2012

# The Effect of Working Memory (n-back) Training on Fluid Intelligence

David Preece
*Edith Cowan University*

2012

# The Effect of Working Memory (n-back) Training on Fluid Intelligence

David Preece
*Edith Cowan University*

# **Use of Thesis**

# The Effect of Working Memory (*n*-back) Training on Fluid Intelligence

David Preece

A report submitted in the partial fulfilment of the requirements for the award of Bachelor of Arts (Psychology) Honours, Faculty of Computing, Health, and Science.

Edith Cowan University

Submitted October 2011

**I declare that this written assignment is my own work and does not include:**
**(i) Material from published sources used without proper acknowledgment; or**
**(ii) Material copied from the work of other students.**

Signed [redacted]

The Effect of Working Memory (*n*-back) Training on Fluid Intelligence

Abstract

The purpose of this thesis is to report on a study that examined the effect of working memory training (using the *n*-back task) on fluid intelligence (*Gf*). Recent research by Jaeggi and colleagues (2008; 2010) found that training in a visualspatial *n*-back task resulted in gains on two different matrix reasoning tests of fluid intelligence (compared to participants who did no task). The present study replicated and extended these results by testing the fluid intelligence construct using a different type of fluid intelligence test, and employing an 'active' rather than 'no-contact' control group to account for motivational effects on intelligence test performance. Fifty eight participants were involved and their fluid intelligence was assessed pre-training using the Figure Weights subtest from the Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV). Participants were randomly assigned to two groups (experimental or active control), and both groups did a training task on their home computer for 20 days, for 20 minutes a day. The experimental group trained using a single *n*-back task whilst the control group completed general knowledge and vocabulary questions. After training, participants were retested using the Figure Weights subtest. Participants' Figure Weights scores were analysed using an analysis of covariance (ANCOVA). The results of this analysis revealed no significant difference between the training groups in terms of performance on the Figure Weights subtest, suggesting that the *n*-back task was not effective in increasing fluid reasoning ability. These findings were in contrast to those of Jaeggi et al. (2008) and Jaeggi et al. (2010) and suggested that differences between the working memory group and control group found in these studies were likely the result of placebo/motivational effects rather than the properties of the *n*-back task itself.

Author: David Preece

Supervisors: Dr Ken Robinson, Dr Ricks Allan

## COPYRIGHT AND ACCESS DECLARATION

I certify that this thesis does not, to the best of my knowledge and belief:

(i) Incorporate without acknowledgment any material previously submitted for a degree or diploma in any institution;

(ii) Contain any material previously published or written by another person except where due reference is made in the text; or

(iii) Contain any defamatory material.

Signed [redacted]

Dated 22-01-2012

## Acknowledgements

I would like thank my supervisors Dr Ken Robinson and Dr Ricks Allan for their endless patience and expert guidance throughout this project. Without their support, completing this study would have been far more onerous. I would also like to express my gratitude to the many participants who gave up their time to provide the data needed for this study.

Table of Contents

# The Effect of Working Memory (*n*-back) Training on Fluid Intelligence

Contemporary theories of intelligence, such as the Cattell-Horn-Carroll (CHC) theory, divide the concept of general intelligence (*g*) into a number of broad abilities. Two prominent broad abilities within this theory are crystallised intelligence (*Gc*) and fluid intelligence (*Gf*) (McGrew, 2009). Crystallised intelligence refers to cultural knowledge and skills which have been acquired through experience, for example, general facts and vocabulary (McGrew, 2009). In contrast, fluid intelligence can be defined as the ability to reason logically, identify relationships and problem solve in relation to novel stimuli (McGrew, 2009).

There has been significant debate within the literature regarding the nature of intelligence and whether these concepts are fixed or dynamic (Sternberg, 2008). Research has indicated a strong hereditary component in intelligence, and attempts to improve fluid intelligence through training have generally resulted in inconclusive results, implying that fluid intelligence is relatively fixed (Baltes, Staudinger, & Lindenberger, 1999; Jensen, 1981; Sternberg, 2008). However, a number of recent studies have found that training in a spatial working memory task, known as the *n*-back, can transfer to significant performance gains on fluid intelligence tests (Jaeggi, Buschkuel, Jonides, & Perrig, 2008; Jaeggi et al., 2010).

## Importance of Fluid Intelligence (*Gf*)

Fluid intelligence is considered one of the strongest predictors of success in educational and professional domains, and has been found to be a key component in learning (Deary, Strand, Smith, & Fernandes, 2007). Considerable attention has therefore been paid to the possibility of improving fluid intelligence, even in the face of early research by prominent intelligence theorists which suggested the ability was essentially fixed (Jensen, 1981). Having reviewed more than 50 years of research that has aimed to improve the general intelligence of children, Jensen (1981) concluded that clear evidence for genuine improvement was still lacking. Jensen’s work emphasised the difference between scores on IQ tests and one’s actual intelligence ability, noting that whilst a number of studies (Heber & Garber, 1973; Ramey & Haskins, 1981) had demonstrated increases to children's IQ scores, these differences typically failed to extend to practical applications of intelligence such as reading ability. This fact, alongside results which suggested IQ gains faded over time, led Jensen to conclude that this early research displayed only the narrow transfer of skills specific to IQ test performance, rather than true increases to general cognitive ability (Jensen, 1981).

**Improving Fluid Intelligence: Working Memory and Transfer**

One can improve performance on typical tasks which measure fluid intelligence simply by practising the test items (Bors & Vigneau, 2003). However, if the type of task/stimuli is changed, such improvements generally fail to transfer to the new task. As highlighted by Jensen (1981), rather than indicating actual increases in fluid intelligence, such patterns reflect the development of very task specific strategies (Jensen, 1981; Morrison & Chein, 2011).

More relevant to real world practical application is the promise of 'far transfer'. This refers to a situation where training in one type of task transfers to performance gains in another; one which is dissimilar to the trained task (Barnett & Ceci, 2002). As noted by Barnett and Ceci (2002), the concept of far transfer is somewhat ill-defined and subjective; that is, how dissimilar (and in what aspects) do two tasks need to be, in order to be classified as 'far'? Nevertheless, the terms 'near' and 'far' transfer might provide a useful framework upon which to view the effects of training tasks (Barnett & Ceci, 2002). Within the context of the current paper and much of the working memory training literature, it is considered by the author that far transfer could be broadly understood as transfer between tasks designed to measure different constructs (such as working memory and fluid intelligence), as opposed to transfer between tasks which measure the same construct (for example, two different measures of working memory capacity). Far transfer may be presumed to occur when performance on the tasks is governed by the same core underlying processes and mechanisms (Conway et al., 2011).

**Working memory.** Based on this concept of shared underlying processes, recent research has shown that improving fluid intelligence may be possible through the training of working memory (Morrison & Chein, 2011). Working memory, which may be defined as a system for the temporary storage and manipulation of information amidst distraction or concurrent processing, is essential for complex cognition (such as reading and problem solving) and can be understood as a mental workspace with a limited capacity (Conway & Getz, 2010). This limited capacity means there is a restriction on the amount of information which can be maintained within the system (Baddeley, 2003).

According to Baddeley's (2003) seminal theory, working memory functions as a system with multiple components: the phonological loop, the visualspatial sketchpad, the episodic buffer, and the central executive. The phonological loop is a limited capacity system used for the short term storage and processing of verbal and auditory (language) information, whilst the visualspatial sketchpad fulfils a similar role for visual and spatial information. The episodic buffer functions as a multimodal backup store for the linking of information from the phonological loop, visualspatial sketchpad and long term memory (Baddeley, 2000). These three components are known as 'slave' systems, and their function is controlled by the central executive (Baddeley, 2003). The central executive has no storage capacity of its own, rather its function is to coordinate and manage the operation of the slave systems. In this respect, the central executive fulfils a crucial role as it regulates processes such as the control of attention and the suppression of irrelevant information in the working memory system (Baddeley, 2003). Whilst other models of working memory have been proposed (for example, Cowan, 1988), a common theme across all models is the emphasis on the importance of these executive attention processes. It is these domain-general executive functions which have been highlighted by a number of researchers as potential vehicles for successful far transfer to fluid intelligence (Conway et al., 2011).

**The link between working memory and fluid intelligence.** Results from a number of correlational studies have indicated that fluid intelligence and working memory are highly related constructs. For example, Kyllonen and Christal (1990) conducted a factor analysis on a battery of working memory capacity and fluid intelligence/reasoning tests. It was found that correlations between the working memory and reasoning factors were between 0.80 to 0.90. Such findings were confirmed in an analysis of the literature conducted by Kane, Hambrick and Conway (2005). They reanalysed 14 studies and reported a median correlation of 0.72, indicating that the two constructs shared approximately 50% of their variance (Kane et al., 2005).

Further evidence for the link between working memory and fluid intelligence has originated from neurological research, which has suggested that overlapping brain regions are involved in both types of tasks. For example, in a review of the literature, Kane and Engle (2002) concluded that common circuitry in the prefrontal cortex is activated in the facilitation of fluid intelligence and working memory. Working memory training has also been found to impact on structural connectivity within this region, demonstrating training induced plasticity in brain regions implicated with fluid intelligence (Olesen et al., 2004; Takeuchi et al., 2010). Jonides (2004) hypothesised that transfer should occur when overlapping cortical regions are involved in both the trained and transfer tasks, and several neuro-imaging studies have supported this proposal (Dahlin et al., 2008; Olesen et al., 2004).

Overlapping activation of cortical regions and the presence of substantial common variance has led to conclusions that fluid intelligence and working memory may share core processes and mechanisms (Conway et al., 2011; Kane et al., 2005). Research has highlighted a number of executive functions as possible candidates for these common mechanisms, including: updating, control of attention, and the inhibition of interference (Conway et al., 2011). These functions serve to create a more efficient working memory system by ensuring its limited capacity is filled with only relevant information (Conway et al., 2011).

Correlational research by Salthouse et al. (2003) and Friedman et al. (2006) have found these executive functions to be significantly correlated with fluid intelligence. Dempster and Corkill (1999) found that tasks which placed heavy demands on attentional control, but little emphasis on memory, still predicted fluid intelligence; suggesting that the control of attention is an important component of the relationship between working memory and *Gf* . Similarly, Gray et al. (2003) evaluated the relationship between interference resolution and fluid intelligence using an *n*-back task. The *n*-back task is a working memory exercise which requires the participant to indicate a match when the current stimulus matches the one a certain number of times back in a sequence (Jaeggi et al., 2008). When the current stimulus is one away from being the correct number back in the sequence, this is known as a 'lure' trial, as the participant must inhibit a familiarity based response. Gray et al. (2003) found that accuracy on 'lure' trials was a better predictor of *Gf* than accuracy on non-lure trials, suggesting that the ability to inhibit interfering information may be an important component for fluid intelligence test performance. As such, there is some evidence linking fluid reasoning ability to a number of functions vital for the efficient operation of working memory (Bunting, 2006).

One theory behind this linkage is that working memory capacity may constrain fluid reasoning ability. For example, Kyllonen and Christal (1990) proposed that performance on reasoning tasks is dependent upon the ability to maintain in memory representations of stimuli and the possible relationships between stimuli. Hence, working memory capacity may facilitate reasoning by determining the number of representations/relationships which can be remembered and manipulated (Kyllonen & Christal, 1990). Carpenter, Just and Shell (1990) also suggested that individual differences in fluid intelligence test performance are accounted for by one's ability to maintain a large number of abstract relationships and goals in working memory. In a related notion, Oberauer (2005) and colleagues argued that memory requires the binding of features together into representations. There is a limit to the number of bindings which can be maintained and this reflects one's working memory capacity. Similarly, Halford, Cowan and Andrews (2007) proposed that fluid intelligence and working memory share common capacity limitations. Therefore, according to these interpretations, the successful stressing and subsequent extension (or increased efficiency) of working memory capacity should result in gains to fluid intelligence (Jaeggi et al., 2008). This hypothesis is the basis for much of the working memory/fluid intelligence training literature, the key findings of which will be reviewed below.

**Working memory capacity training to influence fluid intelligence.** An increasing number of studies have investigated the effect of working memory training on fluid intelligence. Whilst such studies often report that this training improves performance on similar measures of working memory capacity (near transfer), successful transfer to fluid intelligence (far transfer) has garnered more mixed results (Morrison & Chein, 2011).

One of the most common types of tasks used in working memory research are complex span tasks (Kane & Engle, 2002). Complex span tasks involve maintaining lists of items (such as words, numbers or letters) in memory, whilst also completing a distracting processing activity (for example, doing arithmetic or reading sentences). This concurrent processing removes information from the focus of attention and minimises the use of domain-specific strategies (Engle et al., 1999). In the context of training using working memory tasks, it appears that adapting the difficulty of the task is a crucial aspect. The number of items which need to be held in memory should increase or decrease according to the participant's performance, such that participants are always operating close to their capacity limit and stressing the working memory system (Perrig, Hollenstein & Oelhafen, 2009).

One of the earliest studies to use this approach was conducted by Klingberg, Forssberg and Westerberg (2002; experiment 1). They trained 14 ADHD children (aged between 7-15 years) for 4-5 weeks using adaptive visualspatial span, letter span, backward digit span, and reaction time tasks. The children's fluid intelligence was tested before and after training using Raven's Coloured Progressive Matrices (RCPM). Results indicated that children who trained using the adaptive working memory tasks significantly improved their performance on the RCPM relative to a control group; the control group completed versions of the same tasks, but ones which did not adapt the difficulty based on their performance. These results were replicated in a follow-up study using the same training program (Klingberg et al., 2005).

