Usability and the Effects of Interruption in C-TOC:

Self-Administered Cognitive Testing on a Computer

by

Matthew Michael Brehmer

B.Cmp.H, SSP Cognitive Science, Queen's University, 2009

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Master of Science

in

THE FACULTY OF GRADUATE STUDIES
(Computer Science, Sub–Specialisation in Human–Computer Interaction)

The University Of British Columbia
(Vancouver)

August 2011

Abstract

Cognitive Testing on a Computer (C-TOC) is a self-administered web-based computerised cognitive assessment battery. C-TOC's intended scenario of use involves an older adult, who has presented a concern regarding his or her cognitive health, completing the test independently at home, as directed by their family physician or a specialty clinic.

This thesis presents the results of two studies aimed to address the viability of older adults completing the C-TOC test battery in a home setting, first to identify usability issues, and second to understand the effects of interruptions on C-TOC performance.

In *Study 1*, an initial standard evaluation of C-TOC's usability was conducted with representative users and a cross-cultural advisory panel of health professionals. Based on our own observations of participants' interactions with C-TOC, together with subjective reporting measures (interviews, questionnaires, & focus group discussion), several User Interface (UI) design issues were identified. Given these issues, this thesis presents a list of recommendations for improving C-TOC's usability in subsequent versions.

The bulk of the novel contributions presented in this thesis arise from *Study 2*. In this study, we report the findings of a laboratory experiment to investigate the effects of increasingly demanding interrupting tasks on older adults' C-TOC testing performance. Related work has reported interruptions having a range of inhibitory and facilitatory effects on primary task performance. Cognitive ageing literature has suggested that increased interruption workload demand should have greater detrimental effects on older adults' performance, when compared to younger adults.

With 36 participants from 3 age groups (19–54, 55–69, & 70+), we found divergent effects of increased interruption demand on two primary tasks. Results suggest that older and younger adults experience interruptions differently, that increased interruption demand can incur a task resumption cost. However, at no age is test performance, in terms of accuracy, compromised by demanding interruptions. This finding is reassuring with respect to the success of C-TOC, and is promising for other applications used by older adults.

It is our hope that what was learned from both studies will contribute to the development of a usable and valid cognitive assessment test.

Preface

The studies described in this thesis were conducted with the approval of the UBC Clinical Research Ethics Board (CREB): certificate number H09-02293.

The C-TOC.v1 prototype (Appendix A) used in the studies described in Chapter 3 was developed by Claudia Jacova and Hyunsoo (Steve) Lee, collaborating with Ging-Yuek (Robin) Hsiung and Joanna McGrenere.

The usability observation + interview sessions in *Study 1* (Chapter 3) was conducted with Claudia Jacova. My contribution was the identification of issues related to usability, as well as a list of recommendations for addressing these issues, whereas her contribution was the identification of issues related to C-TOC's content and purpose. She also generated Figure 3.1.

The cross-cultural advisory panel session (Chapter 3) was organised by the Douglas College Centre for Health & Community Partnerships (CHCP). While the scope of the session, the questionnaires, and the focus group discussion covered a range of issues, I report predominantly on usability-related issues.

Parts of this thesis appear in a conference paper manuscript[1] where I was the lead author. Joanna McGrenere, Claudia Jacova, & Charlotte Tang provided supervisory assistance.

[1] M. Brehmer, J. McGrenere, C. Tang, & C. Jacova. Effects of Interruptions on Older Adults' Computerised Cognitive Testing Performance. Manuscript in preparation.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AD Alzheimer's Disease

ANOVA Analysis of Variance, a set of statistical techniques to identify sources of variability between groups

ART Aligned Rank Transform, a non-parametric statistical technique that can accommodate repeated measures designs and examine interaction effects (see Wobbrock et al. [59])

C-TOC Cognitive Testing on a Computer, a self-administered assessment test for age-related cognitive impairment (see Jacova et al. [33])

CIND Cognitive Impairment–Not Dementia

COI Cost of Interruption, see Section 2.2.1

FTD Fronto-Temporal Dementia

HCI Human-Computer Interaction

ICT Information & Communication Technology

IM Instant Messaging

LBD Lewy-Body Dementia

MCI Mild Cognitive Impairment

MMSE Mini-Mental State Examination, a clinical screening test for cognitive impairment (see Folstein et al. [20])

MOCA Montreal Cognitive Assessment, a clinical screening test for cognitive impairment; a score of 26 or higher (out of 30) is considered normal (see Nasreddine et al. [40] and Appendix C.3.1)

NAART North American Adult Reading Test, a quick to administer test measuring Verbal intelligence (see Uttl [53] and Appendix C.3.2)

NASA-TLX NASA Task Load Index, an instrument for gauging the subjective mental workload experienced by a human in performing a task (see Hart and Staveland [23] and Appendix C.7)

NCI Not Cognitively Impaired

NPT Neuropsychological Testing

UI User Interface

WM Working Memory

Acknowledgments

I thank my supervisors Joanna McGrenere & Claudia Jacova, as well as Charlotte Tang, for their guidance.

Research assistant Carmen Li conducted a significant portion of the experimental sessions.

Hyunsoo (Steve) Lee developed the original C-TOC prototype, under the direction of Claudia Jacova. C-TOC was developed as a collaboration with Joanna McGrenere and researchers at the UBC-CARD, including Ging-Yuek (Robin) Hsiung, Lynn Beattie, & Howard Feldman. Research coordinators William Wang & Sarah Le Huray provided logistical assistance during both studies. I also thank Benita Mudge, Penny Slack, & the staff at the UBC Alzheimer's Clinic.

The Douglas College CHCP coordinated the cross-cultural advisory panel.

Susan Mele provided logistical assistance at the Kerrisdale Community Seniors' Centre.

LUNCH, MUX, IDRG, and the HCI grad forum members provided support: Karyn Moffatt (statistical tips); Shathel Haddad, Jessica Dawson, Juliette Link, & Mona Haraty (CHI draft readers); Rock Leung, Amirhossein Mehrabian, Gökhan Himmetoğlu, & Jessica Dawson (pilot subjects).

I thank my second reader, Peter Graf.

Finally, I thank my family & Anamaria Crişan for their encouragement and support.

Dedication

My opa, Manfred Brehmer: over 80 years old, a builder, a tinkerer, a musician, a stroke survivor, and a very recent adopter of the personal computer. This work is for him.

Chapter 1

Introduction

The motivation for the research presented in this thesis stems from an initiative to develop Cognitive Testing on a Computer (C-TOC). C-TOC is a self-administered web-based computerised cognitive assessment that individuals will be able to take independently in the comfort of their home [33].

With ongoing advances in modern medicine in developed countries, people are living longer, resulting in an ageing population. This is associated with an increase of older (55+) individuals experiencing cognitive decline and presenting concerns regarding cognitive health [32]. Many will have no impairment, and are merely worried. Some will be experiencing mild cognitive impairment that can be attributed to normal age-related changes (i.e., not dementia). Others may be experiencing a pathological cognitive impairment along the dementia spectrum.

The screening for pathological cognitive decline including Alzheimer's Disease (AD) and related dementias is currently conducted using paper-based tests including the Mini-Mental State Examination (MMSE) [20] and the Montreal Cognitive Assessment (MOCA) [40]. They are administered by health care professionals or trained staff in clinical settings during a visit. There is currently no opportunity to identify potential impairments before a visit. At the UBC Division of Neurology Centre for Alzheimer's & Related Dementias (UBC-CARD) clinic, wait times for in-depth assessment and consultation regarding cognitive concerns ranges between 6 and 24 months [32].

Thus, innovation in cognitive testing is an urgent yet unmet need because of the growing demand for diagnostic services.

