerized cognitive training in cognitively healthy older adults: a systematic review and meta-analysis of effect modifiers. *PLoS Medicine, 11,* e1001756. doi: 10.1371/journal.pmed.1001756

Lane, J. D., & Gast, D. L. (2014). Visual analysis in single case experimental design studies: brief review and guidelines. *Neuropsychological Rehabilitation, 24,* 445-63. doi: 10.1080/09602011.2013.815636.

Levin, J. R., Ferron, J. M., & Gafurov, B. S. (2014). Improved randomization tests for a class of single-case intervention design. *Journal of Modern Applied Statistical Methods, 13,* 2-52.

McAvinue, L. P., Golemme, M., Catorina, M., Tatti, E., Pigni, F. M., Salomone, S. … Rovertson, I.H. (2013). An evaluation of a working memory raining scheme in older adults. *Frontiers in Aging Neuroscience, 5,* 20-27. doi: 10.3389/fnagi.2013.00020

Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A

meta-analytic review. *Developmental Psychology, 49,* 270-291. doi:

10.1037/a0028228

Miyake, A., Friedman, N., Emerson, M., Witzki, A., Howerter, A., & Wager, T.

(2000). The unity and diversity of executive functions and their contributions

to complex 'frontal lobe' tasks: A latent variable analysis. *Cognitive*

Psychology, 41, 49-100. doi: 10.1006/cogp.1999.0734

Mohlman, J., & Gorman, J. M. (2005). The role of executive functioning in CBT: a

pilot study with anxious older adults. *Behaviour Research and Therapy, 43,*

447–465.

Moran, T. P. (2016). Anxiety and working memory capacity: A meta-analysis and

narrative review. *Psychological Bulletin, 142,* 831-864. doi:

10.1037/bul0000051

Morimoto, S., Wexler, B., Liu, J., Hu, W., Seirup, J., & Alexopoulos, G. (2014).

Neuroplasticity-based computerized cognitive remediation for treatment

resistant geriatric depression. *Nature Communications, 5,* 4579-4585. doi:

10.1038/ncomms5579

Morley, S., & Adams, M. (1991). Graphical analysis of single-case time series data.

British Journal of Clinical Psychology, 30, 97-115. doi: 10.1111/j.2044-

8260.1991.tb00926.x

Nock, M. K., Michel, B. D., & Photos, V. I. (2007). Single-case research designs. In

D. McKay (Ed.), *Handbook of research methods in abnormal and clinical*

psychology (pp. 337-350). Thousand Oaks, CA: Sage.

Nolen-Hoeksema, S. (1991). Responses to depression and their effects on the

duration of depressive episodes. *Journal of Abnormal Psychology, 100,* 569-

582.

Nolen-Hoeksema, S. (2000). The role of rumination in depressive disorders and

mixed anxiety/depressive symptoms. *Journal of Abnormal Psychology, 109,*

504-511.

Nolen-Hoeksema, S., Morrow, J., & Fredrickson, B. L. (1993). Response styles and

the duration of episodes of depressed mood. *Journal of Abnormal*

Psychology, 102, 20–28.

Nolen-Hoeksema, S., Wisco, B. E., & Lyubomirsky, S. (2008). Rethinking

rumination. *Perspectives on Psychological Science, 3,* 400-424. doi:

10.1111/j.1745-6924.2008.00088.x

Oberauer, K. (2001). Removing irrelevant information from working memory: A

cognitive aging study with the modified Sternberg task. *Journal of*

Experimental Psychology: Learning, Memory and Cognition, 27, 948-57.

doi: 10.1037/0278-7393.27.4.948

Oberauer, K., Süß, H. M., Schulze, R., Wilhelm, O., & Wittmann, W. W. (2000).

Working memory capacity: Facets of a cognitive ability construct.

Personality & Individual Differences, 29, 1017-1045.

OCEBM Levels of Evidence Working Group (2016). *The Oxford 2011 Levels of*

Evidence: Oxford Centre for Evidence-Based Medicine. Retrieved from:

http://www.cebm.net/index.aspx?o=5653

Onghena, P., & Edginton, E. S. (2005). Customization of pain treatments: Single-

case design and analysis. *Clinical Journal of Pain, 21,* 56-68.

Onraedt, T., & Koster, E. H. (2014). Training working memory to reduce rumination.

PloS ONE, 9, e90632. doi:10.1371/journal.pone.0090632

Owens, M., Koster, E. H., & Derakshan, N. (2012). Impaired filtering of irrelevant

information in dysphoria: An ERP study*. Social, Cognitive and Affective

Neuroscience, 7,* 752-63. doi: 10.1093/scan/nsr050.

Owens, M., Koster, E. H., & Derakshan, N. (2013). Improving attention control in

dysphoria through cognitive training: Transfer effects on working memory

capacity and filtering efficiency. *Psychophysiology, 50,* 297-307. doi:

10.1111/psyp.12010.

Pantzar, A., Laukka, E. J., Atti, A. R., Fastbom, J., Fratiglioni, L., & Bäckman, L.

(2014). Cognitive deficits in unipolar old-age depression: a population-based

study. *Psychological Medicine, 44,* 937-947.

Park, H., Marasculio, L., & Gaylord-Ross, R. (1990). Visual inspection and

statistical analysis in single-case designs. *The Journal of Experimental

Education, 58*, 311-320. doi: http://www.jstor.org/stable/20151823

Parker, R.I., Hagan-Burke, S., & Vannest, K. (2007). Percentage of all non-

overlapping data: An alternative to PND. *The Journal of Special Education,

40*, 194-204.

Philippot, P., & Agrigoroaei, S. (2016). Repetitive thinking, executive functioning,

and depressive mood in the elderly. *Aging & Mental Health, 2,* 1-5. doi:

10.1080/13607863.2016.1211619

Pimontel, M. A., Culang-Reinlieb, Morimoto, S. S., & Sneed, J. R. (2012).

Executive dysfunction and treatment response in late-life depression.

International Journal of Geriatric Psychiatry, 27, 893-9. doi: 10.1002/gps.2808.

Quinn, M., Keil, D., Utke, S., & Joorman, J. (2014). Trait rumination moderates the effect of executive control training. *Journal of Experimental Psychopathology, 5,* 289-301. doi: http://dx.doi.org/10.5127/jep.038713

R Core Team (2013). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria.   Obtained from http://www.R-project.org/.

Reijnders, J., van Heugten, C., & van Boxtel, M. (2013). Cognitive interventions in healthy older adults and people with mild cognitive impairment: A systematic review. *Aging Research Reviews, 12*, 263-75. doi:10.1016/j.arr.2012.07.003

Ricarte, J., Ros, L., Serrano, J. P., Martinez-Lorca, M., & Latorre, J. M. (2016). Age differences in rumination and autobiographical retrieval. *Aging & Mental Health, 20,* 1063-1069. doi: 10.1080/13607863.2015.1060944.

Roberts, M. H., Fletcher, R. B., & Merrick, P. L. (2014). The validity and clinical utility of the hospital anxiety and depression scale (HADS) with older adult New Zealanders. *International Psychogeriatrics, 26,* 325-33. doi: 10.1017/S104161021300197X.

Roberts, H., Watkins, E., & Wills, A. (2013). Cueing an unresolved personal goal causes persistent ruminative self-focus: An experimental evaluation of control theories of rumination. *Journal of Behaviour Therapy and Experimental Psychiatry, 44,* 449-455.

Salthouse, T. A., & Pink, J. E. (2008). Why is working memory related to fluid

intelligence? *Psychonomic Bulletin & Review, 15,* 364-371. doi:

10.3758/PBR.15.2.364

Schweizer, S., Grahn, J., Hampshire, A., Mobbs, D., & Dalgleish, T.J. (2013).

Training the emotional brain: Improving affective control through emotional

working memory training. *The Journal of Neuroscience, 33,* 5301-11. doi:

10.1523/JNEUROSCI.2593-12.2013.

Segerstrom, S. C., Roach, A. R., Evans, D. R., Schipper, L. J., & Darville, A.K.

(2010). The structure and health correlates of trait repetitive thought in older

adults. *Psychology and Aging, 25,* 505-515. doi: 10.1037/a0019456.

Segerstrom, S. C., Stanton, A. L., Alden, L. E., & Shortridge, B. E. (2003). A

multidimensional structure for repetitive thought: What's on your mind, and

how, and how much? *Journal of Personality and Social Psychology, 85,*

909–921.

Sexton, C. E., McDermott, L., Kalu, U. G., Herrmann, L. L., Bradley, K. M., Allan, C.

L., Le Masurier, M., Mackay, C. E., & Ebmeier, K. P. (2012). Exploring the

pattern and neural correlates of neuropsychological impairment in late-life

depression. *Psychological Medicine, 42*, 1195-1202 doi: http://0-

dx.doi.org.lib.exeter.ac.uk/10.1017/S0033291711002352

Shackman, A., Issidoros, S., Maxwell, J., Pizzagalli, D., Lavric, A., & Davidson, R.