However, contradictory results were found in a similar study conducted by Holmes, Gathercole and Dunning (2009). In this investigation, 25 children with ADHD (average age 9 years) completed working memory training on the same tasks as those used by Klinberg et al. (2002) over a minimum of 20 days for 20-25 minutes a day. Fluid intelligence was measured via performance on the matrix reasoning and block design subtests from the Wechsler Abbreviated Scales of Intelligence (WASI) and improvement was compared to a control group who completed a non-adaptive version of the training. Although children in the working memory training group demonstrated significant performance gains on tests of working memory capacity following the training, no improvement to fluid intelligence was present (Holmes, Gathercole, & Dunning 2009). A further study by Holmes, Gathercole, Place, Dunning, Hilton and Elliot (2009), again failed to find transfer to the WASI following the training of 25 ADHD children (aged between 8 and 11 years) for 20-25 days using the same working memory tasks. It is important to note transfer was found when fluid intelligence was measured using the RCPM, but not when it was measured using the WASI. This finding underscores the discrepancies between different measures used as proxies for *Gf*, and indicates that gains following working memory training may have been specific to an aspect of the RCPM, rather than true gains to fluid reasoning ability (Jensen, 1981; Sternberg, 2008).

Whilst much early research focused on the benefits of working memory training in children with ADHD, a number of studies have investigated its application to healthy adult populations. For example, Klingberg et al. (2002; experiment 2) tested four college students using the Raven's Advanced Progressive Matrices (RAPM), then trained them using an adaptive span task for an average of 26 days. Following this, participants were retested and were found to display significantly improved scores on the RAPM (Klingberg et al., 2002). However, the results of this study should be considered with caution given the small sample size and the lack of an appropriate control group (the control group was composed of children with ADHD rather than healthy young adults) (Klingberg et al., 2002).

One of the largest scale studies to utilise working memory training was conducted by Schmiedek, Lovden and Lindenberger (2010). In this study, 101 younger (20-31 years) and 103 older (65-80 years) participants completed a cognitive training regimen consisting of 12 tasks (6 perceptual speed tasks, 3 short term memory tasks and 3 working memory tasks). Each participant completed approximately 100 daily 1 hour training sessions. An advantage of this study was that the fluid intelligence construct was measured using multiple tests, and assessed using a latent variable approach to eliminate the error associated with the individual tests (Schmiedek et al., 2010).

The results for each age group were compared to a control group who completed no training activities. For the younger trained group, no improvement in measures of verbal reasoning was found ($d$ = 0.13), but significant improvement was present for numerical reasoning ($d$ = 0.33), figural/spatial reasoning ($d$ = 0.25) and the RAPM ($d$ = 0.33). For the older trained group, significant improvement in reasoning ability was found only on the RAPM ($d$ = 0.54). Investigation of the latent variable underlying all the reasoning measures

(*Gf*) revealed a small but significant improvement on this latent ability for the younger trained group ($d$ = 0.19), but no significant improvement for older participants ($d$ = -0.02). This differential improvement in fluid intelligence between age groups is consistent with research documenting that brain plasticity declines throughout adulthood, and suggests that cognitive training may be more effective for younger age groups (Kramer et al., 2004).

Nonetheless, whilst this result is encouraging in terms of the promise of successful far transfer, an important consideration in interpreting the results of this study is that the control group completed no activity during the training period. The use of a no-contact control group is quite prevalent throughout the literature reporting occurrences of far transfer after training, and presents a significant limitation to the validity of such studies' findings (Morrison & Chein, 2011). Improvements observed in the trained group may be due to differences in expectations and motivation rather than the properties of the training task itself (Morrison & Chein, 2011). For example, the belief that the training should improve cognitive functioning might increase participants' performance on the post-training administration of the tests (Sternberg, 2008). Similarly, participants who invested considerable effort into completing an extensive training regimen may exert more effort during the post-training assessment in order to validate their previous investment (Morrison & Chein, 2011). These motivational influences will be reviewed in greater detail later, but in relation to the results of Schmiedek, Lovden and Lindenberger (2010), they indicate the possibility that improvements to fluid intelligence may have been due to expectancy/motivational effects rather than the training task itself (Sternberg, 2008).

Another comprehensive study to assess the effect of working memory training on fluid intelligence was conducted by Chein and Morrison (2010) and involved 42 university students. Participants in the experimental group were trained 5 days a week over 4 weeks (30-45 minutes a day) using two working memory capacity tasks (one verbal complex span and one spatial complex span). Before and after the training period, participants were tested on a variety of abilities, including: reading comprehension, cognitive control and fluid intelligence. Fluid intelligence was measured using the RAPM, administered using the standard 45 minute time limit (Chein & Morrison, 2010).

Participants improved their performance compared to a no-contact control group on measures of temporary memory (near transfer) as well as cognitive control and reading comprehension (far transfer), however no transfer to reasoning/fluid intelligence was found. Therefore, after 20 days of effortful training using valid working memory tasks, participants displayed no measurable benefits to fluid reasoning ability (Chein & Morrison, 2010).

Other researchers have similarly failed to find transfer to fluid intelligence following extensive training of working memory capacity. For example, Shavelson, Yuan and Alonzo (2008) randomly assigned 37 middle school students (average age 13.5 years) into either a working memory training group or an active control group. Participants in the working memory group trained for 25 days over 5 weeks (30-40 minutes a day) using a battery of 10 adaptive working memory tasks. Control participants watched science themed videos and completed non-adaptive versions of the training tasks. The students' fluid intelligence was tested before and after training using Ravens Standard Progressive Matrices. Overall, whilst participants in the working memory group demonstrated significant improvement to a measure of short term memory, there was no significant difference in fluid intelligence performance between the groups. However, it must be noted that of the 10 training tasks, the majority were simple span measures. These required participants to store information, but did not involve a concurrent processing/distractor task and as such, placed less emphasis on attentional control compared to traditional complex span tasks (Shavelson, Yuan and Alonzo, 2008). Therefore, the lack of transfer in this study may be attributable to the use of training tasks which did not adequately engage the domain-general executive functions presumed to underlie far transfer (Morrison & Chein, 2011).

In line with this assessment, mounting evidence suggests that one of the most critical elements involved in successful far transfer is the type of working memory task used during the training (Jaeggi et al., 2008; Jaeggi et al., 2010). When considering the many different working memory tasks, the one which has produced the most consistent results in respect to transfer to fluid intelligence is the *n*-back task.

***The n-back task.*** The *n*-back is a working memory task which superficially does not resemble typical items on fluid intelligence tests, suggesting that resulting performance gains on these tests are not the result of direct practice effects (Sternberg, 2008). Popular in neuroimaging research, it involves participants attending to a stream of stimuli and indicating a match when the current stimulus is the same as it was *n* times back in the sequence (Kane et al., 2007). The *n* in *n*-back refers to a number which can be adjusted to manipulate the difficulty of the task; hence in '2-back', the participant must indicate a match when the stimulus is the same as it was two times back in the sequence (Jaeggi et al., 2008).

Stimuli in the *n*-back task can be presented in either visual or auditory modalities. In a common variant of the visualspatial *n*-back, participants are presented with a grid and have to keep track of a visual stimulus which moves to another location in the grid every few seconds (Jaeggi et al., 2008). In an auditory *n*-back task, participants hear stimuli (such as letters from the alphabet) read to them. If the *n*-back task involves only one of these modalities, it is known as a single *n*-back task. A more complex variant, known as the dual *n*-back task, follows the same concept but requires participants to simultaneously keep track of a visual stimulus and an auditory stimulus (Jaeggi et al., 2008).

The *n*-back task has been found to involve a number of the executive functions that have been linked to the relationship between working memory and fluid intelligence, including: the processes of attentional control, updating, and the inhibition of interference

(Conway, Kane, & Engle, 2003; Jaeggi et al., 2009; Kane et al., 2007). This is due to the serial presentation of stimuli which requires the constant updating of representations in working memory for relevant stimuli/goals, and the deletion of representations of stimuli which are no longer relevant (items more than *n* times back in the sequence). Similarly, the *n*-back requires attentional control and the inhibition of familiarity responses to interference; particularly on ‘lure’ trials when the current stimuli matches one recently presented, but not *n* times back (for example, in 3-back, a stimuli which matches the one two times back in the sequence) (Oberaurer, 2005). As such, improvement to fluid intelligence following *n*-back training could be seen to support the view that by training these core mechanisms, it is possible to increase fluid reasoning ability (Conway et al., 2011).

***Training using the n-back task.*** Using a dual visualspatial *n*-back task, Jaeggi et al. (2008) trained participants for 25 minutes a day, for a period of 8, 12, 17 or 19 days (training took place on weekdays, with participants having a break on weekends). Based on participants’ performance, the level of *n* in the task adapted automatically in order to keep participants working at their capacity limit. Results from a matrix reasoning fluid intelligence test (the Bochumer Matrices Test [BOMAT]) administered prior to and again after training, revealed that participants who had received the dual *n*-back training significantly improved their scores on the BOMAT compared to a control group who received no training (Jaeggi et al., 2008). Dosage effects were present, whereby the more days a participant did dual *n*-back training, the greater the improvement in fluid intelligence test performance (Jaeggi et al., 2008). Specifically, there was no significant difference between the training and control group after 8 and 12 days of training, but significant differences emerged after 17 and 19 days of the training. This suggests that maintenance of the training over an extended period of time may be required in order for cognitive changes to become concrete (Jaeggi et al., 2008).

Participants in the Jaeggi et al. (2008) study were also administered (before and after training) traditional measures of working memory capacity and short term memory (reading span and digit span tasks respectively) (Jaeggi et al., 2008). Compared to the untrained group, those who trained using the *n*-back improved significantly on the digit span but not on the reading span. Given that the reading span is considered a valid and reliable test of working memory capacity; such results suggest that improvements in fluid intelligence resulting from the *n*-back may not be attributable simply to increases in working memory capacity (Jaeggi et al., 2008; Conway et al., 2011). This unexpected finding was further explored in a follow-up study by Jaeggi and her colleagues.

In this follow-up study featuring two experiments, Jaeggi et al. (2010) replicated and extended the results of their earlier research using both dual (visualspatial and auditory) and single (visualspatial) *n*-back tasks. In the first part of this study, the correlations between a dual *n*-back task, single *n*-back task, OSPAN (a complex working memory span measure) and two tests of fluid intelligence (the BOMAT and RAPM) were investigated. Overall it was found that both the dual and single *n*-back tasks correlated highly with the fluid intelligence measures ($r$ = 0.44/0.53 between the single *n*-back and RAPM/BOMAT, and $r$ = 0.41/0.40 for the dual *n*-back and RAPM/BOMAT), but shared little common variance with the OSPAN task ($r$ = 0.21 for the single *n*-back and OSPAN, and $r$ = 0.26 for the dual *n*-back and OSPAN). Furthermore, multiple regression analysis revealed that the single *n*-back task was the only significant predictor of matrix reasoning performance. Similarly, both the BOMAT and RAPM were significant predictors of single *n*-back performance, whilst the OSPAN was not (Jaeggi et al., 2010).

These findings are in line with earlier research by Jaeggi and colleagues (2009) and Kane et al. (2007) which investigated the psychometric properties of the *n*-back task. In both these studies, it was found that the *n*-back and complex span tasks were only weakly correlated. Conversely, both tasks were highly correlated with a measure of *Gf*, suggesting that the *n*-back and complex span account for independent variance in the fluid intelligence construct (Jaeggi et al., 2010). This lack of a significant correlation between complex span and *n*-back tasks has been attributed to the different retrieval demands of each task (Jaeggi et al., 2009; Kane et al., 2007). Specifically, *n*-back tasks typically require familiarity based recognition processes to identify target stimuli, while complex span tasks demand the active recall of items with no aid from external cues (Kane et al., 2007).

Given that tasks which are highly correlated are assumed to share similar underlying processes (Morrison & Chein, 2011), it was hypothesised by Jaeggi et al. (2010) that training using the *n*-back tasks would transfer to improvements in performance on the fluid intelligence tests but not on the OSPAN task. This assumption was explored in part two of the study, where participants were pre-tested on the BOMAT, RAPM and OSPAN. They then trained using either the single (visualspatial) *n*-back or dual *n*-back task for 20 days (20 minutes per day), before being retested on the fluid intelligence and working memory capacity measures. In line with this hypothesis, both the single and dual *n*-back groups improved significantly compared to the control group (which did no training task) on both measures of fluid intelligence, but showed no transfer to the OSPAN (Jaeggi et al., 2010). These results indicated that the dual task component of the *n*-back (switching between stimuli streams) was not necessary for transfer, as a single (visualspatial) *n*-back task was similarly effective in generating improvement to fluid intelligence (Jaeggi et al., 2010).

**Criticisms of the studies conducted by Jaeggi and colleagues.**

***Administration of the fluid intelligence tests***. Whilst the specific mechanisms underlying the *n*-back training's transfer to *Gf* in the abovementioned studies are not well understood (Morrison & Chein, 2011), it is evident that the *n*-back training resulted in increased performance on the tests of fluid intelligence (RAPM and BOMAT). However, several aspects of these results have been criticised within the literature. For example, Moody

(2009) was concerned that Jaeggi et al. (2008) did not administer the RAPM and BOMAT intelligence tests using the standard time limit (45 minutes), instead allowing only 10 minutes for each test. Moody argued that this time limit did not give participants the opportunity to reach the more difficult items, transforming the test from an assessment of fluid intelligence into a simple measure of how quickly participants could progress through the easiest items. Given that Chein and Morrison (2010) administered the RAPM using the standard time limit and found no improvement on the measure after extensive working memory training (relative to a control group), there is reason for such concern; although Chein and Morrison did not use the *n*-back task in their training. A companion study (being performed by Vaughan Palmer) will seek to address this issue by conducting research whereby participants who complete *n*-back training will be tested using the RAPM with the standard testing procedure.

***Suitable Control Groups***. Another critical limitation of the studies conducted by Jaeggi et al., (2008) and Jaeggi et al. (2010) is that no-contact control groups were used. As such, it is possible that the differences observed between the *n*-back groups and the control groups were due to placebo/expectancy or Hawthorne effects rather than the specific properties of the training task itself (Sternberg, 2008). In other words, a participant's performance may improve simply as a result of receiving attention from the experimenter (McCarney et al., 2007).

The use of no-contact controls is important given recent research demonstrating the substantial impact of motivational influences on intelligence test performance (Carr & Dweck, 2011). For example, after conducting a study of 508 children who completed the Wechsler Intelligence Scale for Children Revised, Duckworth et al. (2010) concluded that some individuals try harder than others and this variance in motivation significantly affects IQ scores. Furthermore, money or personal gain have been found to effect intelligence test performance (Duckworth et al., 2010). This is of particular relevance to the findings of Jaeggi et al. (2010), given that the *n*-back training group received a monetary reward for participation whilst the control group did not.