The intent of C-TOC is to aid in directing further diagnostic services to those individuals that exhibit pathological decline so they can be diagnosed and managed as promptly as possible. We do not expect C-TOC to be as accurate as exhaustive neuropsychological assessments, but that it should be sufficiently accurate for the purposes of triaging patients.

C-TOC is a novel hybrid testing tool, in part based on non-computerised tests currently administered in clinical settings, and in part consisting of new test paradigms that probe productive and generative skills [32]. In a single 30-minute session, it aims to assess several cognitive faculties, including memory, language, and spatial reasoning, designed to achieve high sensitivity to mild levels of cognitive impairment.

We expect that some test-takers will have no cognitive decline, while others be may be experiencing Mild Cognitive Impairment (MCI) or Cognitive Impairment–Not Dementia (CIND). It is not intended to be used by individuals with moderate to severe levels of cognitive impairment.

1.1 The Viability of C-TOC

The idea of a self-administered, web-based cognitive screening test that older adults might access from home usually elicits the following questions:

1. *Does* C-TOC *produce valid results?* Will C-TOC accurately assess the same cognitive processes assessed by currently administered clinical pencil + paper tests and Neuropsychological Testing (NPT)?

2. *Is* C-TOC *usable?* What barriers will older adults face, in terms of interaction design and familiarity with Information & Communication Technology (ICT)? How will older adults with MCI respond to C-TOC, versus older adults who are cognitively healthy?

3. *Will* C-TOC *work at home?* How will results be different from those gathered in clinical settings? Assuming the answer to the first question above is that C-TOC does indeed produce valid results in clinical

settings, how will interruptions and distractions in home settings affect the validity of C-TOC test results?

Ongoing studies by our collaborators at the UBC-CARD continue to address the first question [32], in which researchers and clinicians are examining the correlations in performance between C-TOC and both established pencil & paper screening tests and NPT.

The research presented in this thesis addresses the second and third questions, which are not only of interest to the developers of C-TOC, but also to the broader Human-Computer Interaction (HCI) research community.

Involving older adults and those with cognitive impairment in the development of ICT is an active research area, and is surveyed in Chapter 2 (Section 2.1). We build upon the methods and findings of prior research to evaluate C-TOC in a standard usability study (*Study 1*, Chapter 3).

The third question addresses threats to C-TOC's validity in the home. We have chosen to focus our scope on the effects of interruptions and distractions, however these are not the only issues that we foresee arising in home settings. Computer literacy, motivation, privacy, and cheating are among these important issues, which we intend to address in future work.

The study of interruptions and their effects on primary tasks is well established (See Section 2.2.1). However, in order to answer the third question, we draw upon two research areas: interruptions in HCI and usable ICT for older adults. Therefore, *Study 2* (Chapter 4) makes contributions to both of these areas, and represents the major contributions of this thesis.

1.1.1 C-TOC, Interruptions, & Older Adults

Since users will be accessing C-TOC at home, it is important to address threats to the validity of C-TOC test results in this context. We have focused our attention on the issue of interruptions and distractions that are pervasive in home environments. The sources of these interruptions and distractions include other individuals, pets, appliances, phones, doorbells, computer applications, and events occurring outdoors in the vicinity of the home. Self-initiated interruptions also occur, such as trips to the washroom or the need for food and refreshment. Interruptions may hinder older adults' progress in completing the test, thus affecting their task performance which will in turn affect the validity of test results, as suggested by cognitive ageing literature [6, 16, 24, 56]. Our current research aim is to understand these effects, which will help inform designs for detecting and mitigating interruptions, specifically in C-TOC, and more generally in other applications designed for the ageing population.

Interruptions are common in everyday life, occurring in all contexts (at work, at home, while driving), affecting all people, young and old. Interruptions can have detrimental effects on ongoing tasks, incurring costs to productivity [41] and causing an increase in errors [21].

The effects of interruptions have been studied in a variety of naturalistic and experimental settings, resulting in implications for the design of applications to support productivity [4], decision-making [50], and vigilance tasks [2]. However, these implications largely focus on younger adults in workplace contexts. This research is summarised in Chapter 2 They have only minimally addressed how the ageing mind is affected by interruptions to ongoing primary tasks, which is the focus of the research presented in Chapter 4.

1.2 Thesis Contributions

The major contributions of this thesis include the finding of divergent effects of increased interruption demand between different age groups for different primary tasks. Results suggest that older and younger adults experience interruptions differently, that increased interruption demand can incur a task resumption cost. However, at no age is test performance, in terms of accuracy, compromised by demanding interruptions. The contributions also include design implications for C-TOC; many of these implications are also promising for the design of other applications used by older adults.

Based on the results of a standard usability study, this thesis also contributes a list of design recommendations for improving the usability of C-TOC in future versions. These recommendations were based on concerns identified during an interview study with patients of the UBC Alzheimer's clinic and during a meeting of a cross-cultural panel of senior health practitioners.

1.3 Overview

This thesis comprises of two studies that were designed to evaluate the usability (*Study 1*) and feasibility (*Study 2*) of C-TOC. Previous work relevant to this research is summarised in Chapter 2. Chapter 3 discusses *Study 1*, the first evaluation cycle of C-TOC. Chapter 4 presents *Study 2*, which investigated the effects of interruptions on older adults' C-TOC performance. Chapter 5 discusses directions for future work and concludes this thesis.

Chapter 2

Background & Related Work

Eight years ago, findings by Selwyn et al. [48] suggested that regular Information & Communication Technology (ICT) use was a minority activity among older adults in developed countries. Now in 2011, the baby-boomer generation is reaching the age of retirement (65+). Statistics Canada estimates there are currently 4 million baby-boomers aged 55–65 in Canada [32]. For many of these individuals, ICT use has been a regular part of their working life. A 2009 study by Ofcom [42] found that adults in their 60s were more likely to be using ICT than adults in their 70s. In the coming years, we can expect this generation to keep using ICT at home and throughout their daily lives. Our research hinges on this, as the intended context of use for C-TOC is in the home.

Given this expectation of ICT in the homes of older adults, we consulted prior research relevant to the viability of C-TOC and the methodology of our studies. Topics include older adults & ICT, cognitive ageing, and the effects of interruptions.

2.1 Older Adults & ICT

Study 1, the evaluation of C-TOC.v1 (Chapter 3), and specifically the clinical usability interview component, is informed by a large body of existing research dealing with the design and evaluation of technology for older users and those with cognitive impairments.

2.1.1 Older Adults in HCI Research

Dickinson, Arnott, and Prior [12] have provided several considerations for planning research studies involving older participants. They address several procedural issues, suggest solutions, and propose possible reasons for these issues. These considerations include expecting that a study session with a senior will take about twice as long as it does with a younger subject and acknowledging that older adults will sometimes give responses that they think the experimenter wants to hear, particularly with regards to topics they are not very comfortable with. The authors note that it is therefore important not to lead older adults to answer in any particular way.

Dickinson et al. [12] also observed that older adults are not as confident in their abilities with technology, and will often attribute problems using the technology to their own shortcomings rather than usability issues; as such, they need to be reassured that they are not being evaluated and that the prototype is being evaluated, and that both positive and negative comments are useful. Finally, seniors will often see their participation as a social event in their calendar, so it is in the experimenter's best interest to make them comfortable and be personable. They will often be talkative and provide a wealth of rich data; however, the experimenter must keep the session on track and so may need to steer them back on task.

It is estimated that the prevalence of Mild Cognitive Impairment / Cognitive Impairment–Not Dementia (MCI/CIND) among older adults (55+) in Canada will be between 15–20% during the next 10 years [32]. As a result, cognitive screening is often involved in HCI research with older adults, particularly when equivalence is desired between experimental groups [12]. This is true of *Study 2* (Chapter 4), in which the MOCA [40] cognitive screening test was administered before the experiment begins. Dickinson et al. [12] emphasise the resonance that cognitive screening may have with older adults, eliciting anxiety that may affect subsequent performance. In particular, they may be nervous about undetected cognitive impairment.