(2006). Anxiety selectively disrupts visuospatial working memory. *Emotion,

6,* 40-61. http://dx.doi.org/10.1037/1528-3542.6.1.40

Shivayogi, P. (2013). Vulnerable populations and methods for their safeguard.

Perspectives on Clinical Research. 4, 53–57. doi: 10.4103/2229-

3485.106389.

Smith, J. D. (2012). Single-case experimental designs: A systematic review of

published research and current standards. *Psychological Methods, 17,* 510-

50. doi: 10.1037/a0029312

Sneed, J. R., Roose, S. P., Keilp, J. G., Krishnan, K. R., Alexopoulos, G. S., &

Sackeim, H. A. (2007). Response inhibition predicts poor antidepressant

treatment response in very old depressed patients. *American Journal of

Geriatric Psychiatry, 15,* 553-63.

Sütterlin, S., Paap, M., Babic, S., Kübler, A. & Vögele, C. (2012). Rumination and

age: Some things get better. *Journal of Aging Research,*

doi:10.1155/2012/267327

Swanson, H. L. (1999). What develops in working memory? A life span

perspective. *Developmental Psychology, 35*, 986-1000.

Synder, H. (2013). Major depressive disorder is associated with broad impairments

on neuropsychological measures of executive function: A meta-analysis and

review. *Psychological Bulletin, 139,* 81–132. doi: 10.1037/a0028727

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,* 619-638.

doi: 10.1080/09602011.2013.824383.

Thomas, A. J., Gallagher, P., Robinson, L.J., Porter, R. J., Young, A.H., Ferrier,

I.N., & O'Brien, J. T. (2009) A comparison of neurocognitive impairment in

younger and older adults with major depression. *Psychological Medicine,

39,* 725-733. doi: http://0-

dx.doi.org.lib.exeter.ac.uk/10.1017/S0033291708004042

Todman, J. (2002). Randomisation in single-case experimental designs. *Advances in Clinical Neuroscience and Rehabilitation, 2,* 18-19.

Topper, M., Molenaar, D., Emmelkamp, P., & Ehring, T. (2014). Are rumination and worry two sides of the same coin? A structural equation modelling approach. *Journal of Experimental Psychopathology, 5,* 363-381. doi: 10.5127/jep.038813

Toril, P., Reales, J. M., Maya, J., & Ballesteros, S. (2016). Video game training enhances visuospatial working memory and episodic memory in older adults. *Frontiers in Human Neuroscience, 10,* 206-212. doi: 10.3389/fnhum.2016.00206

Treynor, W., Gonzalez, R., & Nolen-Hoeksema, S. (2003). Rumination reconsidered: A psychometric analysis. *Cognitive Therapy and Research, 27,* 247-259. doi: http://dx.doi.org/10.1023/A:1023910315561

University of the Third Age. (2015). Retrieved from: http://www.u3a.org.uk/

Van Dillen, L. F., & Koole, S. L. (2007). Clearing the mind: A working memory model of distraction from negative mood. *Emotion, 7,* 715-723. doi: 10.1037/1528-3542.7.4.715

Vecchi, T., & Cornoldi, C. (1999). Passive storage and active manipulation in visuo-spatial working memory: Further evidence from the study of age differences. *European Journal of Cognitive Psychology 11,* 391-406. doi: 10.1080/713752324

von Hippel, W., Vasey, M., Gonda, T., & Stern, T. (2008). Executive function deficits, rumination and late-onset depressive symptoms in older adults.

Cognitive Therapy and Research, 32, 474-487. doi: 10.1007/s10608-006-9034-9

Wanmaker, S., Geraerts, E., & Franken, I. H. (2015). A working memory training to decrease rumination in depressed and anxious individuals: A double-blind randomized controlled trial. *Journal of Affective Disorders, 175,* 310-19. doi: http://dx.doi.org/10.1016/j.jad.2014.12.027

Wanmaker, S., Hopstaken, J. F., Asselbergs, J., Geraerts, E., & Franken, I.H.A. (2014). Decreasing dysphoric thoughts by a working memory training: A randomized double-blind placebo-controlled trial. *Journal of Depression and Anxiety, 3,* 165-168. doi:10.4172/2167-1044.1000165

Watkins, E., Scott, S., Wingrove, J., Rimes, K., Bathurst, N., Steiner, H., … Maliaris, Y. (2007). Rumination-focused cognitive behaviour therapy for residual depression: A case series. *Behaviour Research and Therapy, 45,* 2144-2154. doi: 10.1016/j.brat.2006.09.018

Watson, P. J., & Workman, E. A. (1981). The non-concurrent multiple baseline across-individuals design: An extension of the traditional multiple baseline design. *Journal of Behavior Therapy and Experimental Psychiatry, 12,* 257–259.

Wells, A. (2006). The metacognitive model of worry and generalized anxiety disorder. In G.C. Davey & A. Wells (Eds.), *Worry and its psychological disorders: Theory, Assessment and treatment.* Chichester, UK: John Wiley & Sons.

Wilhelm, O., Hildebrandt, A., Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology, 4,* 433-452. doi: 10.3389/fpsyg.2013.00433

Zigmond, A. S., & Snaith, R. P. (1983). The hospital anxiety and depression scale.

Acta Psychiatrica Scandinavica, 67, 361–370. doi:10.1111/j.1600-

0447.1983.tb09716.x. PMID 6880820.

Zinke, K., Zeintl, M., Rose, N. S., Putzman, J., Pydde, A., & Kliegel, M. (2014).

Working memory training and transfer in older adults: Effects of age,

baseline performance, and training gains. *Developmental Psychology, 50,*

304-315. doi: 10.1037/a0032982

Appendices

Appendix A: Risk of Bias in N-of-1 Trials Scale

As applied to current study.

Items in RoBiNT Scale (Tate et al., 2013)

	Score
Internal validity subscale	
1.Design	1/2
2.Randomisation	2/2
3.Sampling behaviour (all phases	2/2
4.Blinding patient/therapist	0/2
5.Blinding assessors	0/2
6.Inter-rater reliability	1/2
7.Treatment adherence	1/2
External validity and interpretation subscale	
8.Baseline characteristics	1/2
9.Therapeutic setting	1/2
10.Dependent variable (target behaviour)	2/2
11.Independent variable (intervention)	2/2
12.Raw data record	2/2
13.Data analysis	2/2
14.Replication	0/2
15.Generalisation	1/2
Total	**18/30**

Appendix B: Ethics Documentation

Re: 2016/1109 The role of a working memory training programme in reducing repetitive negative thinking in older adults

Leaver, Lisa
Thu 28/01/2016, 15:32

Dear Jodie,
Thank you for addressing these points so thoroughly. I have discussed this with the reviewer and we are now happy to approve your project to go ahead.
Best wishes
Lisa Leaver

Lisa Leaver
Senior Lecturer in Animal Behaviour
University of Exeter
Ext: 4641
www.exeter.ac.uk
Washington Singer, University of Exeter, Exeter, EX4 4QG

2016/1109 The role of a working memory training programme in reducing repetitive negative thinking in older adults

Leaver, Lisa
Wed 10/08/2016, 18:08
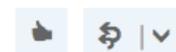
Dear Jodie,
That's fine, I am happy to approve recruitment from other U3A groups using the same methods.
Best wishes
Lisa

Leaver, Lisa and Limond, Jennifer sent automatic replies.

PSYCHOLOGY

College of Life and Environmental Sciences

Study Information Sheet

Does a working memory training programme in reducing repetitive negative thinking in older adults?

WHO IS ORGANISING THIS RESEARCH?

This research is being conducted by Dr Jodie Rawlings, Dr Jenny Limond and Dr Phil Yates at the Mood Disorders Centre at the University of Exeter. The Mood Disorders Centre promotes research, practice and training of benefit to people with mood disorders.

WHO HAS REVIEWED THIS STUDY?

This study has been reviewed and given a favourable opinion by the University of Exeter Ethics Committee.

WHAT THE STUDY IS ABOUT

In this study we are interested in examining whether repeated practice at a computerised working memory task over a 17-26day period improves cognitive control capabilities (as measured by other unpractised tasks) and whether this makes participants less likely to get stuck in negative thinking patterns.

WHAT IS INVOLVED IN TAKING PART?

The study involves participating in three assessment sessions of around an hour each with a member of the research team. Between sessions one and two (for between 10-20days) you will be asked to complete a short questionnaire each day. After the second meeting you would be asked to complete 20-30 minutes of daily practice of the computerised cognitive training task at home for a period of between 17-26days. The assessment sessions can take place at the University of Exeter or in your own home. Everything else outside of this can be completed using your own home computer, so long as you have access to the Internet. If you complete all 36 days of the study then you will be reimbursed £10 for your time and contributions to our research. This will be given to you when you complete the final assessment session.

In each assessment session we will ask you to complete some short tasks designed to measure different aspects of cognitive control. We will also ask you to fill out some brief questionnaires that ask about various aspects of your mood. Finally, you will be asked to give us some details about yourself and any medical history that you may have that could be relevant to the study (for example, if you are taking any medications that might influence your concentration; if you have experienced any mental health difficulties or treatments that might influence how you respond to the cognitive training).