Moreover, Cury et al. (2006) administered part of the Weschler Intelligence Scale for Children to a group of adolescents and then taught them either that intelligence was fixed or that it could be improved through effort/practice. Following the theoretical instruction, participants taught the malleable theory performed significantly better than those taught the fixed theory. Similar findings have been documented by Mueller and Dweck (1998) and suggest that participant beliefs about the effects of training influence their performance on intelligence measures. Considered in its entirety, such research suggests that there is a strong need for future studies to replicate the results of Jaeggi and colleagues utilising an active control group (Sternberg, 2008).

The issue of using an active control group was partially addressed in a recent study by Jaeggi et al. (2011). In this research, one group of school children (average age of 9 years) were trained using a single visualspatial *n*-back task for 15 (15 minute) sessions over a 4-6 week period, whilst the control group completed a series of vocabulary and general knowledge questions. The children were tested prior to and following training with two matrix reasoning tests of fluid intelligence (Raven's Standard Progressive Matrices [RSPM] and the Test of Nonverbal Intelligence [TONI]). Whilst both groups improved their scores from pre to post testing, there was no significant overall difference between the *n*-back and active control groups.  This result is consistent with the view that improvements in previous studies with no-contact controls may have been due to motivational/placebo effects (Jaeggi et al., 2011).

However, subsequent analysis revealed the importance of individual differences in improvement on the working memory training task. Children in the *n*-back group were separated around the median into either a high gain or low gain group, based on how much they improved on the *n*-back task (Jaeggi et al., 2011). Analysis of these groups revealed that relative to the control group, children in the high gain *n*-back group improved significantly on both the RSPM and TONI (whereas there was no significant difference between the control and low gain *n*-back groups). These results indicate that the *n*-back training did transfer to performance gain on the fluid intelligence tests, but only for those children who demonstrated above median improvement on the *n*-back task (Jaeggi et al., 2011). Nevertheless, further research is needed in order to investigate whether *n*-back training produces consistent and meaningful gains in *Gf* performance when the comparison group is matched to the training group in terms of effort and time investment (Morrison & Chein, 2011).

***Extension to other types of fluid intelligence tests.*** A further issue raised regarding the research conducted by Jaeggi and her colleagues (2008; 2010) concerns the fact that the fluid intelligence construct was only measured using matrix reasoning tests. As emphasised by Sternberg (2008), it is important to investigate whether gains from *n*-back training are also present for other types of fluid intelligence tests. This is needed in order to ascertain whether gains reflect actual increases to fluid reasoning ability, or whether the *n*-back is instead only training an element specific to the solving of geometric matrix problems (Morrison & Chein, 2011). Whilst matrix reasoning tests are generally considered one of the best measures of *Gf*, they are only one proxy for fluid intelligence and are by no means a perfect representation of the construct (Lohman & Lakin, 2011). As such, extrapolating gains on this type of test as representative of gains in true fluid reasoning ability is problematic (Sternberg, 2008). Raven himself understood the limitations of his test as a measurement of the *Gf* construct, and emphasised that his matrices tests should not be administered alone when making decisions about students (Raven, Court, & Raven, 1977).

With these concerns in mind, some researchers have suggested there is reason to believe that gains from the *n*-back task may be unique to matrix reasoning tasks. For example, Moody (2009) has argued that the visualspatial *n*-back task used by Jaeggi et al. (2008), Jaeggi et al. (2010), and Jaeggi et al. (2011) was not entirely different to the BOMAT/RAPM, and instead represented a simplified form of exactly the type of detail required to solve the visual matrix analogies. In this respect, Moody's view is that transfer between the visualspatial *n*-back and geometric matrix reasoning tasks is 'near' rather than 'far'. Specifically, the RAPM is made up of a 3 x 3 matrix with 9 visual figures to keep track of; likewise, the visual *n*-back task required participants to remember the position of a square within a 3 x 3 grid with 8 possible locations. Moody has suggested that doing this task practises the storage of spatial information in working memory, which is the skill specifically required for the efficient solving of the geometric matrix problems found in the BOMAT and RAPM. Therefore, Moody proposed that rather than generating actual improvements in fluid reasoning ability, the visual *n*-back may have only facilitated the development of a very domain/task specific skill. Some support for this hypothesis can be found in the results of Jaeggi et al. (2010), who found that a single *n*-back (visual only) was a superior predictor of matrix reasoning performance than a dual *n*-back task (visual and auditory). Such results could be interpreted as indicating that it is only the visualspatial component of the dual *n*-back task (rather than the domain-general underlying processes) which is vital for improving matrix reasoning test performance (Jaeggi et al., 2010).

## Purpose of the Present Study

The purpose of the present study is to address a number of the key reservations outlined in the earlier literature review regarding the results of Jaeggi and her colleagues (2008; 2010), and the effect of working memory training on fluid intelligence. Specifically, this study aims to determine whether gains from the *n*-back are present on non-geometric matrix reasoning fluid intelligence tests, and whether observed gains to fluid intelligence in previous studies can be accounted for by placebo/Hawthorne effects rather than the working memory activity itself (Sternberg, 2008).

The former issue will be addressed by testing the fluid intelligence construct (before and after *n*-back training) using the Figure Weights subtest from the Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV; Wechsler, 2008). This subtest is a measure of fluid intelligence which requires analogical and quantitative reasoning; such reasoning processes involve inductive and deductive logic which can be expressed mathematically (WAIS IV Technical and Interpretive Manual; Wechsler, 2008). Benson, Hulac, and Kranzler (2010) have reported that the Figure Weights subtest was highly correlated with *Gf* at 0.78. As such, the present study will extend the current body of literature by examining whether performance improvements after *n*-back training generalise to tests of fluid intelligence which do not involve geometric matrix reasoning problems.

The problems associated with using a no-contact control group will be addressed by having the control group complete vocabulary and general knowledge questions for the same amount of time as the *n*-back group's training. In doing so, discrepancies between the experimental and control groups in terms of expectancy and effort invested during training will be reduced (Sternberg, 2008). Therefore, the results of the present study will aid in informing whether gains in fluid intelligence test performance seen in the studies of Jaeggi and colleagues were due to the inherent properties of the *n*-back task itself, or whether improvements can be accounted for by placebo/Hawthorne effects (Sternberg, 2008).

Based on the findings of previous studies (Jaeggi et al., 2008; Jaeggi et al., 2010), it is hypothesised that training in the single visualspatial *n*-back task will result in a significant improvement in participants' scores on the Figure Weights subtest when compared to participants who train using the general knowledge and vocabulary activities.

# Method

## Design

The study employed a 2x2 mixed factorial design. The between-subjects factor, training type, had two levels: *n*-back task and general knowledge/vocabulary task. The within-subjects factor, time of test session, also had two levels: pre training and post training. The dependent variable was the raw score on the Figure Weights subtest (out of 27 items). This design supported an initial testing phase, where participants had their performance on the Figure Weights measured; a cognitive training phase, where participants trained with either the *n*-back task or general knowledge/vocabulary questions; and a second testing phase, where participants' performance on the Figure Weights was retested.

## Participants and Power Analysis

A total of 58 participants volunteered to take part in the study and were either students studying psychology at Edith Cowan University or members of the researcher's social group. (7 participants withdrew from the study during the training period, these were not included in the 58 mentioned above). The age of participants ranged from 18 to 66 years (average of 31.7 in the *n*-back group and 28.2 in the control group). Forty one were currently studying at university and all participants indicated they had completed Year 12 or achieved a higher qualification. Participants were randomly assigned to either the experimental (29 total, 15 females) or active control group (29 total, 13 females).

Effect sizes reported by Jaeggi et al. (2010) demonstrated a Cohen's *d* of 0.65, associated with 20 days single *n*-back training on the RAPM. A power analysis using Lipsey (1990) indicated that 25 participants in each group were sufficient to achieve a statistical power of 0.8.

## Materials and Procedure

**Initial testing phase (intelligence testing).** Ethics approval for this project was granted by the Edith Cowan Human Research Ethics Committee. Upon reading the study's information letter (see Appendix A) and signing a consent form (Appendix B), participants were randomly assigned to either the experimental group or active control group.

***Figure Weights subtest.*** At a venue of their choosing (a quiet room at their home or at Edith Cowan University's Joondalup campus), each participant was individually administered the Figure Weights subtest from the WAIS-IV (using the standard procedure specified in the WAIS-IV Administration and Scoring Manual, 2008).

The Figure Weights subtest is a timed test consisting of 27 items; as well as two demonstration items and one sample item. Each item depicts a series of two sided balance scales with weights on the trays. The weights comprise of various different coloured shapes. The scales on the left hand side of the page have weights on both trays and are balanced. From this information participants are tasked with determining the weights of the various coloured shapes relative to one another. The set of scales on the right hand side of the page, whilst balanced, has weights only on the left tray. Participants are asked to select the group of shapes (there are five answers from which to choose) which would balance the scale when placed on the empty tray (Groth-Marnat, 2009; Wechsler, 2008).

It should be noted that in addition to the Figure Weights subtest, participants were also administered the Raven's Advanced Progressive Matrices (RAPM). This was due to participants in the present study being shared with a companion study conducted by Vaughan Palmer; this investigated the effects of cognitive training on performance on the RAPM. Participants were allowed 5 minutes to complete set 1 (12 items), and 40 minutes to complete set 2 (36 items) of the RAPM.

The Figure Weights and RAPM tests were conducted consecutively; however the order in which the two tests were administered was counterbalanced across participants. Testing took approximately 1 hour (15 minutes for the Figure Weights subtest and 45 minutes for the RAPM); with participants given the opportunity to take a short break between each test.

**Cognitive training phase (training tasks).** Once both intelligence tests had been administered, participants were given instructions regarding the task they would complete during the training phase.

***Experimental group: n-back task.*** The cognitive training task for the experimental group was a variant of the single visualspatial *n*-back task used by Jaeggi et al. (2010). It was obtained as an executable download (for Microsoft Windows) from the program creator's website (Brainworkshop, n.d.) and installed on participants' computers by the researcher. The options in the program were adjusted by the researcher in order to make the task equivalent to that used by Jaeggi et al. (2010). Once the program had been installed, participants were given an information booklet describing the *n*-back task (Appendix C) and the researcher then explained in detail how to complete the activity.

The *n*-back program appeared on a white background within a window sized 24.5cm x 18cm and was comprised of a blue square (3.5cm x 3.5cm) moving around the 8 outer areas of a 3 by 3 grid (11.5cm x 11.5cm). For each movement, the blue square appeared inside a grid slot for 500 ms with 2500 ms periods between movements during which no square was visible. Participants were tasked with pressing the 'a' key on their keyboard to indicate a match when the blue square was in the same location on the grid as it was *n* moves previously. For example, in '2-back', a correct match would be when the square was in the same position as it was two moves ago. When the square was not in the same location as it was *n* moves ago, the correct response was for participants to not press any key. Participants had a 3000 ms time frame following the appearance of a stimuli to indicate a match.

Each round consisted of 20 plus *n* square movements (as such, in '3-back' there would be 23 square movements) with 6 position matches per round (when the stimuli matched the one *n* times back). Ten percent of the time the square movement would be a 'lure' trial; this refers to a situation where the current stimulus matched the one $n + 1$ or $n – 1$ movements ago. Following the completion of a round, participants were given feedback (via an on screen indicator) regarding the number of correct and incorrect responses made during that round. The level of *n* was adjusted based on the participant's performance the previous round. If they gave the correct response for at least 90% of the square movements, the level of *n* was increased by 1 for the next round (for example, progressing from '2-back' to '3-back'). Scores between 75% and 89% correct for a round resulted in the level of *n* being maintained, and scoring below 75% resulted in the level of *n* being decreased by 1. In this way, the difficulty of the task adapted to the participant's level of proficiency (Jaeggi et al., 2010). For each new session the participant started at the default level of '2-back', regardless of their performance in previous sessions. These settings and adaptive parameters were the same as those used by Jaeggi et al. (2010) for their single *n*-back training task.

In order to ensure understanding of the task, the researcher completed one round of '2-back' whilst the participant observed. Following this, the participant completed a round of '2-back' themselves, and were encouraged to continue completing rounds until they felt confident doing the activity. The next day, the participant's training commenced. One daily session consisted of 15 rounds of the *n*-back, taking approximately 17-20 minutes. The total training period was 20 sessions over 30 days, one session per day. Participants were able to choose which days they completed sessions during the training period, however it was recommended that the most effective way to complete the training was to do sessions for five consecutive days, then have a two day break.

Participants completed the training independently on their home computer. They recorded their average and highest *n*-back level for each session on a page provided by the researcher (Appendix E). On completion of the 20 sessions, the information was given to the researcher so that performance on the task over time could be analysed.

***Active control group: Vocabulary and general knowledge tasks.*** The active control group rotated through two different tasks for a period of 20 minutes, completing one round of a task, then moving on to the other. The two tasks were based on building vocabulary and general knowledge and were accessed by participants through a web browser on their home computers.

The vocabulary task was an activity known as 'Define Time' (East of the Web, n.d.) and the general knowledge activity was presented in the format of the 'Who wants to be a millionaire' television show (Real Player Games Directory, n.d.). Following the completion of the initial intelligence tests, participants in the active control group were given an information sheet which explained how to complete the vocabulary and general knowledge activities, and the web addresses to access them (Appendix D).

*Vocabulary task (Define Time).* In the 'Define Time' task (East of the Web, n.d), participants were presented with a word at the top of the screen and four possible definitions. Participants made their selection by clicking on a definition with their mouse. If a participant selected the correct definition for a word, 10 points was added to their score and they progressed to another word. If participants selected the incorrect definition, they lost 5 points and were required to select again until they chose the correct definition. The task continued to present participants with new words until the two minute time limit expired, upon which participants were shown their score for that round. Following the researcher's explanation and completion of one round of 'Define Time', participants were given the opportunity to complete a round of the activity as a practice.