2.1.2 Individuals with Cognitive Impairment in HCI Research

The research experience that involves older users with cognitive impairment can be more taxing for both the research coordinators and the participants. Cognitive impairment encompasses CIND, MCI, Fronto-Temporal Dementia (FTD), Lewy-Body Dementia (LBD), and Alzheimer's Disease (AD), among others. For detailed definitions and a review of these clinical conditions and diagnoses, see Dubois et al. [13], Feldman and Jacova [17], and Feldman et al. [18]. As observed in our own study, individuals with CIND can still can provide a great amount of information when interviewed.

Methodologies for Research

Wherton and Monk [58] involved individuals with dementia and their caregivers in open-ended field research aimed at identifying technological opportunities for people with dementia living at home. The design and evaluation of ICT for those with cognitive disabilities can also be conducted in carefully controlled settings without a caregiver present, provided that evaluation sessions are short and relatively low-stress. This is exemplified by a study which had cognitively-impaired individuals use two alternative web browsers in succession, performing carefully chosen browsing and searching tasks; this contributed to research-derived guidelines for cognitively-accessible web design [49]. Interaction with C-TOC.v1 is similarly controlled, as the test battery provides a set of structured tests to be performed in succession.

Involving Caregivers in HCI Research

It is typical for research involving those with cognitive impairment to also involve formal or informal caregivers. This was true of our own evaluation in Chapter 3. While our focus remained on our participants, who are the intended users of C-TOC, several family caregivers were present during some of the clinical interviews, at times contributing to our evaluation. Encouraging the caregiver to participate can also elicit a great amount of useful information in terms of design and evaluation. This has been demonstrated in several studies evaluating memory aid technology for those with memory impairments [25, 35, 60].

2.1.3 Older Adults in ICT Design & Evaluation

Involvement of representative older users in the design and evaluation of C-TOC is preceded by past research involving older users and a spectrum of ICT. Our research approach in *Study 1* (Chapter 3) differs from the following methods in that we did not engage with representative users in group or focus group settings. We nevertheless acknowledge these methods as possible options for future research.

Participatory Design

Massimi, Baecker, and Wu [36] have presented considerations for participatory design with groups of older users, derived from a study in which seniors critiqued and evaluated mobile phones. These considerations include providing alternative activities during group sessions, allowing subgroups of individuals to level out individual differences and deficits, minimising cross-talk during group interviews, providing structured activities, adjusting the pace of activities as necessary, and blending individual and group sessions.

Early Involvement & Mutual Inspiration

Eisma et al. [14, 15] have explored early user involvement and *mutual inspiration* in the context of a user-centred design process. Their *mutual inspiration* strategies included focus groups or group interviews, allowing multiple participants to engage in hands-on activities with prototypes together, and the facilitation of social workshops aimed to build confidence with using ICT. Eisma et al. [14] claim that these strategies are more general than those prescribed by participatory design. While the evaluation of C-TOC.v1 has been conducted in a one-on-one clinical setting, the involvement of groups of participants and their family caregivers could be considered when planning future evaluation.

Web Design for Older Users

Kurniawan and Zaphiris [34] enlisted both older users and HCI experts to establish a set of research-derived web design guidelines for older people. While our own evaluation did not include a panel of HCI experts, future evaluation C-TOC cycles should enlist such a panel to triangulate on a consistent and satisfactory design with a set of representative users.

2.2 Interruptions

Interruptions occurring in the home may have disruptive effects on C-TOC performance, affecting the validity of test results.

This section provides an overview of interruptions and cognitive ageing literature which has contributed to our experimental hypotheses in Chapter 4, Section 4.6, and to our general expectations of how interruptions may affect C-TOC performance in the home.

2.2.1 Interruptions in HCI

Study 2, discussed in Chapter 4, is informed by a well-established body of research investigating the effects of interruptions on primary task performance. We review empirical costs of interruption and discuss factors for predicting these costs.

The Cost of Interruption

The Cost of Interruption (COI) on an ongoing primary task has been examined in naturalistic and experimental settings, and can be defined by several measures.

A coarse COI measurement is the frequency of non-resumption of a primary task following an interruption, leaving the primary task uncompleted [41].

When primary task performance following an interruption is considered, the COI could be measured as the difference in task completion time between uninterrupted and interrupted conditions [63]. This measurement may not capture variation in behaviour immediately following an interruption [39]. A precise measure for examining this interval is the task resumption lag, the time elapsed when switching from an interrupting task back to the primary task [2, 31, 47, 52].

COI can be measured in terms of the difference in a primary task's error rate between uninterrupted and interrupted conditions [21, 43, 50, 51].

Finally, in addition to primary task performance, interruptions occurring at inopportune times can incur a COI in terms of increased reported amounts of stress, frustration, and mental effort, as well as reduced well-being and perceived performance [1].

Many studies, including our own, report on several measures of COI. This is appropriate considering that C-TOC's performance measures, task accuracy and completion time, are also COI measures.

Predicting the Cost of Interruption

Experimental approaches have attempted to isolate factors of primary and interrupting tasks predictive of the COI, however many divergent findings exist in the literature.

Interruption Demand: Increased workload demand of the interrupting task has been shown to be predictive of COI in some cases (Gillie and Broadbent [21], Monk et al. [39]), but not in others. Gillie and Broadbent [21] manipulated the Working Memory (WM) demands of the interrupting task and observed their effects on response accuracy in a WM-dependent computer-based adventure game, which served as their primary task. Interruptions that contained simple one-step mental arithmetic problems were not as disruptive as interruptions containing more complicated multi-step mathematical puzzles. Monk, Trafton, and Boehm-Davis [39] also manipulated the demands of the interrupting task and observed their effects on task resumption time in a simulated VCR programming task, which served as their primary task. Interruptions that contained a simple target pursuit-tracking task (following a target with a mouse as it moves sporadically around the screen) were not as disruptive as interruptions containing a demanding verbal variant of the 'n-back' WM task [44]. The 'n-back involves presenting a sequence of stimuli (in this case, verbally) to an observer, who must respond whenever a stimulus matches what was shown n items prior in the sequence (typically 2 items prior [44]). The task is known to place a high demand on WM. We also adopt this task as a high-demand interruption in our own experiment in Chapter 4.

Low-demand interrupting tasks can sometimes improve performance on a primary task: Zijlstra et al. [63] observed this behaviour when a simulated naturalistic text editing task was interrupted with a relatively simple task, a telephone call requesting irrelevant information. This phenomenon of improved performance on an interrupted primary task, compared to an uninterrupted task, was first reported by Zeigarnik [61] in the 1920s, and is now known as the '*Zeigarnik effect*'.

Oulasvirta and Saariluoma [43] found that if primary task representations can be encoded in WM, increased interruption demand has no effect on primary task performance, otherwise increased interruption workload demand can be disruptive. In their experiment, a reading comprehension primary task was interrupted with a arithmetic verification task, which was deliberately selected as being dissimilar from the primary task. They manipulated the difficulty of the interrupting task (easy and hard arithmetic verification problems) but found no effect on response accuracy in the interrupted primary task.