The home training will involve spending 20-30 minutes each day for between 17-26 days practicing a task on your personal computer. This task will be explained to you in detail when you come to the first assessment session, and will be accessible by logging into a secure website which is hosting the training. You will be assisted with registering with the website during the first session, and also have the opportunity to ask any questions about the website, what the training task involves, and how to complete it. Once you have logged in to complete your daily practice, you will be asked to spend approximately 25 minutes completing the training task. Detailed instructions will be given each time you login before beginning the task and you will have the opportunity to ask any questions about this during the first session prior to beginning the training.

WHAT WILL HAPPEN TO THE INFORMATION YOU GIVE?

All the information that you provide will be kept in a secure place and will remain confidential, and you are free not to answer any particular question if you do not wish to do so. Your answers to the questionnaires and all data gathered by the computer will be identifiable only through an ID number (and not your name). No one else will see this data apart from the research team and we will not communicate any of this information to anybody else. Your name and contact details will be stored separately from any personal information that you provide on the questionnaires.

You have the right to withdraw from the study at any time during the study without explanation, and to request that any data you have contributed thus far be withdrawn from the study.

WHAT WILL HAPPEN TO THE RESULTS OF THE STUDY

When complete, the researchers will communicate the results of the study to the wider community of researchers. This is typically achieved through writing up the results in an academic journal, presenting the results at conferences and other outlets. This will NOT involve identification of individuals who took part. You will be given the option to attend a talk summarising the results of the study and their implications.

WHAT IF THERE IS A PROBLEM?

If you have any questions or experience any difficulties then please contact a member of the research team. If you would like to make a complaint, please contact Dr Jenny Limond (J.Limond@exeter.ac.uk).).

WHAT TO DO IF YOU WOULD LIKE TO TAKE PART

Please contact Jodie Rawlings by emailing jr418@exeter.ac.uk if you are interested in participating.

THANK YOU FOR READING THIS INFORMATION SHEET

Further information and contact details
For further information about the project please contact **Dr Jodie Rawlings**

(jr418@exeter.ac.uk) at the University of Exeter, College of Life and Environmental

Sciences, Psychology, Exeter, EX4 4QG. We will be happy to answer any

questions that you might have.

University of Exeter, College of Life and Environmental Sciences

Title of Project: The role of a working memory training program in reducing repetitive negative

thinking in older adults

Name of Investigators: Dr Jodie Rawlings, Dr Anna Adlam, Dr Phil Yates

Healthy Volunteer's Consent Form

Please read this form and sign it once the above named or their designated

representative, has explained fully the aims and procedures of the study to you

- I voluntarily agree to take part in this study.
- I confirm that I have been given a full explanation by the above named and that I have read and understand the information sheet given to me which is attached.
- I have been given the opportunity to ask questions and discuss the study with one of the above investigators or their deputies on all aspects of the study and have understood the advice and information given as a result.
- I agree to comply with the reasonable instructions of the supervising investigator and will notify him immediately of any unexpected unusual symptoms or deterioration of health.
- I authorise the investigators to disclose the results of my participation in the study but not my name.
- I understand that information about me recorded during the study will be kept in a secure database. If data is transferred to others it will be made anonymous. Data will be kept for 7 years after the results of this study have been published.
- I understand that I can ask for further instructions or explanations at any time.
- I understand that I am free to withdraw from the study at any time, without having to give a reason for withdrawing.
- I confirm that I have disclosed relevant medical information before the study.
- I shall receive an inconvenience allowance of £10. If I withdraw from the study for medical reasons not associated with the study a payment will be made to me proportional to the length of the period of participation, but if I withdraw for any other reason, the payment to be made, if any, shall be at the discretion of the supervising investigator.

Name: ………………………………………………………………………………………

Address: ………………………………………………………………………………..

Telephone number: ………………………………………………………………….

Signature: ……………………………………. **Date:** …………………………….

I confirm that I have fully explained the purpose of the study and what is involved

to:

………………………………………………………………………………………………….

I have given the above named a copy of this form together with the information

sheet.

Investigators Signature: ………………………..   **Date**: ………………………………

Investigators Name:………………………………………………………………………………

Study Volunteer Number: …………………………………………………………………..

Appendix C: Discussion of Ethical and Risk Considerations

Older adults are generally considered to be a vulnerable population for the purposes of research (Shivayogi, 2013). However, participants for this study were exclusively recruited through the U3A, increasingly the likelihood of them being capable of providing informed consent. Care was taken to ensure that participants had a full understanding of the implications of engaging in the research and providing consent. This involved both the provision of detailed information sheets and face-to-face discussion with researchers. Participants with an existing diagnosis of a dementia, and therefore potentially compromised ability to provide informed consent, were not be included in the study.

Much of the research was conducted remotely. Participants were provided with an email address and working hours telephone number to contact the researcher if any concerns arose from participation when the researcher was not present (e.g. difficulties with the technology, unexpected emotional distress prompted by engaging in research). Participants were also contacted weekly to monitor distress levels and solve technical difficulties. Remote completion and storage of data electronically also raises questions of confidentiality. Data were stored in an encrypted format, and codes assigned to participants.

Although initial contact with potential participants was made through the U3A, it was made clear that participation was optional, and participants were free to leave the study at any time. The research area was explained to participants, with no deception necessary. RNT, depression and anxiety are emotive topics and there was a possibility that participants become distressed by questions surrounding these areas. The affective content of the training program also had the potential to be triggering for some individuals. Any concerns about distress were discussed with participants at their initial appointment, and they were encouraged to speak with the researcher if they became distressed at any point during participation. This arose on a few occasions during testing, but all participants were fully supported by the researcher and supervisors, and were made aware that they were able to leave the study if they wished. Contact details for relevant support organisations were provided. Risk was also assessed using the PHQ9 at the initial appointment, with the MDC risk protocol followed where necessary.

Identified risk	Management of risk	Level of risk, in light of management
Maintaining confidentiality and anonymity	- Use of participant codes, not names - Storage of names and codes separate from data - Use of pseudonyms/participant codes - Use of password protected computers - Safe storage of data, in lockable cabinets	Low
Breaking confidentiality (e.g., due to risk of self-harm or injury)	- Clear protocols discussed with participants ahead of consent. - Following MDC protocols for risk. - Use of PHQ9 as screening tool. - Any arising discussed with research supervisor	Low
Participant and researcher safety, when seeing people in their own homes	- Application of NHS lone worker policy - Issues discussed with research supervisors	Low-medium
Loss of data	- Researcher ensures two copies of data - Automatic storage of data online.	Low
Suitability and general management of research project	- Trainee supported by research supervisors - Thesis proposal was evaluated for scientific quality and feasibility	Low
Feasibility of project	- Considered by trainee and supervisors during development of project. - Evaluated in the research proposal assessment process	Low
Sufficient resources to conduct research	- Material resources identified as part of research proposal and evaluated for feasibility. - Appropriate consideration has been given to the number of participants required for projects - Research time allocated in the DClinPsy programme.	Low
Health and Safety	- As NHS employees all trainees receive instruction about health and safety procedures - Incidents managed by University health and safety procedures.	Low

Appendix D: Study Measures

TEST YOUR MEMORY
The TYM Test

PLEASE WRITE YOUR FULL NAME…………………………………………..

TODAY IS   ………………..**DAY**

TODAY'S DATE IS THE : …..…. **OF**   ………….**(MONTH)**   **20**……

HOW OLD ARE YOU?   ………………..**YEARS**

ON WHAT DATE WERE YOU BORN?   ……….. / ………… **(MONTH) 19**……

(10)

PLEASE COPY THE FOLLOWING SENTENCE:

GOOD CITIZENS ALWAYS WEAR STOUT SHOES

…………………………………………………………………………………...

PLEASE READ THE SENTENCE AGAIN AND TRY TO REMEMBER IT

(2)

WHO IS THE PRIME MINISTER ?   …………………   …….……………..

IN WHAT YEAR DID THE 1ST WORLD WAR START?…………………………..

(3)

SUMS

20 – 4   =   ………..
16 + 17   =   ………..
8 x 6   =   ………..
4 + 15 - 17 = ……...

(4)

PLEASE LIST FOUR CREATURES BEGINNING WITH "S"
e.g. Shark
1 S……………………
2 S……………………
3 S……………………..
4 S……………………

(4)

WHY IS A CARROT LIKE A POTATO?...

WHY IS A LION LIKE A WOLF? ....