*General knowledge task.* The second activity was then explained by the researcher. This training activity was presented in the format of the 'Who wants to be a millionaire' television show, whereby participants were required to answer 15 questions of increasing difficulty in order to progress their score from $100 to $1,000,000. For each question, participants were shown four possible answers and had to click the correct answer with their mouse within a 30 second time limit. Three ‘lifelines’ to help participants were provided in the program, including: removing two of the incorrect answers, polling a virtual audience, or phoning a friend (where the computer would suggest an answer and indicate how sure they were out of 100 percent). The round concluded when participants answered a question incorrectly. Five questions were demonstrated by the researcher; showing the participant how to select an answer and how to use each of the three lifelines by clicking the icons presented on the screen.

Active control group participants completed 20 daily sessions of the vocabulary and general knowledge activities. A daily session involved alternating between rounds of ‘Define Time’ and ‘Who wants to be a millionaire’ for a period of 20 minutes (participants started every session with ‘Define Time’). On a sheet provided by the researcher, participants recorded the highest score they achieved in a single round for each of the activities during the daily session (Appendix F). This record sheet was then given to the researcher at the completion of training.

**Post-training retest phase.** Within three days of the completion of the 20 daily sessions of cognitive training, participants were retested on the Figure Weights subtest and RAPM. The order in which the tests were administered was consistent with the pre-training administration for that participant, and the tests were conducted in the same location.

## Results

PASW 18 Statistics package for Windows was used to assess participant performance on the training tasks and on the Figure Weights subtest.

### Data Screening

Participants’ pre-test scores on the Figure Weights subtest, post test scores, and the amount of improvement from pre to post testing was explored for outlying cases. Analysis of box plots revealed two outlying cases (more than 1.5 box lengths above the 75$^{th}$ percentile) in the control group in terms of gain on the Figure Weights subtest (both these cases improved by seven items from pre to post). One of these cases was also identified as an outlier on pre-test score (performing worse than the rest of the group). This case reported to the researcher that they had difficulty concentrating during the pre-test administration due to a toothache after recent root canal therapy. These participants were excluded from the analysis, leaving 27 participants in the control group and 29 in the *n*-back group. Unless otherwise specified, the following results do not include these cases, however each of the main analyses was re-run with these cases included and the results (significance evaluated at $\alpha = .05$) did not change.

**Figure Weights Subtest Performance**

**Split plot analysis of variance**. Initial analysis with a split plot analysis of variance (SPANOVA) was used to assess whether participants improved their performance on the Figure Weights subtest from pre to post administration. Shapiro-Wilk's test of normality indicated that pre-test scores for both the *n*-back and control groups were normally distributed ($p > .05$), but post test scores were not ($p < .05$). However, inspection of box plots indicated that post-test scores were approximately normally distributed and the data was not transformed. Box's M statistic ($p > .001$) and Levene's test for homogeneity of variance were not significant ($p > .05$), indicating that these assumptions were met.

The analysis revealed that overall, participants' post-test scores were significantly higher than their pre-test scores, $F(1, 54) = 9.371$, $p = .003$, partial $\eta^2 = .148$ (See Table 1). This indicates that both training groups significantly improved their performance on the Figure Weights subtest. In order to confirm these results, the analysis was re-run with the two outliers entered back into the sample. The results remained significant, $F(1, 56) = 12.152$, $p = .001$, partial $\eta^2 = 0.178$.

Table 1.

*Means and standard deviations of pre and post-test scores on the Figure Weights subtest for participants in the n-back and control groups. (Outliers removed)*

| Training type | Pre-training FW score | | Post-training FW score | |
|---|---|---|---|---|
| | *M* | *SD* | *M* | *SD* |
| *n*-back | 18.97 | 4.17 | 20.45 | 4.43 |
| control | 19.70 | 3.77 | 20.33 | 3.68 |
| Total | 19.32 | 3.96 | 20.39 | 4.04 |

**Analysis of covariance**. A one way analysis of covariance (ANCOVA) was used to compare the post-training Figure Weights scores of participants undertaking the two different training programs. Participants' pre-training scores on the Figure Weights test were included as a covariate to partial out the effect of participants' initial performance on the Figure Weights test prior to the training.

The assumption of normality was violated for post-test scores in the *n*-back and control groups (Shapiro-Wilk, $p < .05$). However, inspection of histograms indicated that the distributions were approximately normal. Additionally, the Shapiro-Wilk test indicated that the covariate (pre-test scores) did not violate the normality assumption ($p > .05$). As the ANCOVA is considered robust against moderate violations of normality as long as the covariate is normally distributed; no transformation was applied to the data (Allen & Bennett, 2008).

Inspection of scatter plots revealed that there was a linear relationship between the DV and covariate for both the *n*-back and control groups. The interaction between the covariate (pre-test scores) and the training group was not significant, $F(1, 53) = 0.32$, $p = .574$, partial $\eta^2 = 0.006$, indicating that the homogeneity of regression slopes assumption was not violated. Similarly, Levene's test confirmed the assumption of homogeneity of variance was met $F(1, 54) = 3.307$, $p = .075$. Analysis showed that the covariate (pre-training test scores) was significantly related to the dependent variable (post-training test scores), $F(1, 53) = 88.41$, $p < .001$, partial $\eta^2 = 0.625$.

The results of the ANCOVA revealed that after controlling for pre-training scores, post-training scores on the Figure Weights subtest was not significantly related to the type of training the participant completed (*n*-back or control), $F(1, 53) = 1.116$, $p = .296$, partial $\eta^2 = 0.021$ (see Table 2). When the analysis was re-run with the two outlying cases added back into the control group, the difference between the groups remained non-significant, $F(1, 55) = 0.410$, $p = .525$, partial $\eta^2 = 0.007$.

Table 2.

*Estimated marginal means of Figure Weights post-test scores calculated for the covariate (pre-test scores) at 19.32.*

| | Post-test Figure Weights score | |
|---|---|---|
| Training type | *M* | *SE* |
| *n*-back | 20.73 | 0.46 |
| Control | 20.03 | 0.48 |

***Dividing groups into high and low gain.*** Given that Jaeggi et al. (2011) found that transfer to *Gf* was dependent on individual differences in improvement on the *n*-back task, the impact of training task improvement on Figure Weights performance was investigated. Participants in each group were split around the median into either high or low gain groups, depending on how much they improved on the training tasks over the training period. This method was used as it was the same procedure Jaeggi et al. (2011) used to explore their data. As the control group involved completing two different tasks, participants were split around the median based only on the vocabulary activity.

A one way ANCOVA was conducted to compare the post-training Figure Weights scores of participants in the high gain *n*-back group, the low gain *n*-back group, the high gain control group and low gain control group. Pre-test scores on the Figure Weights subtest were included as a covariate to control for the effect of initial performance on the fluid intelligence test. Shapiro-Wilk's statistic indicated that all the groups' pre-test scores were normally distributed ($p > .05$). Similarly, all their post test scores were normally distributed ($p > .05$) except for the high *n*-back gain group ($p < .05$). However, the covariate was normally distributed so no transformation was applied. The interaction between the covariate and gain group was not significant, $F(3, 48) = 0.156$, $p = .925$, indicating the homogeneity of regression slopes assumption was not violated. Scatter plots revealed the relationship between the DV and covariate was linear for each of the groups. Similarly, Levene's test indicated the assumption of homogeneity of variance was met, $F(3, 52) = 1.324$, $p = .276$.

The analysis revealed no significant difference in post-training Figure Weights scores between the groups $F(3, 51) = 0.609$, $p = .612$, partial $\eta^2 = 0.035$. These results indicate that regardless of the degree to which participants improved on the training task, participants in the *n*-back group did not improve their performance on the Figure Weights subtest significantly more than the control group. Indeed, contrary to the results found by Jaeggi et al. (2011), the estimated marginal means of the low gain group for each task type were greater than those of the high gain group, although this difference was not significant (see Table 3).

Table 3.

*Estimated marginal means of Figure Weights post-test scores calculated for the covariate (pre-test scores) at 19.32.*

| | Post-training FW score | |
|---|---|---|
| Training group | *M* | *SE* |
| Low gain *n*-back | 21.11 | 0.65 |
| High gain *n*-back | 20.34 | 0.67 |
| Low gain control | 20.16 | 0.67 |
| High gain control | 19.88 | 0.70 |

**Training task performance**

***N*-back group.** Performance on the *n*-back task was assessed by using a repeated measures analysis of variance (ANOVA). The mean of participants' average *n*-back level during the first two sessions was used as a measure of starting *n*-back performance, and the mean of their average *n*-back level for sessions 19 and 20 was used as a representation of their *n*-back performance at the end of the training period.

Shipiro-Wilk's test of normality indicated that participants' end of training *n*-back scores were normally distributed ($p > .05$), however, start of training *n*-back scores violated the normality assumption ($p < .05$). Inspection of histograms indicated that the data was approximately normally distributed, although with a slight positive skew. Homogeneity of variance was calculated using the $F_{max}$ test. The $F_{max}$ value was less than 10 (1.86) indicating this assumption had been met.

The ANOVA revealed participants' performance at the end of the training period (*M* =

5.45, *SD* = 1.23) was significantly better than their performance at the start of the training period (*M* = 3.56, *SD* = 0.90), $F(2, 14) = 73.84$, $p < .001$, partial $\eta^2 = 0.841$ (sphericity assumed). This indicated that the average *n*-back level participants were operating at increased over the training period (see Figure 1).

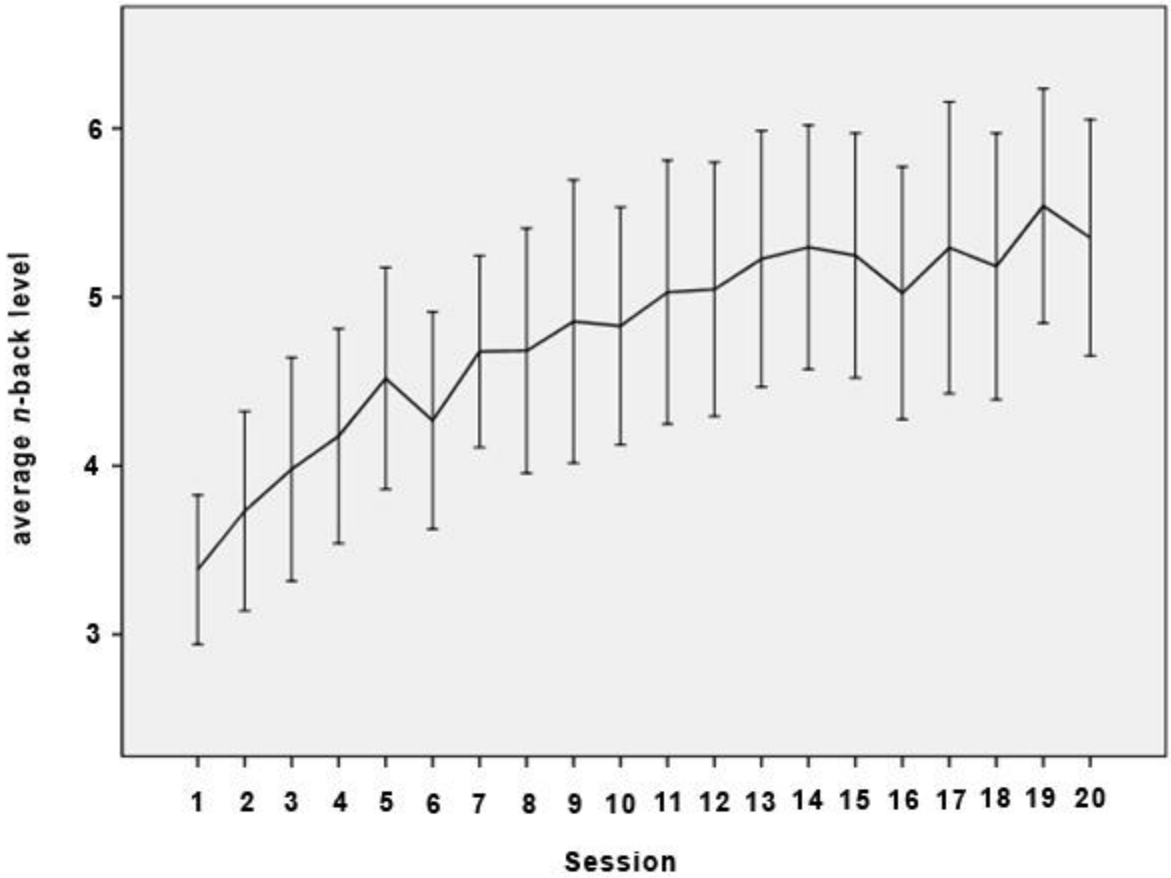



*Figure 1*. Average *n*-back level participants were operating at during each session over the training period. Error bars reflect 95% confidence interval. Note: In plotting their data, Jaeggi et al. (2010) excluded the first 3 rounds of a session as 'warm-up' trials, meaning that low starting levels of *n* did not contribute to the overall average for the 15 rounds. In the present study this was not done; the Brainworkshop program automatically calculated the average for a session based on all rounds (including the first 3). As such, this accounts for why the average level of *n* was slightly lower than that documented by Jaeggi et al. (2010).

**Active control group.** Performance of control group participants on the vocabulary task was also assessed using a one way repeated measures ANOVA. The average of their highest score for the first two sessions of ‘Define Time’ was used as a measure of initial vocabulary performance, and their average for the last two sessions was used to measure performance at the end of the training period.

The results of the ANOVA indicated that performance on the vocabulary task was significantly improved at the end of the training period (*M* = 415.95, *SD* = 155.72) compared to the start (*M* = 232.67, *SD* = 77.65), $F(1, 28) = 74.831$, $p < .001$, partial $\eta^2 = 0.728$ (see Figure 2).