Primary Task Demand: Bailey, Konstan, and Carlis [4] examined the COI for two interrupting tasks on six primary tasks (adding, counting, image comprehension, reading comprehension, registration, & selection), finding memory demand in the primary task at the point of interruption to be most predictive of COI, rather than any demand incurred by the interrupting task. They hypothesised that adding and counting would have highest WM load, selection and registration the lowest, with image and reading comprehension in between. There were two interrupting tasks which imposed a relatively equivalent WM demand, appearing at the midpoint of a primary task trial: (1) reading and summarising a short news article and (2) interpreting stock quotes and deciding whether to buy or sell. Their hypothesis was verified such that there was an increase in primary task completion time that was positively correlated with the WM load of the primary task at time of interruption. Similarly, an effect of primary task workload demand was also found by Speier, Vessey, and Valacich [50], wherein for highly-demanding primary tasks (multi-step production management tasks involving information gathering, calculation, ranking, and planning), simple information acquisition interruption tasks can inhibit performance. Conversely, they also found that interruptions can actually improve performance on low-demand primary tasks (information acquisition, simple calculations), another instance of the *Zeigarnik effect*.

Primary task WM load at the time of interruption may not account for the whole COI. Salvucci [47] explained that interrupting a primary task which involves complex mental states will result in a reconstruction process upon task resumption, a re-formulation of task goals, rather than a purely memory-based process.

Similarity of Primary & Interrupting Tasks: Gillie and Broadbent [21] found that a high degree of similarity between interrupting and primary tasks was predictive of greater COI (i.e., when tasks interfere with one another, engaging the same cognitive processes, placing similar demands on WM). Performance on their primary task, a WM-dependent computer-based adventure game, was negatively affected when an interruption contained a similar task (a free recall task), but not when the interrupting task was dissimilar (a simple arithmetic task). However, Bailey et al. [4] failed to find an effect of similarity between their six primary tasks and two interrupting tasks.

Interruption Duration: Earlier studies did not find interruption duration to be predictive of the COI (Bailey et al. [4], Gillie and Broadbent [21]), however this finding has more recently been disputed. Gillie and Broadbent [21] interrupted a primary task with short (30s) and long (165s) interrupting tasks. Interruption duration did not in itself have an effect on primary task performance, however when the interruption was both long and similar to the primary task, performance was disrupted. Bailey et al. [4] also found no disruptive effect of interruption duration when interrupting tasks took between 10 and 30s to complete.

Conversely, Oulasvirta and Saariluoma [43] found that increased interruption duration was found to incur a greater COI when primary task representations cannot be encoded into WM (the ability to encode task representations was controlled by pacing the presentation of the primary task trials). Monk, Trafton, and Boehm-Davis [39] have also found that increased interruption duration contributes to lower primary task performance on the primary task, a simulated VCR programming task. They examined 6 interruption durations, ranging from 3s to 58s. They found an effect of interruption duration, in terms of longer task resumption times when interruptions were between 3s and 13s. Longer interruptions, up to 1 minute, incur a task resumption cost that increases asymptotically; this increase is not significantly greater than the cost incurred by a 13s interruption. They resolved that longer interruptions result in a decayed activation of primary task goals as a function of interruption duration, as predicted by previous work [2, 52], requiring a reconstruction of task goals [47].

Interruption Frequency: Zijlstra et al. [63] had mixed findings: in some cases the frequency of interruption was positively correlated with performance on a simulated naturalistic text editing primary task, but negatively correlated with reported well-being .

Interruption Lag & Contextual Cues: Prior research [2, 52] has examined the role of the interruption lag, the brief period of time in which an individual is alerted of an imminent interruption but is still focused on the primary task. This addresses the observation that switching to an interrupting task is seldom an immediate action. A short interruption lag, as brief as $1-2$s, may be sufficient for encoding primary task cues and prospective goals [2]. Retrospective rehearsal following an interruption can retrieve these goals [52]. Both Altmann and Trafton [2] and Trafton [52] conducted experiments with a complex military resource-allocation task serving as the primary task. The interrupting task was one of a simulated radar tracking and classification task. They manipulated length of interruption lag, comparing interruption lags of 2, 4, 6, and 8s. They also manipulated whether the primary task was visible during the interruption lag. When the primary task was visible during the interruption lag, there were no differences in COI (in terms of task resumption lag time) between conditions of increasing interruption lag. When the primary task was not visible, 6s and 8s lags incurred a greater COI.

Hodgetts and Jones [26] have provided further support for the role of the interruption lag, finding decreased COI when an interruption lag is provided, but its advantage is lost when contextual visual cues in the primary task are altered following an interruption. Participants were interrupted (with a mood selection task) while completing a Tower of London task [55] with coloured disks. They benefited from a interruption lag lasting 2s, in terms of reduced task resumption lag times. However when returning from an interruption, if the colour of the disks had changed (the position of the disks being unchanged), the advantage of the interruption lag was lost, manifested by longer task resumption lag times, comparable to conditions where no interruption lag was provided. In this example, the colour of the disk, which was not of central importance to the task, acted as a contextual cue for resuming the task.

Empirical studies regarding the availability of the interruption lag and contextual cues have guided the design of task recovery tools for programming tasks [45]. A timeline diagram of recent activity, visible before and after an interruption, was shown to be helpful for reconstructing context and resuming the primary task.

Primary Task Structure: In research examining the effects of Instant Messaging (IM) on ongoing computing tasks (i.e., searching the web), Cutrell et al. [10] reported interruptions (such as IM) occurring early in a task were more costly than those occurring later on. More recently, Iqbal and Bailey [30] examined characteristics of primary task structure for the purposes of predicting the COI at different interruption onsets. Interruptions occurring during subtasks were found to be more disruptive than those occurring at higher-level subtask or task boundaries. Iqbal and Bailey [30] used document editing as a primary task and a stock quote decision task, similar to one used in earlier work [4], as an interrupting task. They found that an interruption occurring in the midst of writing a sentence will be more costly than one occurring after completing a paragraph or section.

Contextual Factors: Observational and simulated naturalistic studies of interruptions have addressed the many contextual, temporal, and social factors found to be predictive of COI [27, 28]. These have included level of task engagement [19], primary task visibility during an interruption [31], and the source and modality of an interruption (i.e., an intrusion by another individual versus a phone call or IM) [51]. While incorporating social and contextual factors is outside the scope of our laboratory study, we acknowledge the impact of interruption frequency and modality on COI and we will be examining these factors in future work (Section 5.2.2).

2.2.2 The Cost of Interruptions for Older Adults

Related work discussed in the previous section was largely carried out with younger adults, and did not analyse participant age as a factor for predicting COI.

The cognitive ageing literature suggests that interruptions will affect older adults to a greater extent than young adults [6, 16, 24, 56]. Many cognitive processes change as we age, and each may contribute to the COI for older adults.Table 2.1 presents a summary of these changes. Normal age-related changes in cognition can be attributed in part to slower processing speed [46]. Changes in cognition may also be attributed to reduced activation of WM [8]. Prospective memory, the ability to remember intentions, is also inhibited in older adults [54], and can be compromised to a greater extent as a result of interruptions [16]. The ability to suppress reactions to distracting or irrelevant information appears to be reduced [24], contributing to further memory interference [57]. Attention switching is also compromised, as is evident when faced with interruptions and distractions [6]. Given these changes, it is no surprise to find that older adults have a reduced capacity for multitasking [56].

However, age-related cognitive changes are not always marked by losses or drops in functioning; there is evidence [5] for increased brain activity in older adults, compensating for age-related losses in functioning by recruiting

Table 2.1: A summary of age-related cognitive changes relevant to how interruptions affect older adults' task performance.

Reference	Key finding
Salthouse [46]	Slower processing speed.
Craik and Byrd [8]	Reduced activation of WM.
Uttl [54]	Prospective memory inhibited.
Farrimond et al. [16]	Prospective memory compromised by interruptions.
Hasher and Zacks [24]	Inhibited attentional modulation (increased distractibility).
West [57]	Inhibited attentional modulation & memory interference.
Clapp and Gazzaley [6]	Inhibited attentional modulation & WM interference by interruptions.
Wasylyshyn et al. [56]	Reduced capacity for multitasking.
Cabeza [5]	Increased brain activity to compensate for some age-related losses.

additional areas of the brain. As a result, they can at times performing as well as younger adults on WM and visual attention tasks [5].