(4)

REMEMBER: GOOD CITIZENS ALWAYS WEAR STOUT SHOES

Please Turn Over

PLEASE NAME THESE ITEMS

1……………..……….	
2……………………..	
3…………………….	
4…………………….	
5……………………..	
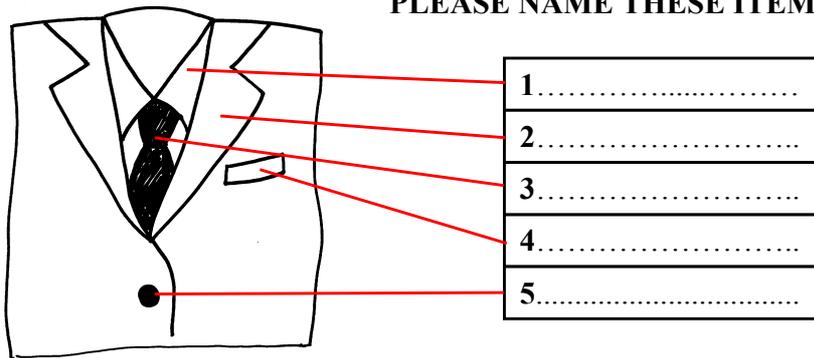
5

PLEASE JOIN THE CIRCLES TOGETHER TO FORM A LETTER (IGNORE THE SQUARES)
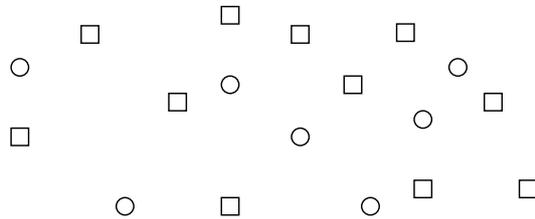
3

PLEASE DRAW IN A CLOCK FACE, PUT IN THE NUMBERS 1 – 12 AND PLACE THE HANDS AT 9.20
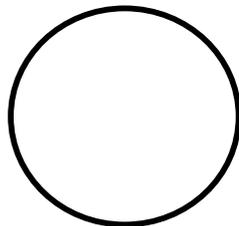
4

WITHOUT TURNING BACK THE PAGE, PLEASE WRITE DOWN THE SENTENCE YOU COPIED EARLIER :

....

6

....

FOR THE TYM TESTER:
HELP GIVEN: NONE/TRIVIAL/MINOR/MODERATE/MAJOR

5

TICK BOX IFANSWERS WRITTEN FOR PATIENT   ☐
www.tymtest.com © jmbrown 2008

/50

PATIENT HEALTH QUESTIONNAIRE (PHQ-9)

NAME:_____ DATE:_____

Over the last 2 weeks, how often have you been bothered by any of the following problems?
(use "✓" to indicate your answer)

	Not at all	Several days	More than half the days	Nearly every day
1. Little interest or pleasure in doing things	0	1	2	3
2. Feeling down, depressed, or hopeless	0	1	2	3
3. Trouble falling or staying asleep, or sleeping too much	0	1	2	3
4. Feeling tired or having little energy	0	1	2	3
5. Poor appetite or overeating	0	1	2	3
6. Feeling bad about yourself—or that you are a failure or have let yourself or your family down	0	1	2	3
7. Trouble concentrating on things, such as reading the newspaper or watching television	0	1	2	3
8. Moving or speaking so slowly that other people could have noticed. Or the opposite — being so fidgety or restless that you have been moving around a lot more than usual	0	1	2	3
9. Thoughts that you would be better off dead, or of hurting yourself	0	1	2	3

add columns [_____] + [_____] + [_____]

(Healthcare professional: For interpretation of TOTAL, please refer to accompanying scoring card).   TOTAL: [_____]

10. If you checked off any problems, how *difficult* have these problems made it for you to do your work, take care of things at home, or get along with other people?	Not difficult at all _____
	Somewhat difficult _____
	Very difficult _____
	Extremely difficult _____

Modified Sternberg shortened for ISSF WM training study

The modified Sternberg task (Oberauer, 2001) has previously been correlated with trait rumination (e.g., Joormann & Gotlib, 2008), and involves working memory updating in the context of negatively valenced stimuli. The modified Sternberg paradigm (Figure 1) presents participants with two lists of three words simultaneously, which they are instructed to remember. One list is presented in red, the other list in blue. The word lists can be of positive or negative, or mixed (positive and negative) valence. A cue is then presented (a red or blue box) indicating which list will be relevant for evaluating the probe word. Finally, a single word (the probe) is presented in black ink and participants are asked to indicate whether this word belongs to the relevant (cued) list. The difference in reaction times (RTs) between decisions to new words and intrusion words (words from the irrelevant list) is interpreted as reflecting the strength of residual activation of words from the to-be-ignored list in working memory.
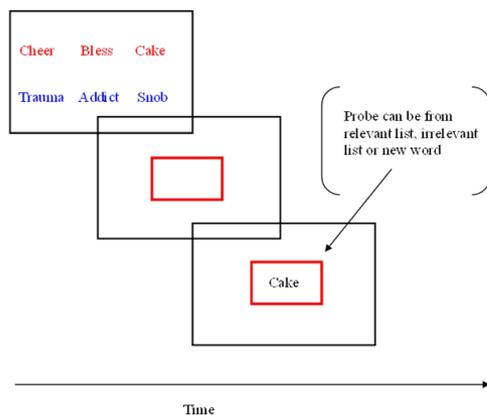
Figure 1: Schematic of the modified Sternberg task

The word lists are presented for 7.8s. A blank screen is then presented for 800ms, followed by a cue displayed for 1s (a red or blue box) indicating which list will be relevant for evaluating the probe. Finally, a single probe word is presented in black and participants are asked to indicate as quickly and accurately as possible whether this word belongs to the relevant (cued) list. The word can come from the relevant list, the irrelevant list, or be a new word that was not presented in either list. The probe word remains on the screen until participants have made their response, using a button press (1 = yes, 2 = no). Participants are instructed to respond as quickly and accurately as possible.

The words lists consist of the positively and negatively valenced nouns selected from the Affective Norms of English Words (Bradley & Lang, 1999) by Joormann and Gotlib (2008). A total of 208 positive words and 208 negative words are used, and these lists were established not to differ in mean arousal rating, or average word length (Joormann & Gotlib, 2008). For each participant, in each block a random sample of words from the word lists is selected without repetition; thus words are only used once within a block, but can be used up to three times across the experiment. The screen position of the red and blue lists and the valence of each colour of list are counterbalanced in each block. Block order and the sequencing of the trials within each block are random.

In the experimental trials, the red and blue lists contain only positive or negative words, and the valence of the two lists differs (i.e., one is positively valenced and one negatively valenced). The valence of the probe word is varied (positive or negative) as is the source of the probe word (relevant list, irrelevant list, or new). Consequently, there are a total of 8 experimental conditions (Table 1). Additionally, a ninth control condition presents lists containing both positive and negative words; this is designed to discourage participants from using the valence of the word lists to cue their responses to the probes.

Table 1

Experimental conditions of Joormann and Gotlib's modified Sternberg task

Condition	Relevant	Irrelevant	Probe
1	Positive	Negative	Relevant
2	Negative	Positive	Relevant
3	Negative	Positive	Irrelevant
4	Positive	Negative	Irrelevant
5	Positive	Negative	New Positive
6	Negative	Positive	New Positive
7	Negative	Positive	New Negative
8	Positive	Negative	New Negative
9	Mixed	Mixed	Positive or Negative

Joormann and Gotlib (2008) found a relationship between rumination and difficulties resolving interference from negatively valenced words amongst depressed individuals (the difference in RT between conditions 4 and 8 above).

The modified Sternberg task as used in this study takes approximately 30 minutes. The critical analysis examines decision latencies to negative intrusion probes and new probes of the same valence (conditions 4 and 8 in the table below). In order to adapt the task to be a suitable length for the original study, the focus owas n the key conditions of interest (conditions 4 and 8) and retained conditions 1, 2, and 9 in order to prevent stimulus valence from biasing responding. Conditions 3, 5, 6, and 7 were therefore removed/

Joormann and Gotlib presented each experimental condition four times in each block and the control condition eight times in each block and included three blocks of forty trials within their task, with a break between each block. Retaining equivalent trial frequencies, the task comprised a total of 28 trials within a block. Two continuous blocks of trials as opposed to three blocks were presented, with breaks between each block. Thus, the adapted task comprised 56 trials in total, as opposed to 120, and did not include additional time between blocks. The task took approximately 14 minutes to complete.

Perseverative Thinking Questionnaire (PTQ)

Instruction: In this questionnaire, you will be asked to describe how you *typically* think about negative experiences or problems. Please read the following statements and rate the extent to which they apply to you when you think about negative experiences or problems.

		never	rarely	sometimes	often	almost always
1.	The same thoughts keep going through my mind again and again.	0	1	2	3	4
2.	Thoughts intrude into my mind.	0	1	2	3	4
3.	I can't stop dwelling on them.	0	1	2	3	4
4.	I think about many problems without solving any of them.	0	1	2	3	4
5.	I can't do anything else while thinking about my problems.	0	1	2	3	4
6.	My thoughts repeat themselves.	0	1	2	3	4
7.	Thoughts come to my mind without me wanting them to.	0	1	2	3	4
8.	I get stuck on certain issues and can't move on.	0	1	2	3	4
9.	I keep asking myself questions without finding an answer.	0	1	2	3	4
10.	My thoughts prevent me from focusing on other things.	0	1	2	3	4
11.	I keep thinking about the same issue all the time.	0	1	2	3	4
12.	Thoughts just pop into my mind.	0	1	2	3	4
13.	I feel driven to continue dwelling on the same issue.	0	1	2	3	4
14.	My thoughts are not much help to me.	0	1	2	3	4
15.	My thoughts take up all my attention.	0	1	2	3	4

Short form Perseverative Thinking Questionnaire (SF-PTQ) used for daily measures:

		never	rarely	sometimes	often	almost always
1.	The same thoughts keep going through my mind again and again.	0	1	2	3	4
2.	I can't stop dwelling on them.	0	1	2	3	4
3.	I get stuck on certain issues and can't move on.	0	1	2	3	4
4.	I keep thinking about the same issue all the time.	0	1	2	3	4
5.	I feel driven to continue dwelling on the same issue.	0	1	2	3	4

Working Memory Training Programme Protocol

Stimulus sets: a list of seven negative words randomly selected from a pool of 35 words matched on length and frequency across conditions (35 negative words taken from the ANEW battery). A new word set is presented on each training session. Words selected for each new session can have been used in previous training sessions but not the session immediately before that day's training (session n-1).