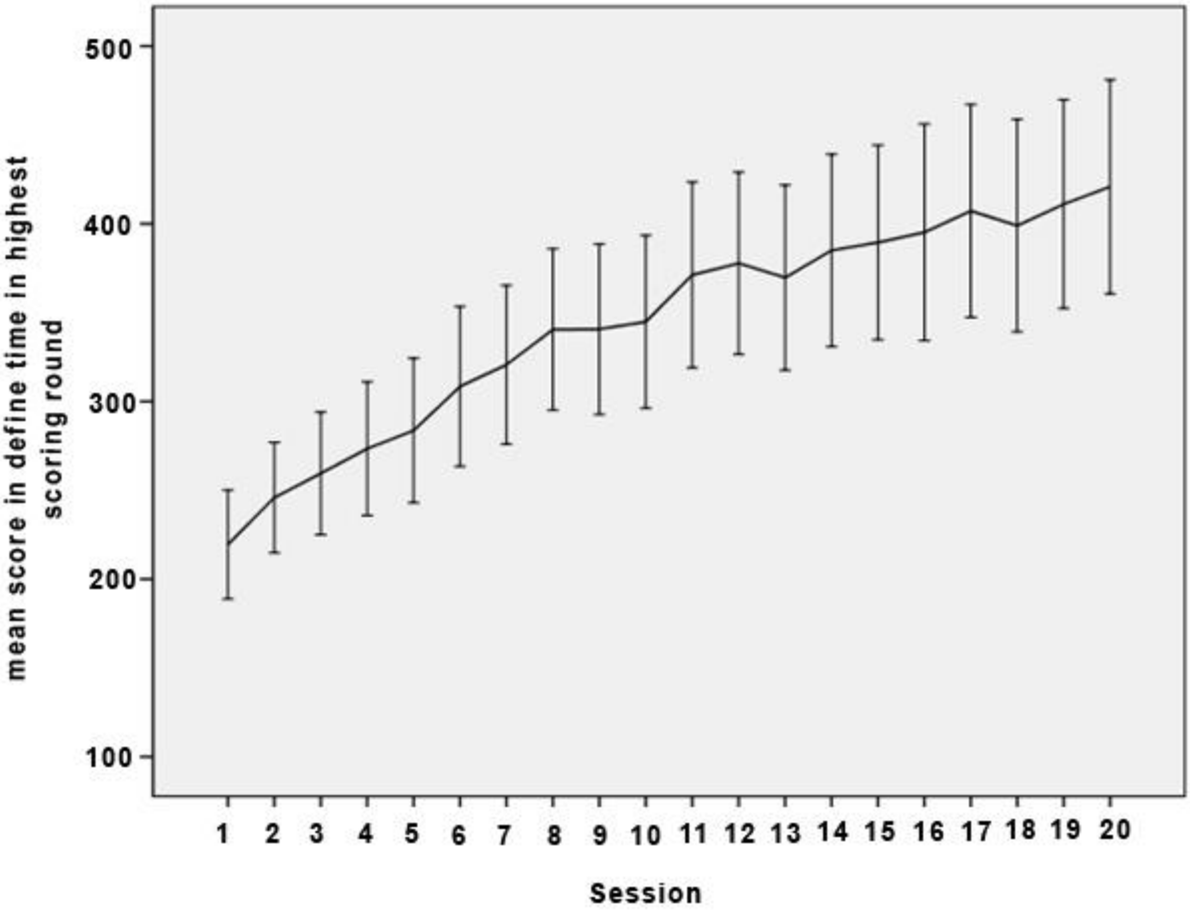



*Figure 2*. Highest point total reached by participants during a round of ‘Define Time’ (the vocabulary task) for each session over the training period. Error bars reflect 95% confidence interval.

The same method was used to calculate participant improvement on the general knowledge task. This analysis revealed that participants performed significantly better on the general knowledge task at the end (*M* = 10.05, *SD* = 1.92), compared to the start of the training period (*M* = 6.24, *SD* = 2.17), $F(1, 28) = 78.749$, $p < .001$, partial $\eta^2 = 0.738$ (sphericity assumed) (See Figure 3).

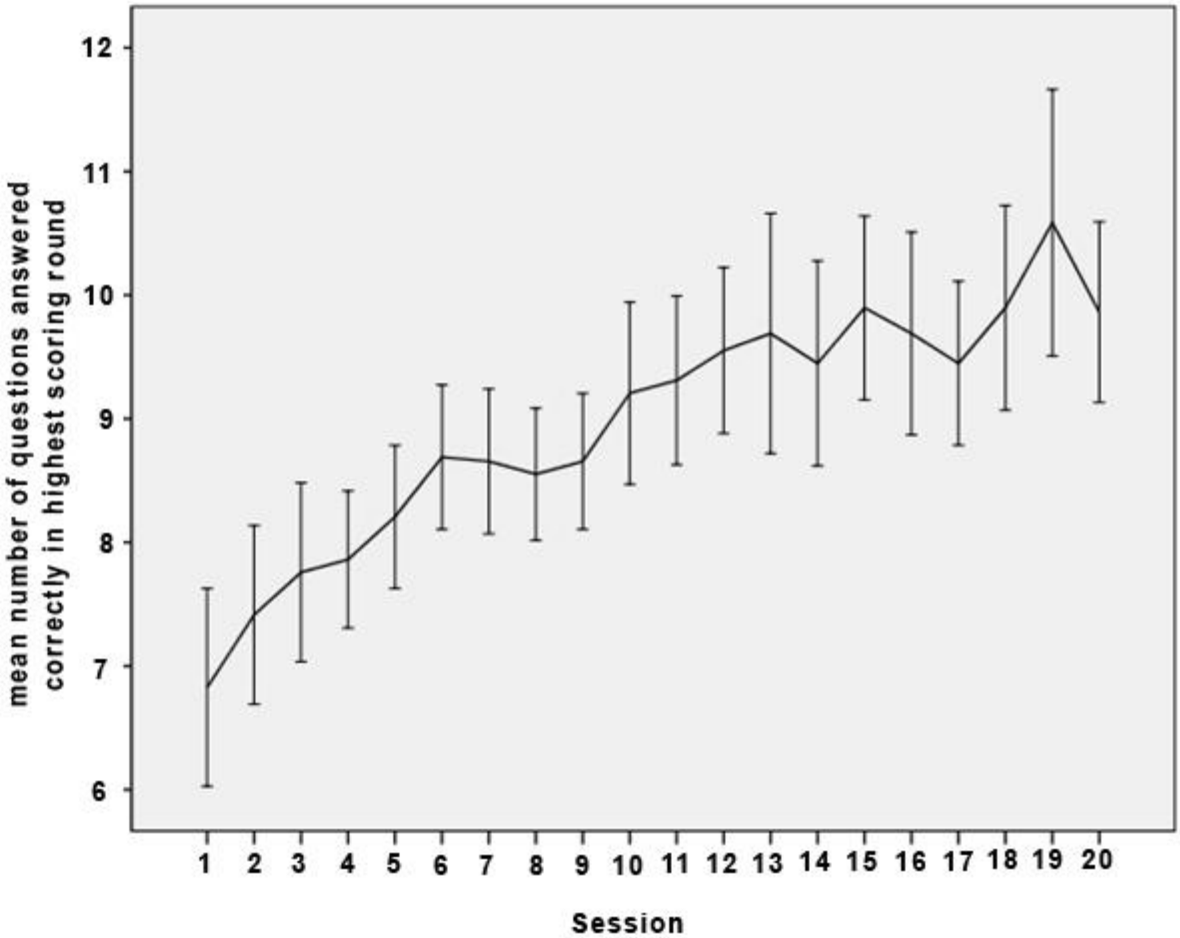



*Figure 3.* General Knowledge performance over the 20 sessions (highest number of questions correctly answered in a round). Error bars reflect 95% confidence interval.

**Regression Analysis**

Lastly, a standard multiple regression analysis was conducted on participants in the *n*-back group in order to see how well gain on the *n*-back task, age, and gender predicted improvement on the Figure Weights subtest.

Shapiro-Wilk's test indicated that amount of gain on the *n*-back task was normally distributed ($p > .05$). Conversely, the distribution of age was not normal (Shipiro-Wilk, $p <$

.05), however skewness (0.435) and kurtosis (0.845) values indicated age was approximately normally distributed. The maximum Mahalanobis distance value (9.917) did not exceed the critical value for *df* = 3 (16.266), indicating that multivariate outliers were not an issue in the sample.

Inspection of the normal P-P plot of regression standardised residuals, and the scatter plot of standard residuals against standardised predicted values, indicated that the assumptions of normality, linearity and homoscedasticity of residuals had been met. Collinearity tolerance statistics were greater than 0.02 for all variables, suggesting multicollinearity was not an issue for the predictor variables.

The results of the standard regression analysis showed that in combination, participant gender, age and *n*-back gain accounted for 10.1% of the variance in Figure Weights score improvement. The variance accounted for by these predictors was not significant, $F(3, 25) = 0.935$, $p = .438$, $R^2 = .101$, adjusted $R^2 = -.007$. For each predictor, unstandardised and standardised regression coefficients, and squared semi-partial correlations are presented in Table 4.

Table 4.

*Unstandardised (B) and Standardised (β) Regression coefficients, and Squared Semi-Partial Correlations ($sr^2$) for Each Predictor in a Regression Model Predicting Improvement on the Figure Weights Subtest from Pre to Post Administration.*

| Variable | *B* | β | $sr^2$ |
|---|---|---|---|
| *n*-back gain | .905 | .336 | .094 |
| age | .007 | .035 | .001 |
| gender | .260 | .047 | .002 |

*$p < .05$

These results suggest that the amount a participant gained on the *n*-back task, their age, and their gender, could not account for a significant amount of the variance in their improvement on the Figure Weights subtest.

## Discussion

It was hypothesised that participants who trained using the *n*-back working memory task would significantly improve their performance on the Figure Weights subtest, relative to participants who did a control task. The results of the current study did not support this hypothesis. Statistical analysis indicated that participants improved significantly from pre to post administration. However, there was no significant difference between the *n*-back and control groups in terms of improvement on the measure of fluid intelligence, and even those participants who demonstrated the highest degree of improvement on the *n*-back task did not display significant transfer to Figure Weights performance (compared to the active control). Such findings suggest that the *n*-back working memory task was not a successful training activity in terms of increasing one's fluid reasoning ability.

These results are in contrast to those of Jaeggi et al. (2008), Jaeggi et al. (2010) and Jaeggi et al. (2011) who found the *n*-back task improved performance on tests of fluid intelligence relative to controls. This raises several questions: Why was no transfer found in the present study? What is responsible for this contrasting finding? In order to answer such questions, three key aspects of the present study will be evaluated: the nature of the training task, the way fluid intelligence was measured, and the composition of the control group. These aspects will be addressed in turn throughout the following sections, and the resulting implications for the literature will be discussed.

### The *n*-back Training Task

Firstly, it must be explicitly understood that the properties of the single visualspatial *n*-back task used in the present study were identical to those of the *n*-back task used by Jaeggi et al. (2010). The number of square movements per round (20 + *n*), the number of rounds per session (15), the duration of each session (approximately 20 minutes) and the total number of sessions completed (20) were the same as in the training paradigm used by Jaeggi et al. (2010). The thresholds for increasing, maintaining and decreasing the level of *n* were identical, as was the amount of interference stimuli (*n* + 1 or *n* – 1) (10 percent). Parallel to the procedure used by Jaeggi and her colleagues (2010), participants started each new session from the default level of ‘2-back’, and the total training period was approximately 1 month. As such, the implementation of the *n*-back task in the present study was equivalent to that of the *n*-back training used by Jaeggi et al. (2010).

The lack of transfer to fluid intelligence in the present study is therefore unlikely to be attributable to an ineffective version of the training task, given that a functionally identical *n*-back training regime produced successful transfer in earlier studies (Jaeggi, et al., 2010). Rather, these findings suggest that by addressing the key reservations highlighted by Sternberg (2008) and Moody (2009) (namely the use of a different fluid intelligence test and an active control group), the apparent effectiveness of the *n*-back training regime on fluid intelligence was reduced/eliminated. Therefore, attention must turn to these experimental manipulations.

## Measuring the Fluid Intelligence Construct

One of the most distinct differences between the present study and that of Jaeggi et al. (2010) was the type of test used to measure the fluid intelligence construct. The present study used the Figure Weights subtest, rather than the geometric matrix reasoning tests (RAPM and BOMAT) used as proxies for fluid intelligence by Jaeggi and her colleagues (2010). As mentioned previously, the Figure Weights subtest has been found to correlate highly with *Gf* (0.78) and is a measure of quantitative and analogical reasoning (Benson, Hulac, & Kranzler, 2010). In contrast to the RAPM and BOMAT, the Figure Weights subtest places less emphasis on visualspatial patterns, and instead requires reasoning with the quantitative information underlying each stimulus (Groth-Marnat, 2009).

Additionally, unlike the RAPM, the stimuli of the Figure Weights subtest are not presented within a 3 x 3 grid/matrix layout, providing further dissimilarity from the visualspatial *n*-back used by Jaeggi et al. (2010). In the opinion of the current researcher, these elements make the Figure Weights less susceptible to training gain related to non-*Gf* processes (such as spatial storage) after *n*-back training. Therefore, considered in isolation, the lack of transfer to the Figure Weights subtest in the current study could be seen as support for one of the arguments proposed by Moody (2009); that the *n*-back may not improve actual fluid reasoning ability, but rather just improve a skill specific to the efficient solving of geometric reasoning problems (practice with storing spatial information).

However, the results of the companion study provide evidence against this direct conclusion. The companion study used the same methodology and participants as the current project, but investigated the effect of the training regime on the RAPM rather than the Figure Weights subtest. As in the present study, whilst both groups improved from pre to post, no significant transfer to fluid intelligence was found in the *n*-back group relative to the active control group; in fact, on the RAPM, the control group improved significantly more than the *n*-back group. Consequently, it would appear that the *n*-back task was unsuccessful in significantly training even an aspect unique to the intricacies of matrix reasoning tests.