While older adults have been found to be distractible [24] and have a reduced capacity for task switching [56], few studies have directly addressed the COI for older adults.

Clapp and Gazzaley [6] conducted an experiment to compare WM performance on a facial comparison task between young and old adults in interrupted conditions. The primary task was one in which two images of human faces were compared over a short 500ms delay. In a control condition, the delay was unfilled. In an interruption condition, an interrupting task was given during this delay. The interrupting task was a facial discrimination task in which a single image of a face was shown. Participants were asked if they believed the face to belong to a male over 40 years old. They were to press one of two keys which corresponded to a binary response to this question. Clapp and Gazzaley [6] found that older adults perform disproportionately worse than young adults on the primary task when interrupted.

In terms of attentional modulation, they asserted that older adults attend to interrupting stimuli more than younger adults.

In a naturalistic prospective memory task, Farrimond, Knight, and Titov [16] found that older adults' ability to remember intentions in a primary task was inhibited when it was interrupted with a demanding cognitive task. The primary task involved a simulated street scene and a series of shopping errands. In an interrupted condition, participants completed four 1-minute semantic or verbal fluency tasks during the ongoing primary task. Their results showed that older adults' ability to recall shopping errand cues was reduced in the interruption condition.

2.2.3 Evaluating the Effect of Interruptions on Older Adults Taking C-TOC

To our knowledge, no previous studies have directly compared the COI in the setting of computer-based cognitive testing involving multiple primary task types with older adults. Likewise, design of task resumption cues for reducing the COI has been directed at younger adults in the workplace [37, 45], and may not apply to cognitive testing for older adults. Naturalistic primary tasks used in previous work include event-based adventure games [21], VCR programming [39], production management decision-making tasks [50], resource allocation tasks [2, 39, 52], document and media editing [30, 63], and simulated shopping errands [16]. The task structure of C-TOC tests are considerably different from these tasks. As mentioned above, characteristics of primary task structure have been found to be predictive of COI [30]. Therefore our methodology in *Study 2* (Chapter 4) was designed with careful consideration of the task structure of C-TOC tasks, and does not replicate any single methodology from previous work. Like the Tower of London task [55] used by Hodgetts and Jones [26], our primary tasks are abstract rather than naturalistic, however well-defined for cognitive assessment purposes, designed to engage a specific cognitive process.

Older adults in home settings are also not likely to experience the types of interruptions faced by younger adults in the workplace. This is particularly true of the naturalistic interruptions used in prior research described above, which include radar tracking tasks [2, 39, 52], information acquisition tasks [50], article summarisation tasks [4], and stock quote decision tasks [4, 30]. As such, these naturalistic interruptions do not lend well to our research. As with many of the experiments cited above, our examination of interruptions on older adults' C-TOC performance in Chapter 4 balanced realism and generalisability for increased precision. Our interrupting tasks do not correspond to specific naturalistic interruptions that occur in the home (i.e., intrusions by other individuals in the household, pets, appliances, phones, doorbells, computer applications, outdoor events occurring in the vicinity of the home, and trips to the kitchen or washroom). They are abstract, meant to simulate different levels of WM demand posed by the many possible types of domestic interruptions. Abstract interrupting tasks represent a range of naturalistic interruptions and can be precisely manipulated in terms of their duration and the amount of WM demand they impose. As described above, abstract interruptions have been used in previous experiments, and include the *n*-back' WM task [39], verbal and semantic fluency tasks Farrimond et al. [16], and arithmetic verification tasks Oulasvirta and Saariluoma [43]. It was with this mindset that we designed our own methodology in *Study 2* (Chapter 4).

Chapter 3

Study 1: A Usability Evaluation of C-TOC.v1

This chapter discusses the results of the first evaluation cycle of the C-TOC project. The first cycle of evaluation focused primarily on usability issues, and was conducted in individual clinical and focus group settings.

The planned iterative development of the C-TOC screening tool includes three evaluation cycles, each preceded by the completion of a new version of the tool, as shown in Figure 3.1.

During each evaluation cycle, clinical usability interviews are conducted with patients at the UBC Division of Neurology Centre for Alzheimer's & Related Dementias (UBC-CARD). These interviews are conducted in pairs by a cognitive scientist and a computer scientist specialising in Human-Computer Interaction (HCI).

Concurrently, a cross-cultural advisory panel of community health and social workers convenes once during each evaluation cycle to discuss the content and usability of the current version of the tool. Researchers from various disciplines, including cognitive science, behavioural neurology, and computer science also attend.

The first cycle of evaluation was carried out between May and July 2010, and is the focus of the remainder of this chapter.

Observations regarding the usability of C-TOC.v1, as well as comments

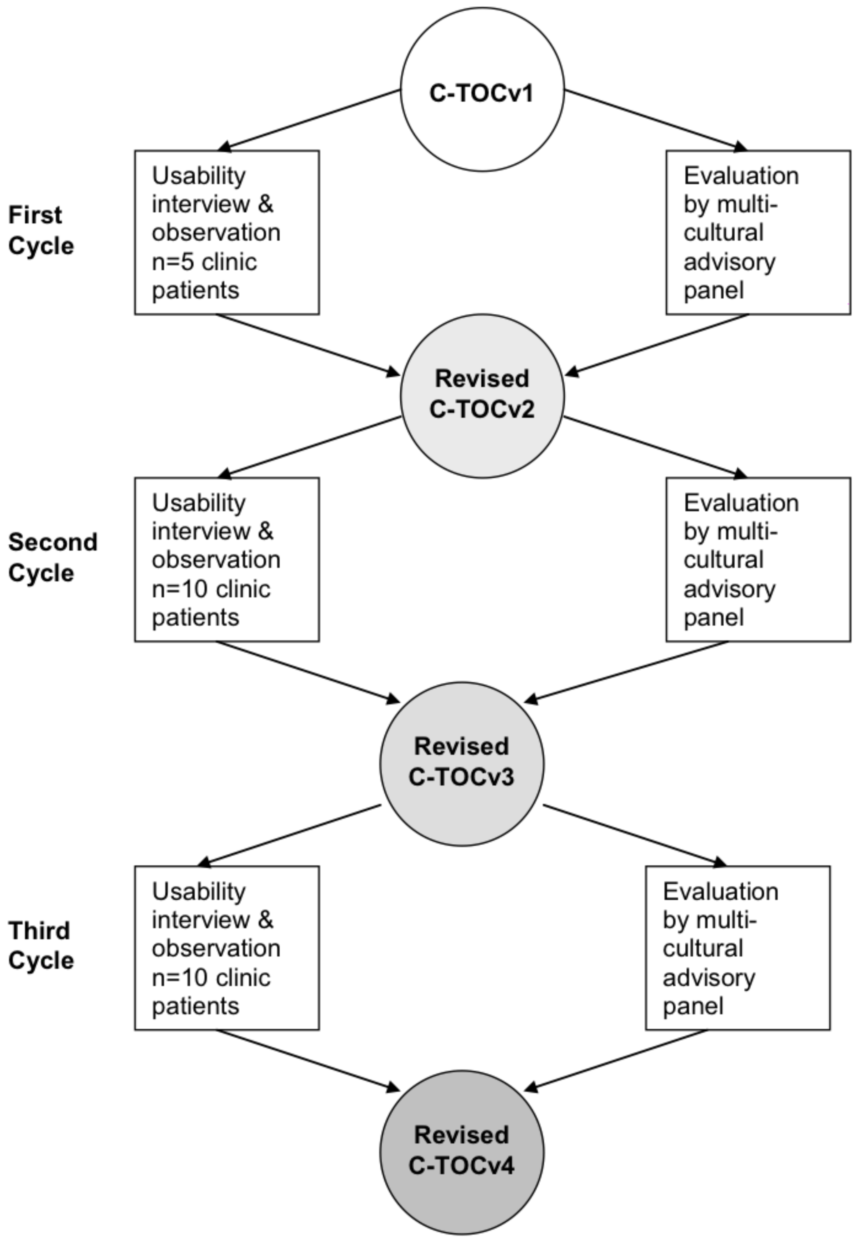


Figure 3.1: The C-TOC evaluation cycle (credit: Claudia Jacova).