Example words: cruel, dead, fear, grief, lonely, mad, pain

The word sets are ordered randomly (i.e., not using the first letter as an alphabetical cue).

Prior to completing each training session, participants spend 2 minutes rehearsing the stimulus set in order. They then complete two minutes of cued recall where two letters from each word are presented with the words in the correct order and they must type in the letters to complete the words in order. Finally, they are then tested on free recall to the criteria of recalling the correct order of the words with 100% accuracy for 3 consecutive trials.

Input 3 words. Number of steps per trial varies adaptively starting at a baseline of 2 steps. Each step involves on of the following updates:

(1) Recall = recall word in the position indicated.

(2) Transform = +1 or +2 letters working through the first letters of the word set (e.g. dead, +2 = grief)

(3) Substitution = replace previous word with a different word (either presented word or the word in a consecutive location as indicated by the arrow instruction).

At final recall, a single word is presented in one of the three locations and participants are required to indicate whether this is the correct word for this location.

The test word may be:

(a) The correct word in the correct location = yes (40% trials within a block)

(b) The word in that location on step n-3 (lures; trials with >2 steps only; 60% of these trials within a block; because we are training to ignore previously relevant information i.e., 3 = b, and 2 = a)

I The correct word for a consecutive location (only for n=2 trials; with 3 = c, 2 = a)

Each training session begins with 2 practice trials of 2 steps each. The number of steps in a trial varies adaptively. Within a block of 5 trials with n steps each, if 3+ trials are correct, in the next block each trial will have n+1 steps (i.e., must achieve over 50% correct to proceed to next level); if 3+ trials are incorrect, in the next block each trial will have n-1 steps unless n = 2 in which case the number of steps remains at 2. If a participant makes the same response (i.e., all 'yes' or all 'no') for all 5 trials within a block, the next block will have n-1 trials unless n=2 in which case they will remain at n=2.

Participants will complete 10 blocks of trials per training session. Each session will commence at n-1 the highest n correctly achieved on the previous days training (e.g., if the maximum level where 3 or more trials within a block were correct on the previous day was 6, then the next day training would commence at n=5).

Following each day's training a final test of recall of word order for the day's stimulus set will be completed as memory check.

On each trial, subjects are presented with an initial set of three words, each presented in a separate rectangular frame on the screen (in a single row). Thus, subjects always have to remember (and then update) three items. Previous research (cf. Oberauer et al., 2000) has shown this set size to provide an intermediate difficulty level ideal for the study of individual differences, as floor or ceiling

effects are avoided. Switching is held constant by moving to a new (randomly chosen) item on each step.

All updating operations are cued by the display of the appropriate prompt in the frame that is to be updated. Subjects type the result of the update (i.e., by entering the correct word) at every updating step. The updating steps are followed by a test where participants are presented with a single word and must indicate Y/N as to whether it is the correct word in the correct location.

To generate a sequence of steps for a given trial, updating conditions will be selected from the 3 options randomly without replacement; once all 3 have been exhausted this process will repeat until the number of conditions required for the number of steps in that trial have been selected (e.g., select 3 in turn randomly without replacement, then select 2 more in turn randomly without replacement from the original 3 if a given trial requires 5 steps). The selected conditions will then be randomised to generate a sequence with the constraint that there will be no immediate repetitions of any given condition within the sequence.

Conditions involving retrieval require subjects to retrieve the most recent word of the cued frame from memory to perform the current operation. In contrast, this word is provided with the cue in the no-retrieval conditions, meaning that the operation can be executed without retrieval of the previous content of the frame. Transformation conditions involve a transformation of the selected word by alphabet arithmetic. Only positive operations of +1 and +2 are used (with equal probability), and the alphabet is wrapped around such that Z+1=A. Substitution conditions result in the replacement of memory content with new information, and are a component of all conditions as we are targeting updating in the training.

In condition R-Tno-S, subjects are presented with an arrow that linked one frame with another, indicating that they should retrieve and then copy the word from one frame to the other as indicated by the arrow, thus requiring a retrieval and a substitution but no transformation (other than a spatial "relocation").

Each trial is initiated by a key press. The starting words are presented in their frames, all at once, for 2 s. Then the frames are cleared, and the first updating instruction is displayed in a randomly selected frame. Subjects have to carry out the operation required by the instruction, type the result within a specified time limit (10 s), and remember the result as the new content of that frame from there on. If no response is detected within the time limit, an error is recorded and the next step commenced with display of a new instruction in a new frame. No feedback is given after a response, and responses are not echoed on the screen. The instruction for each step is presented immediately after the response of the previous step. After n such steps, final test is indicated by the presentation of a single test word in one of the frames, and subjects are required to indicate whether this is the remembered content for that frame (Y/N). The inter-trial interval is 2.5 s.

The practice session is a supported first training sessions. The session takes approximately 30 min.

There will be a time limit of 10s per frame to enter a response.

The following variables are recorded for each training session:

- RT and accuracy for each step in each trial and what condition the step required

- RT and accuracy for test on each trial; which conditions featured in the trial; how many steps were involved; what the correct response was; whether or not it was a lure trial; and the trial duration.

- Number of blocks completed at each set size.

- Accuracy of recall of the word lists at pre-training testing for each training session

- Length of each training session

- Capacity for the training day (defined as: highest level achieved, i.e., the highest level where 3 or more trials within the block were correct).

WMTP Evaluation Questionnaire

How easy was the training programme to use?				
1	2	3	4	5
Quite difficult		Neither easy nor difficult		Very easy

How useful was completing the training programme?				
1	2	3	4	5
Not at all useful				Very useful

How enjoyable did you find completing the training programme?				
1	2	3	4	5
Not at all enjoyable				Very enjoyable

Any other comments you have about the programme?:

Appendix E: Raw data

Table 1

Key Baseline and Outcome Measures

Participant Number	Age	Gender	Baseline TYM	PHQ	HADS A	HADS D	PTQ	Outcome HADS A	HADS D	PTQ	Eval 1	Eval 2	Eval 3
301	68	0	45	7	11	6	30	11	6	24	1	5	3
302	65	1	49	6	8	13	30	7	12	33	1	2	1
303	73	0	45	5	10	7	33	8	10	31	1	2	1
304	65	0	50	14	19	8	54	14	5	42	4	3	1
305	71	1	45	4	7	0	29	7	1	37	4	4	4
306	77	0	47	15	16	4	36	7	3	22	3	5	4

Table 2

Daily SF-PTQ Ratings

301		302		303		304		305		306	
A	10	A	11	A	13	A	20	A	11	A	12
A	5	A	8	A	12	A	20	A	11	A	6
A	4	A	6	A	11	A	20	A	11	A	5
A	2	A	4	A	12	A	20	A	11	A	8
A	4	A	4	A	13	A	20	A	11	A	8
A	4	A	4	A	13	A	20	A	10	A	15
A	2	A	16	A	12	A	20	A	9	A	9
A	3	A	6	A	12	A	20	A	9	A	3
A	3	A	8	A	15	A	20	A	9	A	15
A	7	A	9	A	10	A	20	A	6	A	9
B	10	A	7	A	6	A	15	A	6	A	7
B	10	A	4	A	10	A	15	A	6	A	10
B	8	A	6	B	14	A	12	A	8	A	11
B	9	A	15	B	6	A	15	A	6	A	7
B	8	A	7	B	10	A	20	B	6	B	10
B	7	A	9	B	8	A	10	B	6	B	5
B	10	B	7	B	8	A	10	B	6	B	10
B	7	B	7	B	10	B	15	B	6	B	12
B	7	B	6	B	10	B	15	B	6	B	18
B	7	B	6	B	10	B	10	B	6	B	18
B	8	B	10	B	10	B	10	B	6	B	3
B	7	B	6	B	10	B	10	B	6	B	9
B	7	B	9	B	10	B	16	B	6	B	8
B	10	B	9	B	14	B	15	B	6	B	15
B	13	B	9	B	10	B	10	B	6	B	4
B	7	B	9	B	10	B	10	B	5	B	4
B	7	B	9	B	12	B	10	B	8	B	1
B	10	B	5	B	11	B	15	B	6	B	2
B	6	B	7	B	10	B	10	B	6	B	2
B	10	B	6	B	10	B	15	B	6	B	2
B	8	B	6	B	11	B	15	B	6	B	2
B	19	B	6	B	10	B	10	B	8	B	6
B	8	B	6	B	10	B	10	B	6	B	4
B	7	B	6	B	10	B	10	B	6	B	5
B	9	B	6	B	10	B	10	B	8	B	5
B	8	B	6	B	10	B	10	B	6	B	1

* A indicates baseline phase, B indicates intervention phase

Table 3

WMTP Performance as Indicated by Maximum Capacity for the Day

Day	P301	P302	P303	P304	P305	P306
1			2	2	5	
2	2	2	2	6	4	
3	2	2	4		5	
4	3	2	7	6	4	
5	4		10	9	5	
6	2	3	13	10	4	
7		4	16	11	3	6
8	2	2	19	11	3	9
9	5	2	6	11	2	12
10		2	9	11	6	15
11		3	12	10	7	6
12	6		6	9	6	9
13	7	3	9	8	6	6
14	8	2	12	8	5	9
15	8	5		9	5	12
16		4	6		5	15
17	2	2	9	6	4	18
18	2		10	7	4	
19	2		13	6	3	6
20	2		16		4	9
21	2		15		5	
22	2		16			
23	2		16			
24	2		15			
25	2					
26	2					

Please note, participants were randomised to varying lengths of WMTP intervention.