These findings illustrate the dangers of extrapolating the results of single tests as representative of true gains in the fluid intelligence construct (Jensen, 1981); as training tasks may develop skills specific to the solving of a single test, rather than improving fluid intelligence (Shipstead et al., 2010). The control tasks used in the present and companion studies were crystallised intelligence tasks (vocabulary and general knowledge) and under CHC theory are unlikely to increase reasoning ability (McGrew, 2009). Therefore, although outside the scope of the present thesis, one explanation for the control group outperforming the *n*-back group on the RAPM (a pattern not seen on the Figure Weights subtest) is that these tasks facilitated performance specific to completing the RAPM. Due to the administration procedure of the RAPM, quickly recognising and selecting the correct answer for previously solved items could be advantageous in allowing an individual more time to solve the remaining items (Hamel & Schmettmann, 2006). It is possible that the nature of the vocabulary task conditioned participants to scan and select the correct answer as quickly as possible, facilitating efficient strategy use on the RAPM.

Nevertheless, returning to the basis for this study, the effectiveness of the *n*-back task; results suggest that the *n*-back task was not successful in generating transfer to either the Figure Weights subtest or a matrix reasoning test previously (and successfully) used by Jaeggi et al. (2010). Therefore, the weight of current evidence indicates that the reason for these divergent findings is likely to lie not with the type of fluid intelligence test used, but with the other major difference between the present study and those conducted by Jaeggi and colleagues; the nature of the control group.

**Use of ‘Active’ rather than ‘No-contact’ Control Group**

Studies by Jaeggi and colleagues (2008; 2010) which found successful transfer to fluid intelligence after *n*-back training, utilised a no-contact control group. This meant that whilst participants in the *n*-back group completed up to 20 days of effortful working memory training, control group participants did no activity. Whilst this controlled for simple retest/practice effects, the use of a no-contact control group left open the possibility that gains in the *n*-back group were not due to the properties of the task, but rather the result of placebo and Hawthorne effects (Morrison & Chein, 2011). In other words, simply by engaging in an effortful task which they believed to be improving their intelligence and by being exposed to the research environment, participants were motivated to achieve superior scores at the post-test (Sternberg, 2008). Equally, as emphasised by Orne (1962), participants in a no-contact control may anticipate the purpose of the study (that they are not expected to improve) and consequently be less motivated to perform well during the post-training test measures. Such factors can substantially increase the difference between experimental and control groups in medical and behavioural research (Orne, 1962). Shipstead et al. (2010) and Sternberg (2008) specifically highlighted these issues as very real concerns for the validity of the results of Jaeggi et al. (2010), and the findings of the current paper suggest these concerns were well founded.

In the present study, placebo and Hawthorne effects were accounted for by having the control group complete vocabulary and general knowledge questions. One model which may provide a useful account of these effects is proposed by Geers, Weiland, Kosbab, Landry and Helfer (2005). In a series of studies, Geers and colleagues showed that placebo effects were related to beliefs/expectations about the outcome of an intervention. The effect of these expectations on performance was related to an individual’s goals, motivation and self regulation. According to this model, when expectations for a training task fulfil goals, self regulation can unconsciously guide thoughts, cognitions, and behaviours to confer with the achievement of this goal (Geers et al., 2005). As such, in the context of the current research, a participant’s goal of increasing their intelligence through cognitive training may have unconsciously influenced their thoughts and behaviours to exert more mental effort during the IQ test (Shipstead et al., 2010).

Given that this study used an identical training regime to that used by Jaeggi et al. (2010), the crucial difference between the study designs appears to be that Jaeggi et al. did not use an active control group. Therefore, the lack of significant transfer to *Gf* in the present study supports the notion that transfer in these previous studies (Jaeggi et al., 2008; 2010) was likely the result of placebo/Hawthorne effects rather than actual increases to fluid reasoning ability (Shipstead et al., 2010).

These conclusions are consistent with the large body of literature indicating that performance on IQ tests is not simply a static number, but rather, a range of motivational factors can significantly influence one's score (Carr & Dweck, 2011). For example, it has been found that an individual's beliefs about whether intelligence is fixed or malleable impacts on academic and intelligence test performance (Cury et al., 2006). Similarly, in a phenomenon known as stereotype threat, researchers have found that an individual's concerns about their performance and how others will view them can significantly degrade intelligence test scores (Carr & Dweck, 2011). Such research demonstrates the significant impact that participants' beliefs and expectations can have on measures of their intellectual functioning (Carr & Dweck, 2011).

**Implications and Consistency with Prior Research**

The purpose of the current study was to explore the effectiveness of the *n*-back task as a method of improving fluid intelligence. As such, the pertinent finding is that, relative to an active control group, the *n*-back task demonstrated no significant transfer to the Figure Weights subtest; which is a valid measure of fluid reasoning ability (Bensen et al., 2010). These findings may have substantial implications for the working memory training literature. A number of researchers (such as Jaeggi et al., 2010) had ascribed gains on fluid intelligence tests after *n*-back training to the improvement of executive functions such as attentional control, updating and interference inhibition (Conway et al., 2011). The results of the present study are inconsistent with this view and suggest that training these mechanisms through *n*-back training may not significantly improve fluid reasoning ability (Colom et al., 2010).

Furthermore, the present result is in keeping with the findings of Chein and Morrison (2010), who trained participants using complex span tasks which also involved the abovementioned executive functions. Following the training, improvements were found to

Stroop task performance (indicating increased cognitive control), but these gains did not generalise to fluid intelligence (as measured by the RAPM). Such results provide additional support for the conclusions of Jensen (1981), that the fluid intelligence construct is largely immune to change.

The lack of an effect on *Gf* in the current study reinforces the concerns of Shipstead et al. (2010) and Sternberg (2008) in suggesting that transfer to *Gf* tests in previous studies was perhaps not the result of actual gains to fluid reasoning ability, but was instead due to differences in motivation associated with completing a training task believed to be improving performance. Consistent with this assessment, the majority of studies documenting successful transfer to *Gf* have used no-contact controls or no control group (Shipstead et al., 2010), and other researchers have failed to find transfer after working memory training was compared to an active control (Colom et al., 2010; Holmes, Gathercole, & Dunning, 2009; Shavelson et al., 2008).

However, Jaeggi et al. (2011), whilst failing to find an overall effect between an *n*-back training and an active control group in terms of transfer to fluid intelligence, reported significant transfer when their sample was divided into those who improved most on the *n*-back (high gain group) and those who improved less (low gain group). The high gain group showed significant transfer to fluid intelligence relative to the active control and low gain groups. This led Jaeggi and colleagues (2011) to conclude that improvement on the *n*-back was related to increased fluid intelligence scores. It is important to note the differences in the present study, which used this same method of splitting the groups into high and low gain. Whilst Jaeggi et al. (2011) used matrix reasoning tests, the current project used the Figure Weights subtest and found no difference in performance between any of the groups. Furthermore, analysis showed that *n*-back gain was not a significant predictor of *Gf* improvement. It would be interesting to reanalyse the data of Jaeggi et al. (2011) and perform median splits on the active control group as well, in order to ascertain whether those high gain active control group participants also showed a significant transfer to fluid intelligence (indicating a motivational component behind the test improvement) (Carr & Dweck, 2011). Moreover, Jaeggi et al. (2008) found that the size of the transfer to *Gf* increased with higher dosages, and they suggested that a training regime which stressed working memory capacity for longer may have produced larger improvements in *Gf* test performance. An equally viable explanation is that the higher dosage training participants were more motivated (Carr & Dweck, 2011; Morrison & Chein, 2011). Hence, taken as a whole, the present data and that of the literature are consistent with a motivational, rather than specific training explanation (Morrison & Chein, 2011).

**Age and transfer to fluid intelligence.** When considering the working memory training studies which have used an active control group, it is also important to note that those which have demonstrated successful transfer to *Gf* involved children or adolescent participants. No published working memory training studies featuring an active control group have demonstrated successful transfer to *Gf* in adults (Morrison & Chein, 2011). This pattern is consistent with research documenting that brain plasticity decreases with age, and suggests that the potential to improve fluid intelligence through working memory training may be greater in children (Garber & Heber, 1982). Therefore, one possibility is that true gains in fluid reasoning ability through working memory training are attainable in child participants, but not in adults. This is in accordance with the hypothesis of early research with low functioning children, which sought to implement interventions while children were young in order to maximise the effectiveness of the programs (Heber & Garber, 1973; Ramey & Haskins, 1981). Future research comparing the effectiveness of working memory training programs across age groups would be useful in establishing whether there is substance behind this pattern, and for which groups (if any) fluid reasoning ability can be increased.

## Limitations

The results of the current study fulfil a crucial role in further examining the beneficial applications of the *n*-back task; however, a number of limitations must be noted.

**Training was unsupervised.** Firstly, due to the number of participants and extended nature of the training period, it was not feasible for participants to be directly supervised by the researcher during the 20 training sessions. As such, it is possible that some participants may have neglected to do the training and fabricated their scores; meaning they received a lesser dosage of the training. Whilst this is a valid concern, it is unlikely to have been an issue in the present study. It is probable that the majority of participants volunteered to take part because they were genuinely interested in the prospect of increasing their intelligence; they were not required to participate (to fulfil university course requirements) or given any monetary incentives. Given this intrinsic motivation, it follows that participants would adhere to the training regime in order to achieve their goal of increased intelligence (Carr & Dweck, 2011). In support of this conclusion, participants in both groups significantly improved their performance on the task across the 20 sessions, and followed a similar pattern of learning on the *n*-back task to participants in the study by Jaeggi et al. (2010).

**Lack of a 'no-contact' control group.** Secondly, the scope of the present project did not allow for the inclusion of both an active control and a no-contact control group; hence, only an active control group was included. Therefore, whilst it is assumed that the present study failed to find significant transfer to *Gf* because placebo effects were controlled for, the precise impact of having control group participants complete a placebo task cannot be determined from the present data. The results of future research would be strengthened by including both active and inactive control groups to account for this concern (Shipstead et al., 2010).

## Future Directions

Whilst it has been stressed that the present results give credence to the view that previous studies may not have demonstrated true gains to fluid reasoning ability; given the range of divergent findings in the literature, further research is required in order to disentangle the effects of working memory training on fluid intelligence (Morrison and Chein, 2011). The findings of the current study highlight that it is of paramount importance for this research to consider motivational influences; a factor which has been largely overlooked in the existing body of research (Shipstead et al., 2010). Only after these influences have been adequately accounted for, can the precise effects of working memory training on fluid intelligence begin to be understood (Carr & Dweck, 2011). As such, future research should include active control groups, and ideally utilise observer ratings of participant motivation or self-report intrinsic motivation questionnaires (in order to control for differences in engagement/effort in the training tasks and transfer measures) (Duckworth et al., 2010).

**Measuring *Gf* using latent variable approach.** Another important step for this research area is to move away from the measurement of *Gf* with individual tests. As emphasised by Jensen (1981) and Sternberg (2008), extrapolating gains on one test as representing gains in fluid intelligence ability is problematic, as performance is affected by elements specific to the intricacies of that test. As mentioned earlier, this is clearly demonstrated by the differential improvement in the RAPM and Figure Weights subtest in the present and companion studies; two tests purported to measure the same construct (Groth-Marnat, 2009). Hence, it is of concern that the current body of literature has almost exclusively used geometric matrix reasoning tests as a sole proxy for *Gf* (Lohman & Lakin, 2011). It will therefore be crucial for future research to measure *Gf* using a battery of reasoning tests and assess transfer to the latent variable (*Gf*) common to all tests measured. Doing so will allow for a clearer understanding of the effects of working memory training on the underlying fluid intelligence construct, rather than performance specific to one type of test (Jensen, 1981). At this point in time, only one study has used this approach, and did so without accounting for motivational factors on performance (Schmiedek, Lovden, & Lindenberger, 2010). The Concept Formation and Analysis Synthesis tests (from the Woodcock-Johnson-III Tests of Cognitive Abilities) are examples of other *Gf* measures which may prove useful additions to a battery of fluid reasoning tasks in future studies (Woodcock, McGrew, & Mather, 2001).

**Practical benefits of training.** In order for working memory training to have merit, it will also be necessary to show that any gains on *Gf* tests extend to practical applications of fluid intelligence, such as academic and professional achievement (Sternberg, 2008). Although improvement of scores on IQ tests is undoubtedly noteworthy, it is of little practical significance unless gains also extend to real world applications of fluid intelligence (Jensen, 1981). Answers to such questions are outside the scope of the current project, but are necessary for the usefulness of working memory training to be fully understood (Sternberg, 2008).

## Conclusion

Whether or not fluid intelligence can be improved, and to what degree, is a highly controversial subject (Jensen, 1981; Sternberg, 2008). Recent research has generated optimism that working memory training using the *n*-back task may result in improvement in fluid intelligence (Jaeggi et al., 2008). However, the results of the present study shed doubt on such claims and lend support to the concerns of Shipstead et al. (2010), Sternberg (2008) and Morrison and Chein (2011). Their proposal, that increases to fluid intelligence in previous studies were possibly the result of motivational/expectancy effects rather than the properties of the training task (*n*-back), is consistent with the lack of result reported in this study. It is therefore important to question the validity of previous studies which did not include an active control group to account for motivational effects on *Gf* test performance (Shipstead et al., 2010). Further research is required in order to more clearly understand whether it is possible to improve fluid intelligence through such training (Morrison & Chein, 2011), but the present study has served to highlight the fact that motivational influences on performance must be accounted for if the nature of this link is to be properly understood (Carr & Dweck, 2011). By doing so, this study makes a valuable contribution to the current body of literature striving to determine and understand the effect of working memory training on fluid intelligence.