Table 3.1: C-TOC.v1 tests and processes tested.

Test name	Processes tested
PICTURE-WORD PAIRS	memory encoding, language
WORD RECOGNITION	cued recognition, language
TEMPORAL ORIENTATION	temporal awareness
SYMBOL-DIGIT MATCHING	processing speed, attention
SIMILARITIES	abstraction
PATTERN CONSTRUCTION	visuospatial skills, executive functioning
PATTERN RECALL	spatial recall
SENTENCE COMPREHENSION	language comprehension, verbal WM
TRAILS	processing speed, executive functioning
ARITHMETICS	attention, WM, processing speed
MISPLACED OBJECT SEARCH	spatial memory encoding
MISPLACED OBJECT RECALL	spatial associative recall
SENTENCE PRODUCTION	language generation
SQUARE PUZZLES	problem solving, spatial WM
GO-STOP	inhibition, executive functioning

and suggestions put forward by study participants formed the basis for the recommendations in Section 3.5. These findings also provide potential directions for future work, discussed in Section 5.2.

3.1 C-TOC.v1

The first evaluation cycle made use of C-TOC.v1, a pilot version of the test battery implemented in a series of interactive Microsoft PowerPoint files. All participants in this study interacted with the prototype on a PC running Microsoft Windows XP. The User Interface (UI) of C-TOC.v1 supports several forms of interaction, including hyperlinks and the ability to move items around the screen.

C-TOC.v1, described in detail in Appendix A, is comprised of an introduction followed by 15 individual tests, listed in Table 3.1. The unfamiliar reader will need to consult Appendix A in order to fully appreciate the recommendations presented in this chapter.

Each C-TOC test contains several trials, including one or two training

Table 3.2: Clinical usability interviewees.

id	age	sex	occupation	diagnosis	MMSE	ICT exp /	use
s1	76	M	education	NCI	30	high	high
s2	70+	M	finance	MCI / AD	23	mod	high
s3	71	F	health care	MCI / AD	20	low	none
s4	68	F	health care	MCI / FTD	29	mod	low
s5	71	F	health care	CIND	30	mod	mod
s6	69	M	education	CIND	28	mod	mod
s7	81	M	agriculture	AD	21	none	none

trials. In these training trials, feedback is given to the user after the trial is completed. Should the user respond to a training trial incorrectly, the trial is repeated until a correct response is made.

3.2 Participants

3.2.1 Clinical Usability Interviewees

Seven clinical usability interviews were conducted between May 17 and July 5, 2010. 4 male and 3 female interviewees ranged in age from 68 to 81 years old ($M = 72.7$). All interviewees were of European descent. Each interviewee lived at home with a companion or caregiver. Interviewee details are listed in Table 3.2.

All interviewees were referred by the UBC Alzheimer's clinic, and at one time had been referred to the clinic to evaluate memory and/or thinking complaints. They did not receive monetary compensation for participating in this study. Diagnoses included Not Cognitively Impaired (NCI), MCI/CIND, early stage FTD, and mild Alzheimer's Disease (AD). The Mini-Mental State Examination (MMSE) was administered to each participant at an earlier visit; their scores ranged from 30 (no impairment) to 20 (at the threshold of mild to moderate impairment). For a review of these clinical conditions and diagnoses, see Dubois et al. [13], Feldman and Jacova [17], and Feldman et al. [18]. Also of note, this range of diagnoses is quite representative of early and recently-assessed UBC-CARD patients.

Some of the computer use background questions of the interview script (See Appendix B.1) were adapted from previous HCI research with older users carried out by Goodman, Dickinson, and Syme [22]. Each interviewee was asked about their Information & Communication Technology (ICT) usage; which ranged from atypically high (s1, a retired physics professor who uses a computer often to prepare presentations and publications) to none (s7, a farmer and salesman who had never previously used a computer).

All but one interviewee (s7) had completed some form of post-secondary education, and had encountered computers in some capacity or another throughout their working years (all but one interviewee, s5, are currently retired). Four reported their first exposure to computers occurring over 20 years ago; another reported first learning to use a computer 15 years ago; the sixth interviewee could not recall when a computer was first used.

All seven interviewees had a personal computer in their home, however in some cases their spouse was the primary computer user. Computer use ranged between often (up to 10 hours a week) (s1, s2), daily (s5, s6), infrequently (i.e., occasionally responding to emails at the suggestion of a spouse) (s4), and not at all (s3, s7); s7, the oldest interviewee, had never before used a personal computer, despite one being owned and used by his spouse.

Of those who used computers daily or often, email and web browsing were stated as the most heavily used computer applications. Interviewee s1 claimed substantial use of document preparation and graphics software, such as Adobe Illustrator, for academic publications and presentations. Interviewee s6 claimed to regularly use most of the applications in the Microsoft Office 2007 suite. Interviewee s2 regularly plays games on a computer. Interviewee s4 had prior experience as a consultant for the redesign of a corporate website, and acknowledged that this experience would be useful for the needs of addressing C-TOC's usability concerns.

None of the interviewees had prior experience with computerised testing, such as DriveABLE, a test meant to simulate aspects of driving a motor vehicle.

None of the interviewees claimed to ever rely on help documentation for learning how to use new computer hardware or software applications.

Some of the interviewees (s2, s3, s4) were accompanied by caregivers or companions, who were also able to provide feedback and additional background information.

While some of the interviewees stated additional medical problems including hand tremors and eyesight difficulties, these did not affect the interaction or the course of the interview. In one case, interviewee s3 and her caregiver withdrew from the interview after completing the C-TOC introduction and the first two tests. This participant had the lowest MMSE score and was too challenged by C-TOC and computer usage. This was due to fatigue; anti-psychotic medication to alleviate agitation rendered many seemingly simple tasks difficult and wearisome for this interviewee. As a result, the majority of results and observations were gathered from the remaining six interviewees.

3.2.2 Cross-Cultural Advisory Panel

The cross-cultural advisory panel focus group for the first evaluation cycle took place on May 27, 2010 at Douglas College in Coquitlam, BC. The panel included five individuals representing the Vietnamese, South Asian, Chinese, Japanese, and Latin American communities. The panel was selected by the Douglas College Douglas College Centre for Health & Community Partnerships (CHCP). The selection criteria was such that panel members' occupations were related to health care, nursing, and social work within their respective community groups. All of the panel members have immigrated to Canada. Also present were the cultural advisory panel coordinators, the project manager from the CHCP, and investigators from the disciplines of cognitive science, behavioural neurology, and computer science. In total, ten users interacted with the C-TOC.v1 prototype.

3.3 Setting & Procedure

3.3.1 Clinical Usability Interview & Observation

Clinical usability interviews were conducted using a PowerPoint compatible laptop PC with an attached mouse, with the C-TOCv.1 prototype folder previously loaded. Interviews with a single interviewee were conducted by two experimenters, and took place in examination rooms at the Alzheimer's clinic, lasting approximately 2 hours in duration. Caregivers were permitted to stay in the examination rooms. Interaction with the prototype was videotaped with the consent of the interviewee, and all interaction was carefully observed by the experimenters. The interview script is provided in Appendix B.1; interviewees were asked about their familiarity with ICT, as well as about general and specific test-specific usability issues encountered while interacting with C-TOC.