Table 4

WM Performance as Measured by the Modified Sternberg Test

Participant	Time 1			Time 2		
	Intrusion cue mean RT	Novel cue mean RT	Difference RTs	Intrusion cue mean RT	Novel cue mean RT	Difference RTs
301	2763	2850.625	-87.625	3056.857	1857.125	1199.732
302	3843.143	2674	1169.143	3437.625	1717.75	1719.875
303	1963.286	1595.875	367.411	1648.143	972.5	675.643
304	3098	1997.125	1100.875	3597.75	1852	1745.75
305	2111.333	2052.125	59.208	2376.857	1564.143	812.714
306	3476.667	2175.125	1301.542	2701.167	1712.857	989.31

Table 5

SPPS Output for Repeated-measures t-test of WM Performance on the Modified Sternberg

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	MSDiffinRTatBaseline	651.7590	6	611.59930	249.68437
	MSDiffinRTatOutcome	1190.5040	6	455.39062	185.91244

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	MSDiffinRTatBaseline & MSDiffinRTatOutcome	6	.543	.266

Paired Samples Test

		Paired Differences							
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper	t	df	Sig. (2-tailed)
Pair 1	MSDiffinRTatBaseline – MSDiffinRTatOutcome	-538.74500	528.25138	215.65772	-1093.1108	15.62082	-2.498	5	.055

Appendix F: Larger Format Multiple Baseline Graphs
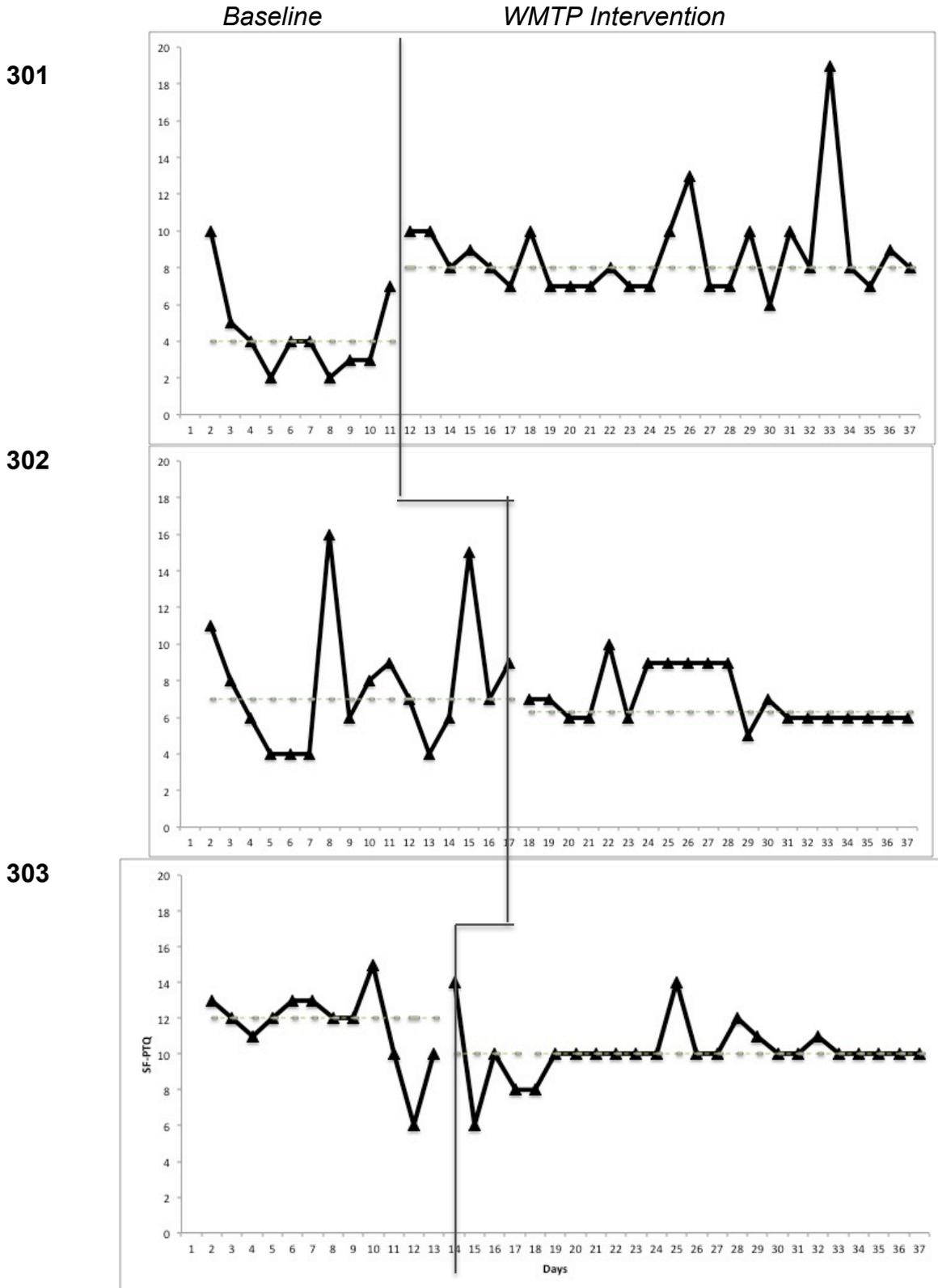
304
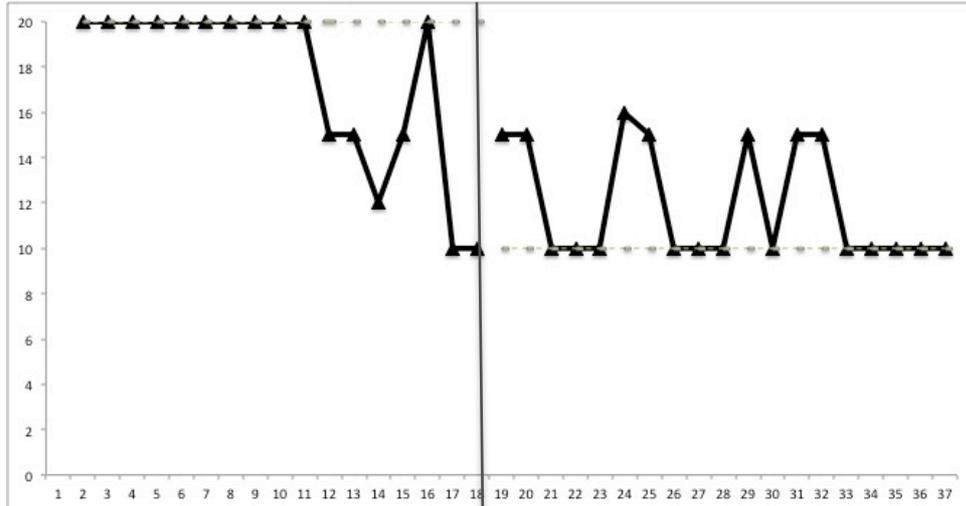
305

306
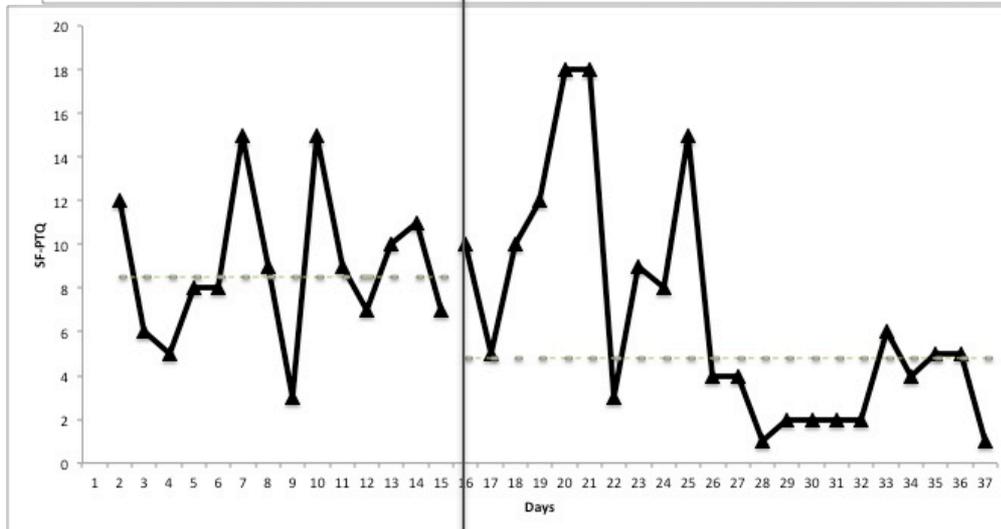
Appendix G: Visual Trend Analysis for Multiple Baseline Graphs