References

Allen, P., & Bennett, K. (2008). *SPSS for the health and behavioural sciences*. Australia: Thomson.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences, 4*, 417-423. doi:10.1016/S1364-6613(00)01538-2

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews, 4,* 829-839. doi:10.1038/nm1201

Baltes, P. B., Staudinger, U. M., & Lindenberger, U. (1999). Lifespan psychology: Theory and application to intellectual functioning. *Annual Review of Psychology, 50*, 471-507. doi: 10.1146/annurev.psych.50.1.471

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*, 612-637. doi: 10.1037/0033-2909.128.4.612

Benson, N., Hulac, D. M., & Kranzler, J. H. (2010). Independent examination of the Wechsler Adult Intelligence Scale – Fourth Edition (WAIS-IV): What does the WAIS-IV measure? *Psychological Assessment*, *22*, 121-130. doi: 10.1037/a0017767

Bors, D. A. & Vigneau, F. (2003). The effect of practice on Raven's advanced progressive matrices. *Learning and Individual Differences, 13,* 291-312. doi:10.1016/S1041-6080(03)00015-3

Brainworkshop. (n. d.). A dual n-back game. Retrieved from http://brainworkshop.sourceforge.net/download.html

Bunting, M. (2006). Proactive interference and item similarity in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 32*, 183-196. doi: 10.1037/0278-7393.32.2.183

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What one intelligence test measures: a theoretical account of the processing in the Raven progressive matrices test. *Psychological review, 97,* 404-431. doi: 10.1037/0033-295X.97.3.404

Carr, P. B., & Dweck, C. S. (2011). Intelligence and motivation . In R. J. Sternberg, & S. B. Kaufman (Eds.), *The Cambridge Handbook of Intelligence* (pp. 748-770). USA: Cambridge University Press.

Chein, J. M., & Morrison, A. B. (2010). Expanding the mind's workspace: Training and transfer effects with a complex working memory span task. *Psychonomic Bulletin and Review, 17,* 193-199. doi: 10.3758/PBR.17.2.193

Colom, R., Quiroga, M. A., Shih, P. C., Martinez, K., Burgaleta, M., Martinez-Molina, A., Romain, F. J., Requena, L., & Ramirez, I. (2010). Improvement in working memory is not related to increased intelligence scores. *Intelligence, 38,* 497-505. doi:10.1016/j.intell.2010.06.008

Conway, A. R., & Getz, S. (2010). Cognitive ability: Does working memory training enhance intelligence? *Current Biology, 20,* 362-364. doi: 10.1016/j.cub.2010.03.001

Conway, A. R., Getz, S., Macnamara, B., & Engel de Abreu, P. (2011). Working memory and intelligence. In R. J. Sternberg, & S. B. Kaufman (Eds.), *The Cambridge Handbook of Intelligence* (pp. 394-419). USA: Cambridge University Press.

Conway, A. R., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Science, 17,* 547-552. doi:10.1016/j.tics.2003.10.005

Conway, A. R., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin and Review, 12,* 769-786. doi: 10.3758/BF03196772

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information processing system. *Psychological*

*Bulletin,* 104, 163-191. doi: 10.1037/0033-2909.104.2.163

Cury, F., Elliot, A. J., Da Fonseca, D., & Moller, A. C. (2006). The social-cognitive model of achievement motivation and the 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology, 90,* 666-679. doi:10.1016/j.jesp.2007.07.003

Dahlin, E., Neely, A.S., Larsson, A., Backman, L., & Nyberg, L. (2008) Transfer of learning after updating training mediated by the stratum. *Science, 13*, 1510-1512. doi: 10.1126/science.1155466

Deary, I. J., Strand, S. S., Smith, P., & Fernandes, C. (2007). Intelligence and educational achievement. *Intelligence, 35,* 13-21. doi:10.1016/j.intell.2006.02.001

Dempster, F. N., & Corkill, A. J. (1999). Interference and inhibition in cognition and behaviour: Unifying themes for educational psychology. *Educational Psychology Review, 11,* 1-88. doi: 10.1023/A:1021992632168

Duckworth, A. L., Quinn, P. D., Lynam, D. R., Loeber, R., & Stouthamer-Loeber (2010). Role of test motivation in intelligence testing. *Proceedings of the National Academy of Sciences, 108,* 7716-7720. doi: 10.1073/pnas.1018601108

East of the Web. (2003). Definetime [Adobe Flash Player game]. Retrieved from http://www.eastoftheweb.com/games/DefineTime1.html

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short term memory and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General, 128,* 309-311. doi: 10.1037/0096-3445.128.3.309

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., and Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science, 17*, 172-179. doi: 10.1111/j.1467-9280.2006.01681.x

Garber, H., & Heber, R. (1982). Modification of predicted cognitive development in high-risk children through early intervention. In D. K., Detterman & R. J., Sternberg, R. J (Eds.), *How and how much can intelligence be increased* (pp. 121-140). USA: ABLEX Publishing Corporation.

Geers, A. L., Weiland, P. E., Kosbab, K., Landry, S. J., & Helfer, S. G. (2005). Goal activation, expectations, and the placebo effect. *Journal of Personality and Social Psychology, 89*, 143-159. doi: 0.1037/0022-3514.89.2.143

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience, 6,* 316-322. doi: 10.1038/nn1014

Groth-Marnat, G. (2009). Handbook of psychological assessment (5th ed.). Wiley: New Jersey: USA.

Halford, G. S., Cowan, N., & Andrews, G. (2007). Separating cognitive capacity from knowledge: A new hypothesis. *Trends in Cognitive Sciences, 11*, 236-242. doi: 10.1016/j.tics.2007.04.001

Hamel, R., & Schmettmann, V. D. (2006). The 20 minute version as a predictor of the Raven advanced progressive matrices test. *Educational and Psychological Measurement, 66*, 1039-1046. doi: 10.1177/0013164406288169

Heber, G., & Garber, R. (1973). The Milwaukee project: Early intervention as a technique to prevent mental retardation. Retrieved from http://eric.ed.gov/PDFS/ED080162.pdf

Holmes, J., Gathercole, S. E., Place, M., Dunning, D. L., Hilton, K. A., & Elliot, J. G. (2009). Working memory deficits can be overcome: Impacts of training and medication on working memory in children with ADHD. *Applied Cognitive Psychology, 24,* 827-836. doi: 10.1002/acp.1589

Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science, 12,* 9-15. doi: 10.1111/j.1467-7687.2009.00848.x

Jaeggi, S. M., Buschkuel, M., Jonides, J. & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Procedures of the National Academy of Sciences, 105*, 6829-6833. doi: 10.1073/pnas.0801268105

Jaeggi, S. M., Buschkuehl, M., Perrig, W. J., Meier, B. (2009). The concurrent validity of the *n*-back task as a working memory measure. *Memory, 18,* 394-412. doi: 10.1080/09658211003702171

Jaeggi, S. M., Studer-Luethi, B., Buschkuehl, M., Yi-Fen Su, Jonides, J., & Perrig, W. J. (2010). The relationship between n-back performance and matrix reasoning – Implications for training and transfer. *Intelligence, 38*, 625-635. doi: 10.1016/j.intell.2010.09.001

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences,108,* 10081-10086. doi: 10.1073/pnas.1103228108

Jensen, A. R. (1981). Raising the IQ: The Ramey and Haskins study. *Intelligence, 5,* 29-40. doi:10.1016/0160-2896(81)90015-5

Jonides, J. (2004). How does practice make perfect? *Nature Neuroscience*, *139*, 10-11. doi: 10.1016/j.neuroscience.2005.07.002

Kane, M. J., Hambrick, D. Z., & Conway, A. R. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin, 131*, 66-71. doi: 10.1037/0033-2909.131.1.66

Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent variable approach to verbal and visual spatial memory span and reasoning. *Journal of Experimental Psychology: General, 133*, 189-217. doi: 10.1037/0096-3445.133.2.189

Klingberg, T., Forssberg, H., & Westerberg, H. (2002). Training of working memory in children with ADHD. *Journal of Clinical and Experimental Neuropsychology, 24,* 781-791. doi: 10.1076/jcen.24.6.781.8395

Klingberg, T., Fernell, E., Olesen, P., Johnson, M., Gustafsson, P., Dahlstrom, K., Gillberg, C. G., Forssberg, H., & Westerberg, H. (2005). Computerized training of working memory in children with ADHD – A randomized, controlled, trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 44,* 177-186. doi:10.1097/00004583-200502000-00010

Klingberg, T. (2010). Training and plasticity of working memory. *Trends in Cognitive Sciences, 14,* 317-324. doi:10.1016/j.tics.2010.05.002

Kramer, A. F., Bherer, L., Colcombe, S. J., Dong, W., & Greenough, W. T. (2004). Environmental influences on cognitive and brain plasticity during aging. *The Journals of Gerontology, 59,* 940-957. doi: 10.1093/gerona/59.9.M940

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working memory capacity?! *Intelligence, 14,* 389-433. doi:10.1016/S0160-2896(05)80012-1

Lohman, D. F., & Lakin, J. M. (2011). Intelligence and reasoning. In R. J. Sternberg, & S. B. Kaufman (Eds.), *The Cambridge Handbook of Intelligence* (pp. 419-442). USA: Cambridge University Press.

McCarney, R., Warner, J., Iliffe, S., van Haselen, R., Griffin, M., & Fisher., P. (2007). The Hawthorne effect: a randomized, controlled trial. *BCM Medical Research Methodology, 30,* 1-8. doi: 10.1186/1471-2288-7-30

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence, 37,* 1-10. doi: 10.1073/pnas.0803396105

Moody, D. E. (2009). Can intelligence be increased by training on a task of working memory? *Intelligence, 37,* 327-328. doi: 10.1016/j.intell.2009.04.005

Morrison, A. B., & Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychological Bulletin Review, 18,* 46-60. doi: 10.3758/s13423-010-0034-0

Mueller, C. M., & Dweck, C. S. (1998). Praise for intelligence can undermine children's motivation and performance. *Journal of Personality and Social Psychology, 75*, 33-52. doi: 10.1037//0022-3514.75.1.33

Oberauer, K. (2005). Binding and inhibition in working memory: Individual and age differences in short-term recognition. *Journal of Experimental Psychology: General, 134,* 368-387. doi: 10.1037/0096-3445.134.3.368

Olesen, P. J., Westerberg, H., & Klingberg, T. (2004). Increased prefrontal and parietal activity after training of working memory. *Nature Neuroscience*, *7,* 75-79. doi:10.1038/nn1165

Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 17, 776-783. doi: 10.1037/h0043424

Perrig, W. J., Hollenstein, M., & Oelhafen, S. (2009). Can we improve fluid intelligence with training on working memory in persons with intellectual disabilities? *Journal of Cognitive Education and Psychology, 8,* 148- 164. doi: 10.1891/1945-8959.8.2.148

Raven, J. C., Court, J. H., & Raven, J. (1977). *Standard progressive matrices*. London: H. K. Lewis.

Real Player Games Directory. (n.d.). Who wants to be a millionaire [Adobe Flash Player game]. Retrieved from http://www.box10.com/who-wants-to-be-a-millionaire.html

Salthouse, T. A., Atkinson, T. M., & Berish, D. E. (2003). Executive functioning as a potential mediator of age-related cognitive decline in normal adults. *Journal of Experimental Psychology: General, 132,* 566-594. doi: 10.1037/0096-3445.132.4.566

Schmiedek, F., Lovden, M., & Lindenberger, U. (2010). Hundred days of cognitive training enhance broad cognitive abilities in adulthood: Findings from the COGITO study. *Frontiers in Aging Neuroscience, 2*, 1-10. doi: 10.3389/fnagi.2010.00027

Shavelson, R. J.,Yuan, K., & Alonzo, A. (2008). On the impact of computer training on working memory and fluid intelligence. In D. C. Berliner & H. Kuermintz (Eds.), *Fostering change in institutions, environments, and people: A festschrift in honor of Gavriel Salomon* (pp. 35-48). New York: Routeledge.

Shelton, J. T., Metzger, L. R., & Elliot, E. M. (2007). A group administered lag task as a measure of working memory. *Behavior Research Methods, 39,* 482-493. doi: 10.3758/BF03193017

Shipstead, Z., Redick, T. S., & Engle, R. W. (2010). Does working memory training generalize? *Psychologica Belgica, 50*, 245-276. Retrieved from http://psychology.gatech.edu/

Sternberg (2008). Increasing fluid intelligence is possible after all. *Procedures of the National Academy of Sciences, 105*, 6791-6792. doi: 10.1073/pnas.0803396105

Takeuchi, H., Sekiguchi,A., Taki, Y., Yokoyama, S., Yomogida, Y., Komuro, N., Yamanouchi, T., Suzuki, S., & Kawashiwa, R. (2010). Training of working memory impacts structural connectivity. *The Journal of Neuroscience, 30,* 3297-3303. doi: 10.1523/JNEUROSCI.4611-09.2010

Wechsler, D. (2008). Wechsler Adult Intelligence Scale- 4$^{th}$ ed. San Antonio, Tx: NCS Pearson, Inc.

Woodcock, R.W., McGrew, K.S., & Mather, N. (2001). *Woodcock-Johnson III tests of cognitive abilities*. Itasca, IL: Riverside Publishing.

Appendix A

# Information Letter to Participants

*Improving Intelligence through Cognitive Training*

Thank you for your interest in our project. We are postgraduate students completing Honours in Psychology at Edith Cowan University. You are invited to take part the research project that we are conducting as part of the requirements for our degree. The research project has ethics approval from the Faculty of Computing, Health and Science Ethics Sub-Committee.

This project aims to investigate whether intelligence can be improved by training in a specific type of cognitive task. It had previously been thought that intelligence was relatively fixed, however, some recent studies have indicated that certain types of training can improve this ability.

**If you choose to take part in the study you would be required to:**

**1)** Do two tests of intelligence at a location of your choosing (takes just over one hour). This can occur at the time/date most convenient for you.

**2)** Train using the cognitive activity. The activity is done on the computer and can be done at any location (e.g.: home). The training schedule requires you to do the activity for approximately 20 minutes a day, for 20 days after you have completed the intelligence test.

**3)** Re-do the tests of intelligence (takes just over 1 hour). Ideally this would occur no more than 3 days after the cognitive task training is completed.

(Note: The days and time at which you complete the training activity are up to you. However it is important for learning to have some form of routine, and the 20 days does need to be completed within approximately 1 month. The recommended training schedule involves 5 days of training followed by a 2 day break; for example, complete a training session each weeknight and have a break from training on the weekend).

**Benefits to participants:** Benefits gained from participating in the study include possibly improving your intelligence during the training period, and gaining access to a program that you can use to practise these skills in the future. You would also gain experience in answering the type of questions used in intelligence tests, which may be an advantage to you if you are required to do a full IQ /aptitude test in the future (e.g.: as part of a job application).