3.3.2 Cross-Cultural Advisory Panel

The cross-cultural advisory panel made use of a Douglas College computer lab. The C-TOC prototype was loaded on to each computer. The panel's interactive session lasted for 1.25 hours, during which time each panel member interacted with the entire test battery and concurrently completed a questionnaire addressing C-TOC's content and usability. The questionnaire included general and test-specific questions. The questionnaire is provided in Appendix B.2).

The interactive session was followed by a focus group discussion with the panel members lasting another 1.25 hours. The focus group questions are provided in Appendix B.3).

3.4 Results

This section presents noteworthy findings from the clinical usability interviews and the cross-cultural advisory panel. Given the difference in one-on-one time with participants between the interview study and the cross-cultural advisory panel session, it should not be surprising that the interview data are richer. Hence the majority of the findings arose from the interviews, however additional unique findings from the panel session are mentioned where appropriate.

3.4.1 Reaction to C-TOC

The majority of interviewees had a positive reaction to the idea of a computerised cognitive test battery before experiencing the prototype, and to the idea that the test is self-administered and taken from home.

Of the interviewees who completed the test battery, four claimed they would be comfortable and happy to take the test from home on their own computer. By contrast, interviewee s1 remarked that he would view such an activity as a chore and would prefer the human interaction associated with taking a test in a clinical setting.

There was no unanimous agreement among interviewees with regards to whether they could imagine themselves independently accessing and completing a web-based version of the C-TOC test battery, four interviewees claimed that their test performance would be as good or better at home. One interviewee with mild FTD, s4, claimed that she would likely be more distracted at home: her mind would wander, and thus she would need her caregiver in the room to keep her on task. Additional motivation could originate from the test itself by means of using a timer, a clock icon, or a progress bar. Other interviewees expressed worry over the use of a timer or clock icon, as it could induce additional anxiety and negatively affect the test results.

The targeted duration of the C-TOC test battery (30 min) was perceived as appropriate, however acknowledging that the test could potentially take longer for those who are more severely impaired. However, it is also likely that those with cognitive impairment may rush through the test battery.

Many tests were considered to be fun, appropriately challenging, and an appropriate means of cognitive assessment (from the patient's perspective).

Fatigue and eye strain was encountered during some of the clinical interviews, as up to 2 hours was required to complete the prototype battery and respond to interviews regarding each component of the battery.

Finally, despite the extensive list of design issues and recommendations that follow, there were elements in C-TOC that were perceived as being well-designed. Many tests were seen as easier than analogous pencil and paper tests: on a computer these tests were seen as being more intuitive. Participants praised the use of images and photographs throughout the test battery. There were also elements that elicited no negative feedback, such as the help menu.

3.4.2 User Interface

The introductory slides contain a number of drop-down menus for collecting user information, which are not interactive in the current prototype; this caused confusion for some of our panelists, as the distinction between which UI elements were interactive and which were not was not immediately obvious.

3.4.3 Mouse Interaction

Interviewees and panelists with moderate to high computer expertise were all accustomed to the conventional drag-and-drop mouse action. The mouse action required to move objects does not match the conventional drag-and-drop mouse action used in most current desktop computing environments, and thus there is no transfer from previous learning. The point-and-click mouse action required by the prototype to move objects on screen (clicking once on an object, moving the object along with the cursor, and clicking again to leave the object in place), was difficult for our interviewees and panelists, often resulting in verbally expressed confusion. Interviewee s4 called this interaction 'pretty stupid', but later remarked that explicitly instructing a user to do something other than the conventional drag-and-drop mouse action could be used as a means of assessment in terms of whether the user understands written instructions or is merely skimming the instructions.

Problems with moving objects could be avoided by implementing the test battery on a touch-screen display.

There was also frustration with regard to the number of unprompted mouse clicks required for advancing in the test battery (i.e., there was no instruction to 'click to continue'). Clear navigation buttons on all screens were requested.

3.4.4 Instructions

Interviewee s4 expressed uncertainty while reading test instruction screens. Without a demonstration of the test task or immediate training, she reportedly felt as though she were 'diving in blind' to many of the tests, despite performing well on the tests and declaring that instructions were well understood upon completion of training trials.

We also observed that too much information is overwhelming, and may be particularly difficult for users with a cognitive impairment.

Not all parts of the screen were given the same amount of dedicated attention by all interviewees: instructions delivered at the bottom of the screen may not be obvious. Interviewees favoured instructions at the top of the screen.

Instructions in small font were not universally acknowledged. Interviewee s7 read many of the instructions aloud, but did not notice many lines of instruction displayed in smaller font sizes.

3.4.5 Vocabulary

Interviewee s1 expressed concern over the use of words 'battery' and 'practice', as they may be ambiguous to some users.

'Done' and 'Give Up' buttons, used in several C-TOC tests (i.e., SENTENCE COMPREHENSION (Appendix A.9), PATTERN CONSTRUCTION (Appendix A.7), PATTERN RECALL (Appendix A.8)), may be distracting and can be misunderstood. Additionally, these buttons were in the same physical location as the 'Next' button seen on other screens in the test battery.

3.4.6 Content

Interviewee s6 considered the use of photographic images used in several tests (WORD RECOGNITION (Appendix A.3), MISPLACED OBJECT SEARCH (Appendix A.12), MISPLACED OBJECT RECALL (Appendix A.13), & SENTENCE PRODUCTION (Appendix A.14) to be quite pleasant, and would have enjoyed more photographic images throughout the rest of the test battery.

Interviewee s6 proposed a 'Thank you' page at the end of the test battery, a reward for completing the test battery.

Some interviewees assumed that all tasks were timed, which is not true.

It was argued by cross-cultural panelists that feedback of some form be provided to the user upon completion of the tests, such as the total time elapsed, the number of items correct, and where the user scored relative to population norms or past times the test was taken, whether the user's performance is considered normal or below-normal. It was also suggested that upon completion of the tests, a user is directed to community resources related to mental health awareness, regardless of their performance.

3.4.7 Training Items

More training items were requested, both in terms of practicing mouse actions (clicking on and moving objects on screen), and more training items for each test.

3.4.8 Assistance

When asked about the C-TOC Help Menu (Appendix A.1, Figure A.6), interviewees had very little to say with regards to its contents; it was well-received. Interviewee s7, who has no computer use literacy, expected C-TOC to assist him when he was stuck with a computer- or test-related problem.

The inclusion of a 'I don't understand' button for each test was requested, which would serve to invalidate the results of a single test whose instructions are not well understood.

Cross-cultural panelists suggested that supplemental audio instructions should be offered as an optional feature for all instructions throughout the C-TOC test battery. Other supplemental instruction formats were also suggested, including flow diagrams, animated step-by-step instructions, or comic-strip-style images.

3.4.9 Demographic & Cultural Issues

Cross-cultural panelists agreed that the C-TOC test battery assumes that the user is well-educated, well-integrated (culturally and linguistically), and of a high socioeconomic status.

We learned that some ethnocultural communities would likely be reluctant to the idea of taking a computerised cognitive test at a doctor's office, particularly the Vietnamese communities. Members of the South Asian community would likely not be receptive to a self-administered computerised cognitive test taken from home. Community representatives from these communities cited privacy concerns as the cause of this predicted reluctance.

A previously unrealised use case scenario was brought up by a South-Asian panelist in the discussion, in which older co-located users at a community centre assist one another with setting up and taking the test, with the caveat that they would not assist in the actual test content material.

3.4.10 Test-Specific Observations

There were many observations, suggestions, and comments specific to individual C-TOC test components. These are reflected in the list of recommended changes in Section 3.5. The cross-cultural advisory panel questionnaire (Appendix B.2) results revealed that interaction difficulties may be most prevalent with SENTENCE COMPREHENSION (Appendix A.9), PATTERN CONSTRUCTION (Appendix A.7), PATTERN RECALL (Appendix A.8), and ARITHMETICS (Appendix A.11) tests.