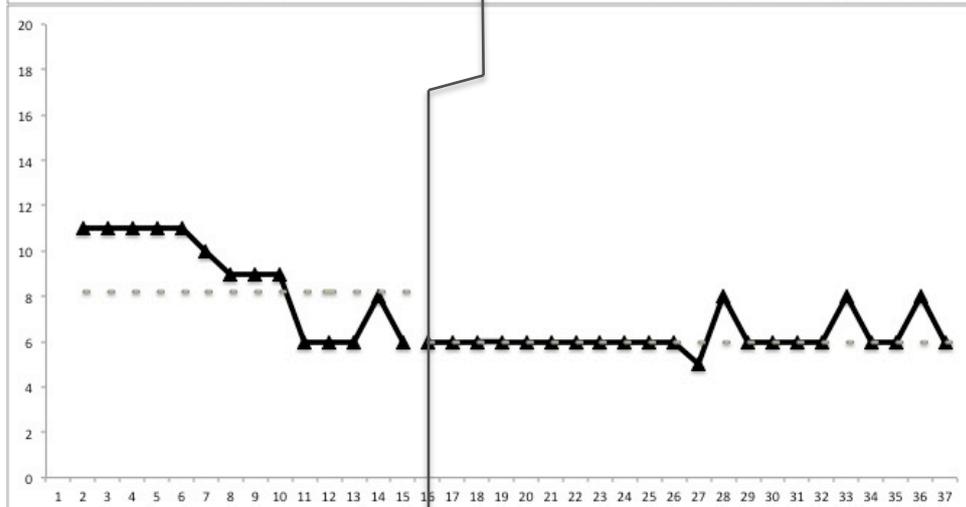*Baseline* *WMTP Intervention*

301 **302** **303** **304** **305** **306**

Split middle trend line

Days
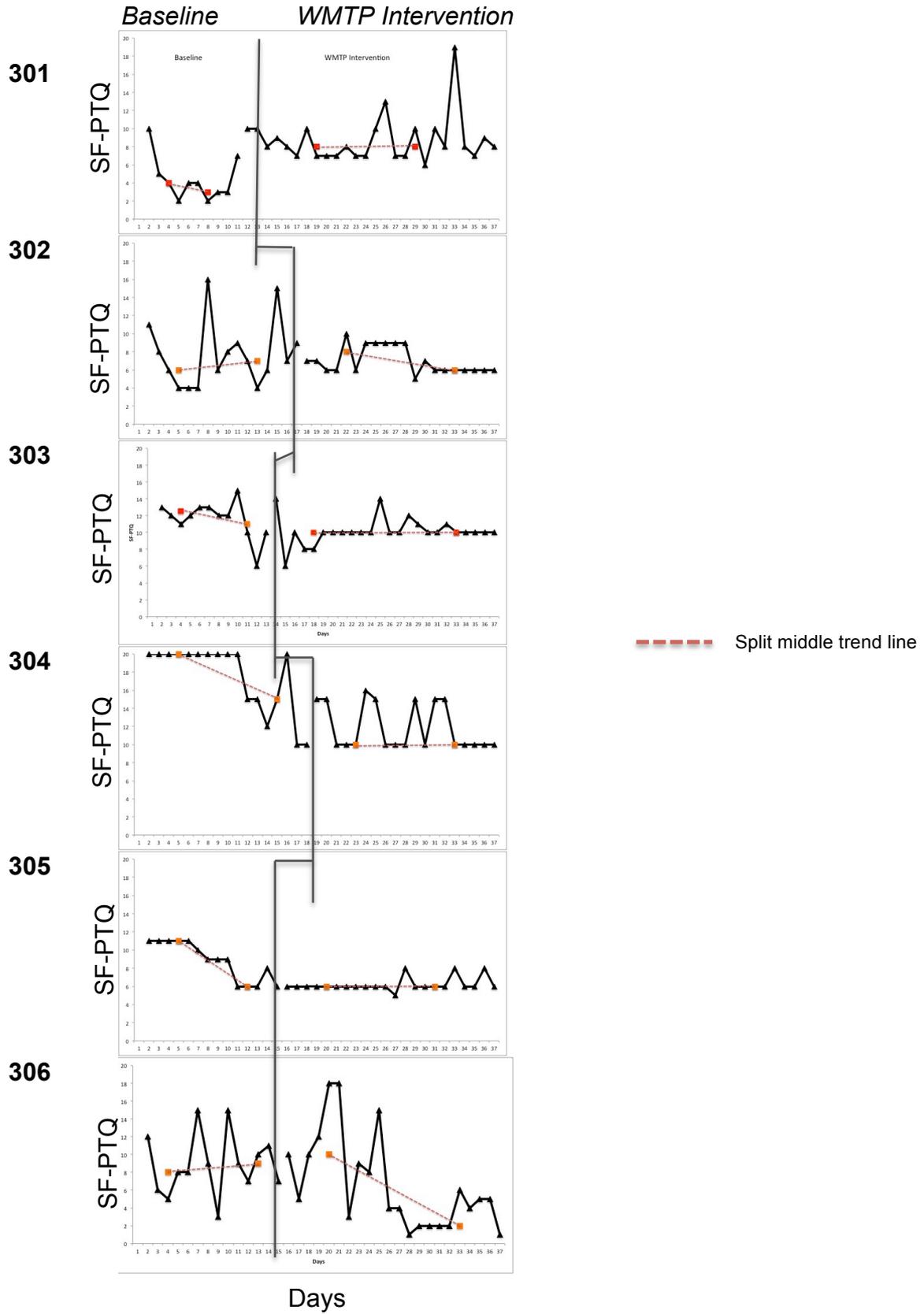
Appendix H: R and SPSS Outputs for SCED Analysis

> library(RcmdrPlugin.SCDA)

Loading required package: SCVA

Loading required package: SCRT

Loading required package: SCMA

Rcmdr Version 2.3-2

Warning message: package 'Rcmdr' was built under R version 3.3.2

> selectdesign(design = "MBD")

[1] 11 11 11 11 11 11

> pvalue.random(design = "MBD", statistic = "A-B", number = 1000)

[1] 0.88

Figure 1: R output for randomisation tests.

ABPhase * Sorted Crosstabulation

			Sorted A	Sorted B	Total
ABPhase	A	Count	46	37	83
		Expected Count	31.9	51.1	83.0
	B	Count	37	96	133
		Expected Count	51.1	81.9	133.0
Total		Count	83	133	216
		Expected Count	83.0	133.0	216.0

(46/83)-(37/133)=0.554 – 0.278 = 0.276

Phi = .28

Figure 2: SPSS output for PAND analysis.

Appendix I: Dissemination statement

The results of this study will be disseminated to interested parties

through feedback, journal publication and presentation.

Dissemination to Participants

Findings will be disseminated to participants and the wider U3A group

through a presentation at a group meeting. Participants will be informed that the

study is complete and will be invited to request individualised feedback with the

researcher if they wish to.

Journal Publication

The study will be submitted for publication with the Journal of Cognition

and Emotion (Impact factor 1.57). See Appendix F for the journal's 'Instructions

to Authors'.

Presentation

On 12th June 2017, my research findings were presented to an

academic audience, for peer review, as part of the Doctorate in Clinical

Psychology at the University of Exeter.

The findings will also be presented to the U3A as described above.

Appendix J: Cognition & Emotion Instructions for Authors

About the journal

Cognition and Emotion is an international, peer reviewed journal, publishing high-quality, original research. Please see the journal's Aims & Scope for information about its focus and peer-review policy.

Please note that this journal only publishes manuscripts in English.

This journal accepts the following article types: Full articles; Brief Articles; Registered Reports of Replication (RRR) studies. The Journal also considers theoretical papers and literature reviews as long as these form a major contribution to our understanding of the interplay between emotion and cognition.

Manuscripts that describe the findings of one experiment should typically be submitted as a Brief Article. The main text of a Brief Article should contain no more than 4000 words and should include a maximum of 2 tables or figures and 25 references.

Registered Replication Reports are manuscripts describing the findings of a study designed to directly or conceptually replicate empirical findings published previously. Unlike the more conventional process where a full report of empirical research is submitted for peer review, RRRs can be considered as proposals for empirical research, which are evaluated on their merit prior to the data being collected. For information on how to prepare Registered Reports of Replication (RRR) submissions see: http://explore.tandfonline.com/page/beh/pcem-registered-reports-of-replication-studies/pcem-rrr-instructions-for-authors.