**Potential risks/discomfort to participants:** There are no foreseeable risks to participants in this study, other than the inconvenience of having to commit to the training and testing time.

**Confidentiality/use of data from study:** All data gathered during the course of the study will be kept confidential, and will not be discussed with anyone outside the research team. The results may be published for scientific purposes, but will not include your name or any other personal information. Should you choose to participate, once your scores for the fluid intelligence tests have been collected, your name will be removed from the answer sheets and replaced with a code (e.g.: name ‘John S’, changed to ‘Participant A1’). As such, if you choose to participate in this study, you will not be able to be personally identified with the data.

**Choice to participate in the study:** You are under no obligation to participate in this study. I understand that taking part in this research requires a significant time commitment, so please be aware that no consequences, punishment or loss of benefits will occur as a result of a decision to not participate.

If you do choose to take part in the study, you are free to withdraw at any time without explanation. If you elect to discontinue your participation at any time during the study, please contact David Preece, Vaughan Palmer or Dr Ken Robinson as soon as possible.

**Contacts:** If you would like to take part in the project please contact David or Vaughan through email or phone (details below).

Please note that a summary or the full thesis can be provided on request if you would like to know the results of the study.

Thank you for your time,
yours sincerely,

David Preece Vaughan Palmer

Listed overleaf are the contact details if you have any questions about the research project or require further information.

| Contact Details | |
|---|---|
| Investigator: | David Preece |
| email | |
| Phone | |
| | |
| Investigator | Vaughan Palmer |
| email | |
| Phone | |
| | |
| Supervisor | Dr Ken Robinson |
| email | |
| Phone | |
| School | Edith Cowan University (Joondalup) |
| Faculty | Computing, Health and Science |

Should you have any concerns regarding this research project and you would like to contact an independent research ethics officer, you may contact:

| Contact Details | |
|---|---|
| Research Ethics officer | Kim Gifkins |
| email | |
| Phone | |

Appendix B

# Participant Demographic Information Sheet

Name: ______________________________

Age (years): ______________

Gender: ________________

Contact number: ____________________________

Email address: ______________________________________________

Study number: _______________________ (researcher to provide)

**What is your highest qualification?**

- ❑ < Year 12
- ❑ Year 12
- ❑ University degree
- ❑ Other qualification after year 12

**Are you currently studying at:**

- ❑ University
- ❑ Another Institution (TAFE, other colleges)

# Participant Informed Consent Form

**Project Title:** *The Effect of Cognitive Training on Fluid Intelligence*

| Contact Details | |
|---|---|
| Student in charge of project | David Preece |
| email | |
| Phone | |
| Supervisor | Dr Ken Robinson |
| email | |
| Phone | |
| School | Edith Cowan University (Joondalup) |
| Faculty | School of Psychology and Social Science |

**Consent:**
If you have read the details in the information letter and are interested in participating in the study, please indicate this by signing the statement below:

*I have been provided with the project information letter and understand the purpose of this study, what I would be required to do as a participant, and how the resulting data will be used. Given this information, I am volunteering to be a participant in 'The Effect of Cognitive Training on Fluid Intelligence' study.*

**Participant's Name:** ________________________________

**Participant's Signature:** ____________________________ **Date:** ___________________

**Researcher's Name:** ______________________ **Researcher's Signature:** __________________

**Date:** ___________________

Appendix C

# Instructions for the *n*-back task

The n-back task is done on the computer and will be given to you on a CD or emailed to you by the researcher. Alternatively, you can download the program from the link on this web address: http://brainworkshop.sourceforge.net/download.html

The program is required to be installed on your desktop, it is fairly small and only requires a few minutes to download/install.

Once the program has been installed, double click on the desktop icon to run it. This will open up the 'main menu'. Before starting press the 'c' key on your keyboard to go into the options menu. Scroll down to the 'use audio' option and press the 'space bar' key to change this option to 'no'. Doing this means you only have to keep track of a visual stimulus, rather than having to keep track of BOTH a visual and audio stimulus. Press 'enter' to apply these changes to the options.

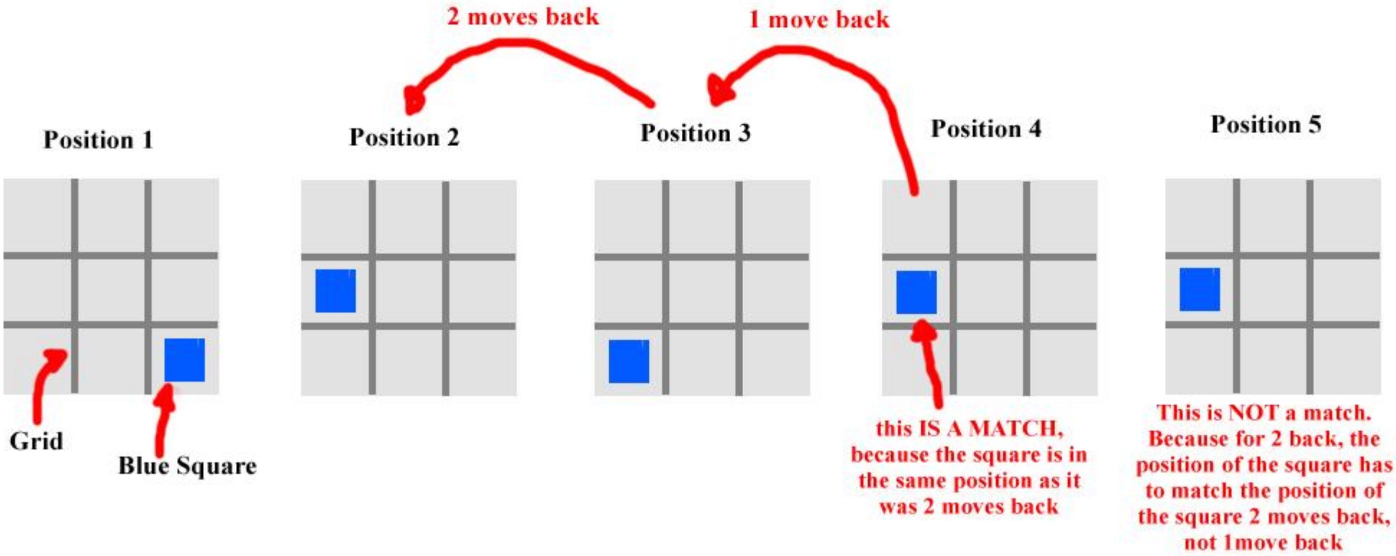
You should now be back on the main menu, press the 'space bar' key to start the activity when you are ready. This should bring you to the 'game' screen with a 3 by 3 grid in the middle of the screen, press the 'space bar' key when you are ready to start the activity. On the top right hand side of the screen, the program displays how many 'sessions' you have completed, each session is made up of around 21-29 movements of the square. Do the activity until you have completed 15 sessions for that day. This should take around 17-20 minutes to complete. In the top right hand of the 'game screen' is the phrase 'PoNB average' and a number (eg: 0.00), this refers to the overall average level of *n* you were operating at during the 15 sessions. As you do the activity, the n-back level you were at for each of the 15 sessions will be recorded on the right hand side of the screen (an example of one of these messages would be: #12 Po4B 60%. This means that on session 12 you were on n-back level 4 and got 60% right). After you have finished the 15 sessions for the day, please note down the PoNB average number, and the highest n-back level you reached that day and give this information to the researcher.

In terms of doing the activity, press the 'a' key to indicate a match. When you need to indicate a match and how to do the task will be explained below.

## How to do the *n*-back task:

The n-back task is a working memory activity where you have to keep track of the position of a blue square. This blue square moves around positions on a grid. The aim of the task is to press a button to indicate a match when the stimulus is in the same position on the grid as it was *n* times earlier. What the *n* in *n*-back means will be explained in the following paragraphs.

For example; if you are playing *2*-back, you are required to indicate a match when the blue square is in the same position on the grid as it was *2* moves ago. When the blue square is not in the same position as it was 2 moves earlier, you simply don't press any button. (See picture bellow for a diagram explanation of *2*-back).




The level of *n* increases or decreases depending on your performance on the task. What this means is that if you are doing *2*-back (have to remember the position of the square 2 times back) and perform well, the program will change the value of *n* up to 3-back. For 3-back, the task is harder because you now have to remember where the square was 3 moves ago, rather than 2. Conversely, if you don't perform well on the task, the program may decrease the level of *n* to make it slightly easier. In this way the task remains challenging, but not too difficult.

The default starting level is *2*-back. Do not be discouraged if you have trouble with the task at first, it is a difficult working memory activity. The easiest level is *1*-back where you have to indicate a match when the square is in the same position as it was 1 move earlier. Levels of *n* progress upwards in increments of 1. There is no limit to what level of *n* you can reach. However, regardless of which level you reached in the previous days training, each daily training session should start at 2-back.

For example, in 6-back, you are required to indicate a match when the position of the square is the same as it was 6 moves ago. If you indicate a match when the square is in the same position as it was 5 moves ago this is incorrect, and if you indicate a match when it is in the same position as it was 7 moves ago this is also incorrect. It has to be in the same position as it was exactly 6 moves ago to be a correct match.

As you can see, keeping track of exactly where the square was, and how many moves ago it was there can get difficult on higher levels of *n* as you have to remember further and further back in the sequence.

# Questions (see the next page for information on the main key commands you will need to know)

If you have any questions regarding this task please don't hesitate to contact me. I know from experience that wrapping your head around exactly what the *n*-back task is and what it requires you to do, can be confusing.

,Dave

**Email:**
**Phone:**

# Relevant Commands / Controls

On the ‘game screen’ (the screen with the 3 by 3 grid in the middle)

**Ctrl + C = clear the n-back session history**. Sometimes the program will store the n-back levels reached from the previous day’s session. So that these numbers don’t effect the average n-back number for the current days sessions, simply hold Ctrl and press the ‘c’ key to clear the history from the right hand side of the ‘game screen’. You will know this has been successful from the ‘PoNB average’ displayed up the top right changing to read as ‘PoNB average: 0.00’.

**Press ‘m’ key = Manual mode**. For the purposes of this study, each daily session should start at ‘2-back’ and progress up from there. Sometimes the program will save the n-back level you reached at the end of the previous days training and try to use this as the starting level for the next day’s training. **To make sure that you start each days sessions at 2-back press the ‘m’ key when on the game screen.** This will take you into ‘manual mode’ (note that the title up the top of the screen changes from ‘Position n-back’ to ‘Manual mode: Position n-back’). From here **press the F1 and F2 buttons to manually decrease (F1) /increase (F2) the level of n that you are at until the title at top says 2-back**. Be careful not to adjust any of the other settings as this could affect the results of the study. **Once you have reset the level of n back to 2-back to start the days training press ‘m’ again**. **This returns you out of manual mode, back into standard mode (note that the title up the top changes back to ‘Position n-Back’. It is important to do this as you have to be in standard mode for the level of n to increase/decrease automatically during the 15 sessions based on your performance, which is required for the training to be effective.**

# Creating a User Profile:

On the main menu there is an option to create a user profile. Don’t do this, just use the default profile. Because the program is set up with custom options, the options only take effect when the default profile is being used.

If different people are doing the training on the same computer, when the new person is about to start their sessions, press ‘Ctrl + c’ on the game screen if another person’s scores are still saved on the program. This will clear their session scores and stop their scores from contributing to the new person’s scores.

Appendix D

# Instructions for Vocabulary/Knowledge task

These training tasks are designed to increase an aspect of intelligence.
The two intelligence activities are available online through websites and require Adobe Flash Player to play. You will need to rotate through these activities (do one once then switch to the other, then switch back) for a period of 20 minutes.

## Vocabulary activity

This activity called 'Define Time' requires you to select the correct definition for a word. Bonus points are awarded depending on how fast/accurate you are. At the end of the time limit you will be given your total score for the round. You will need to note down your highest total score achieved for each daily session. On the main menu, you can compare your scores to those of others who have tried the task. This task can be accessed through the website:
**http://www.eastoftheweb.com/games/DefineTime1.html**

## General knowledge activity

This activity is presented in a 'Who Wants to Be a Millionaire' style. At the end of the daily session, you will be required to note down the highest 'money level' you reached. This task can be accessed through the website:
**http://www.box10.com/who-wants-to-be-a-millionaire.html**

If you have any questions, concerns or problems regarding these tasks or what you are required to do, please don't hesitate to contact David Preece or Vaughan Palmer

**Name:** David Preece
**Email:**
**Phone:**

**Name:** Vaughan Palmer
**Email:**
**Phone:**

# Appendix E

| Training Session Number | Average N-Back level (PoNB Average, found in top right corner of program screen) | Highest n-back level reached | Date |
|---|---|---|---|
| Session 1 | | | |
| Session 2 | | | |
| Session 3 | | | |
| Session 4 | | | |
| Session 5 | | | |
| Session 6 | | | |
| Session 7 | | | |
| Session 8 | | | |
| Session 9 | | | |
| Session 10 | | | |
| Session 11 | | | |
| Session 12 | | | |
| Session 13 | | | |
| Session 14 | | | |
| Session 15 | | | |
| Session 16 | | | |
| Session 17 | | | |
| Session 18 | | | |
| Session 19 | | | |
| Session 20 | | | |

Name: ___________________________

Please email your scores after each session to either David or Vaughan so we are able to see how participants are going.

Appendix F

| Training Session Number | Highest Define Time score | Highest money level reached in Millionaire quiz | Date |
|---|---|---|---|
| Session 1 | | | |
| Session 2 | | | |
| Session 3 | | | |
| Session 4 | | | |
| Session 5 | | | |
| Session 6 | | | |
| Session 7 | | | |
| Session 8 | | | |
| Session 9 | | | |
| Session 10 | | | |
| Session 11 | | | |
| Session 12 | | | |
| Session 13 | | | |
| Session 14 | | | |
| Session 15 | | | |
| Session 16 | | | |
| Session 17 | | | |
| Session 18 | | | |
| Session 19 | | | |
| Session 20 | | | |

Name: ___________________________

Please email your scores after each session to either David or Vaughan so we are able to see how participants are going.