Peer review

Taylor & Francis is committed to peer-review integrity and upholding the highest standards of review. Once your paper has been assessed for suitability by the editor, it will then be double blind peer-reviewed by independent, anonymous expert referees. Find out more about what to expect during peer review and read our guidance on publishing ethics.

Preparing your paper

All authors submitting to medicine, biomedicine, health sciences, allied and public health journals should conform to the Uniform Requirements for Manuscripts Submitted to Biomedical Journals, prepared by the International Committee of Medical Journal Editors (ICMJE).

Word limits

Please include a word count for your paper.
A typical full article for this journal should be no more than 8000 words; this limit does

not include tables, figure captions, endnotes, footnotes; this limit includes references. A typical brief article for this journal should be no more than 4000 words; this limit does not include tables, endnotes, footnotes, figure captions; this limit includes references.

Style guidelines

Please refer to these style guidelines when preparing your paper, rather than any published articles or a sample copy.

Please use British -ise spelling consistently throughout your manuscript.

Please use double quotation marks, except where "a quotation is 'within' a quotation". Please note that long quotations should be indented without quotation marks.

Formatting and templates

Papers may be submitted in any standard format, including Word and LaTeX. Figures should be saved separately from the text. To assist you in preparing your paper, we provide formatting templates.

Word templates are available for this journal. Please save the template to your hard drive, ready for use.

A LaTeX template is available for this journal. Please save the template to your hard drive, ready for use.

If you are not able to use the templates via the links (or if you have any other template queries) please contact authortemplate@tandf.co.uk

References

Please use this reference guide when preparing your paper. An EndNote output style is also available to assist you.

Checklist: what to include

1. **Author details**. Please ensure everyone meeting the International Committee of Medical Journal Editors (ICJME) requirements for authorship is included as an author of your paper.Please include all authors' full names, affiliations, postal addresses, telephone numbers and email addresses on the cover page. Where available, please also include ORCiDs and social media handles (Facebook, Twitter or LinkedIn). One author will need to be identified as the corresponding author, with their email address normally displayed in the article PDF (depending on the journal) and the online article. Authors' affiliations are the affiliations where the research was conducted. If any of the named

co-authors moves affiliation during the peer-review process, the new affiliation can be given as a footnote. Please note that no changes to affiliation can be made after your paper is accepted. Read more on authorship.

2. A non-structured **abstract** of no more than 200 words. Read tips on writing your abstract.

3. **Graphical abstract** (optional). This is an image to give readers a clear idea of the content of your article. It should be a maximum width of 525 pixels. If your image is narrower than 525 pixels, please place it on a white background 525 pixels wide to ensure the dimensions are maintained. Save the graphical abstract as a .jpg, .png, or .gif. Please do not embed it in the manuscript file but save it as a separate file, labelled GaphicalAbstract1.

4. You can opt to include a **video abstract** with your article. Find out how these can help your work reach a wider audience, and what to think about when filming.

5. up to 5 **keywords**. Read making your article more discoverable, including information on choosing a title and search engine optimization.

6. **Funding details**. Please supply all details required by your funding and grant-awarding bodies as follows:
For single agency grants: This work was supported by the[Funding Agency] under Grant [number xxxx].
For multiple agency grants: This work was supported by the [funding Agency 1]; under Grant [number xxxx]; [Funding Agency 2] under Grant [number xxxx]; and [Funding Agency 3] under Grant [number xxxx].

7. **Disclosure statement**. This is to acknowledge any financial interest or benefit that has arisen from the direct applications of your research. Further guidance on what is a conflict of interest and how to disclose it.

8. **Supplemental online material.** Supplemental material can be a video, dataset, fileset, sound file or anything which supports (and is pertinent to) your paper. We publish supplemental material online via Figshare. Find out more about supplemental material and how to submit it with your article.

9. **Figures.** Figures should be high quality (1200 dpi for line art, 600 dpi for grayscale and 300 dpi for color, at the correct size). Figures should be saved as TIFF, PostScript or EPS files. More information on how to prepare artwork.

10. **Tables.** Tables should present new information rather than duplicating what is in the text. Readers should be able to interpret the table without reference to the text. Please supply editable files.

11. **Equations**. If you are submitting your manuscript as a Word document, please ensure that equations are editable. More information about mathematical symbols and equations.

12. **Units.** Please use SI units (non-italicized).

Using third-party material in your paper

You must obtain the necessary permission to reuse third-party material in your article. The use of short extracts of text and some other types of material is usually permitted, on a limited basis, for the purposes of criticism and review without securing formal

permission. If you wish to include any material in your paper for which you do not hold copyright, and which is not covered by this informal agreement, you will need to obtain written permission from the copyright owner prior to submission. More information on requesting permission to reproduce work(s) under copyright.

Disclosure statement

Please include a disclosure of interest statement, using the subheading "Disclosure of interest." If you have no interests to declare, please state this (suggested wording: *The authors report no conflicts of interest*). For all NIH/Wellcome-funded papers, the grant number(s) must be included in the disclosure of interest statement. Read more on declaring conflicts of interest.

Clinical Trials Registry

In order to be published in a Taylor & Francis journal, all clinical trials must have been registered in a public repository at the beginning of the research process (prior to patient enrolment). Trial registration numbers should be included in the abstract, with full details in the methods section. The registry should be publicly accessible (at no charge), open to all prospective registrants, and managed by a not-for-profit organization. For a list of registries that meet these requirements, please visit the WHO International Clinical Trials Registry Platform (ICTRP). The registration of all clinical trials facilitates the sharing of information among clinicians, researchers, and patients, enhances public confidence in research, and is in accordance with the ICMJE guidelines.

Complying with ethics of experimentation

Please ensure that all research reported in submitted papers has been conducted in an ethical and responsible manner, and is in full compliance with all relevant codes of experimentation and legislation. All papers which report *in vivo* experiments or clinical trials on humans or animals must include a written statement in the Methods section. This should explain that all work was conducted with the formal approval of the local human subject or animal care committees (institutional and national), and that clinical trials have been registered as legislation requires. Authors who do not have formal ethics review committees should include a statement that their study follows the principles of the Declaration of Helsinki.

Consent

All authors are required to follow the ICMJE requirements on privacy and informed consent from patients and study participants. Please confirm that any patient, service user, or participant (or that person's parent or legal guardian) in any research, experiment, or clinical trial described in your paper has given written consent to the inclusion of material pertaining to themselves, that they acknowledge that they cannot be identified via the paper; and that you have fully anonymized them. Where someone is deceased, please ensure you have written consent from the family or estate. Authors

may use this Patient Consent Form, which should be completed, saved, and sent to the journal if requested.

Health and safety

Please confirm that all mandatory laboratory health and safety procedures have been complied with in the course of conducting any experimental work reported in your paper. Please ensure your paper contains all appropriate warnings on any hazards that may be involved in carrying out the experiments or procedures you have described, or that may be involved in instructions, materials, or formulae.

Please include all relevant safety precautions; and cite any accepted standard or code of practice. Authors working in animal science may find it useful to consult the International Association of Veterinary Editors' Consensus Author Guidelines on Animal Ethics and Welfare and Guidelines for the Treatment of Animals in Behavioural Research and Teaching. When a product has not yet been approved by an appropriate regulatory body for the use described in your paper, please specify this, or that the product is still investigational.

Submitting your paper

This journal uses ScholarOne Manuscripts to manage the peer-review process. If you haven't submitted a paper to this journal before, you will need to create an account in the submission centre. Please read the guidelines above and then submit your paper in the relevant Author Centre, where you will find user guides and a helpdesk.

Please note that *Cognition and Emotion* uses Crossref™ to screen papers for unoriginal material. By submitting your paper to *Cognition and Emotion* you are agreeing to originality checks during the peer-review and production processes.

On acceptance, we recommend that you keep a copy of your Accepted Manuscript. Find out more about sharing your work.

Publication charges

There are no submission fees or page charges for this journal.

Color figures will be reproduced in color in your online article free of charge. If it is necessary for the figures to be reproduced in color in the print version, a charge will apply.

Charges for color figures in print are £250 per figure ($395 US Dollars; $385 Australian Dollars; €315). For more than 4 colour figures, figures 5 and above will be charged at £50 per figure ($80 US Dollars; $75 Australian Dollars; €63). Depending on your location, these charges may be subject to local taxes.

Copyright options

Copyright allows you to protect your original material, and stop others from using your work without your permission. Taylor & Francis offers a number of different license and reuse options, including Creative Commons licenses when publishing open access. Read more on publishing agreements.

Complying with funding agencies

We will deposit all National Institutes of Health or Wellcome Trust-funded papers into PubMedCentral on behalf of authors, meeting the requirements of their respective open access (OA) policies. If this applies to you, please tell our production team when you receive your article proofs, so we can do this for you. Check funders' OA policy mandates here. Find out more about sharing your work.

For more information on license options, embargo periods and APCs for this journal please search for the journal in our journal list.