3.5 Recommendations

Based on the observations made during the clinical interviews and cultural advisory panel focus group, design recommendations can be formulated for

C-TOC.v2, many of which have now been implemented.

The following list contains recommendations relevant to the C-TOC test battery's usability, based on the observations, comments, and suggestions emanating from both the clinical interviews and from the cultural advisory panel focus group. The general recommendations are also applicable to any computerised task. As such, they are given precedence and reported first. They are followed by recommendations specific to individual tests.

3.5.1 General Recommendations

User Interface

- Unprompted mouse clicks for the purpose of advancing to the next screen should be replaced with navigation buttons or prompts to click the mouse. Consistent and easily identifiable navigation buttons should be present on all pages, enabled and disabled as necessary or when permitted. Alternatively, a 'See Previous Screen' button may be helpful following screens containing detailed information of instructions (provided this does not interfere with the validity of a test).

- 'Done', 'Next', and 'Give Up' buttons may be distracting or misunderstood. An alternative means of discontinuing tasks must be implemented; some users will attempt to persevere and never achieve the correct solution for a task. This is particularly true for open-ended tasks that require a 'Done' click to advance to the next trial. Buttons that have very different meanings (i.e., 'Next' and 'Give Up') should not have a similar appearance or be placed in the same location on different screens.

- An indicator of progress (i.e., individual test completion, test battery completion, or both), displayed throughout the test battery, or at regular intervals between tests, may serve to motivate a user, and allow her to gauge how long she can expect to spend completing the remainder of the test battery.

- Mouse-over, mouse-down, disabled, and enabled states should be implemented for all interactive elements of the test battery. This applies to navigation and selection buttons, drop zones, shapes, lines, and other drag-able objects. These states will provide important feedback about on-screen to users, indicating which items are available for clicking or moving, and which item is currently being manipulated. All enabled interactive elements should be highlighted when the cursor hovers over them. This will help users distinguish between non-interactive elements from interactive puzzle elements (i.e., those that can be manipulated or moved) in the PATTERN CONSTRUCTION (Appendix A.7) and PATTERN RECALL (Appendix A.8) tests.

- Long check-lists should be segmented into sections, limiting the number of items per screen to 4 or 5.

Mouse Interaction

- Conventional drag-and-drop should be used, resolving the problem of negative skill transfer from conventional drag-and-drop to point-and-click move operations.

Instructions

- Whenever instructions are displayed, they should be presented in a consistent manner in terms of position, colour, and large font size. Sentence clauses should be displayed on separate lines to maximise readability.

- Multi-step text instructions could be delivered incrementally, analogous to a incremental reveal of information in a PowerPoint presentation, with the current step highlighted, allowing the user to refer back to previous steps as necessary without losing her place at the current step.

- Supplemental audio instructions could be offered as an optional feature for all instructions presented throughout the application. Other instruction formats are also possible, such as flow diagrams, animated step-by-step instructions, or comic-strip-style images. Eliminating text instructions altogether and introducing a task via a guided training task may also be a feasible alternative.

Content

- The inclusion of a 'I Don't Understand' button for each test could serve to invalidate the results collected for a single test where instructions are not well understood, but would not affect the data collected from other tests.
- The user should be able to distinguish which tasks are timed and which are not (i.e., displaying a stopwatch icon). A solution would need to be sensitive to the potential anxiety timed tasks may induce. This would also reduce the likelihood of users taking too much time to complete test components.
- A 'Thank you' page at the end of the test battery could reward users for completing. In C-TOC, a 'Thank you' page could also serve to direct users to community resources related to mental health awareness.

Training Items

- More training items are needed, both in terms of practicing mouse actions (clicking on and moving objects on screen), and more training items for each novel task (in C-TOC.v1, there were 1-2 training items per test).

Assistance

- Feedback is often expected. It may need to be explained why feedback is given for certain tasks (i.e. training problems, encoding tasks), but not others (i.e. recall tasks).

- A description of why any particular test or task is useful or necessary, located via a help menu, may serve to motivate a test-taker, giving her a better understanding of what a given test is measuring.
- A 'Give Up?' prompt could appear after a pre-defined period of inactivity.
- Open-ended tasks, such as TRAILS (Appendix A.10) require a means of informing users that they are stuck, either in the form of a prompt or some other form of feedback, appearing after a pre-defined period of inactivity.

3.5.2 Test-Specific Recommendations

The following recommendations are specific to individual C-TOC test components, which are described in Appendix A. Due to their specificity, they may not apply to other computerised tasks.

C-TOC Introduction

Appendix A.1 contains a description of the C-TOC Introduction.

- The checklist of items to be completed before the test begins should be shortened or split between multiple pages. Check boxes should also be made larger, and should not occlude any of the text.
- The navigation buttons in the introduction should be more immediately obvious to users. The help bubble/cloud pointing to the 'Next' button on the first page (Figure A.2) was not obvious to all participants.
- The ambiguous background question of 'highest degree achieved' (Figure A.4) should be changed to avoid any possible confusion.

The Picture-Word Pairs Test

Appendix A.2 contains a description of the PICTURE-WORD PAIRS test.

- Additional prompts are needed after selecting an image in order to proceed. Alternatively, a navigation button should become visible and enabled. This is also needed on feedback pages.
- Unambiguous and prototypical images corresponding to words used in the test are needed. Many of the current images are ambiguous.

The Word Recognition Test

Appendix A.3 contains a description of the WORD RECOGNITION test.

- Clearer instructions to click on the numbered box beside word responses are needed, or alternatively allow the user to click on the word response. Enlarge the numbered response button.

The Temporal Orientation Test

Appendix A.4 contains a description of the TEMPORAL ORIENTATION test.

- Include an 'I don't know button' for each question asked.
- Provide an alternative layout for selecting the day of the month. One participant disliked the arrangement of dates in a grid 7-columns wide, which reminded her of a calendar. This was misleading and confusing as the current month may not have started on a Sunday. A drop-down menu or counter buttons could be more appropriate.

The Symbol-Digit Matching Test

Appendix A.5 contains a description of the SYMBOL-DIGIT MATCHING test.

- Instructions and screen layout for this test should be revised to reduce confusion as to the location of the image response bar and image source bar. Where to perform the requested action must be made more obvious.

The Similarities Test

Appendix A.6 contains a description of the SIMILARITIES test.

- This task requires an example to guide users, prior to completing a training problem.
- Instructions should be placed at the top of each screen, rather than at the bottom.
- Any ambiguity or confusion resulting from viewing 'Sorry, Try Again' screen (Figure A.9) during the training items must be addressed.

The Pattern Construction Test

Appendix A.7 contains a description of the PATTERN CONSTRUCTION test.

- The target pattern and destination pattern areas must be clearly labeled and obvious to the user, potentially with the use of colour-coding. Where to perform the requested action must be made more obvious.
- 'Target pattern' could be re-worded as 'Desired pattern'. Alternatively, drop use of these words altogether, i.e., 'Reproduce the blue pattern with the green shapes'.

The Pattern Recall Test

Appendix A.8 contains a description of the PATTERN RECALL test.

- No target shape area with a '?' is required. The user's working area can be maximised to fit the entire screen, aside from the area occupied by the instructions box. However, a different layout between this test and PATTERN CONSTRUCTION may cause other confusion; further prototyping and testing is needed to address this issue.

The Sentence Comprehension Test

Appendix A.9 contains a description of the SENTENCE COMPREHENSION test.

- Potentially ambiguity in instructions such as 'Stack squares on top of one another' could be resolved by providing a visual example or animation of what is meant by the instruction (in this instance, a stack was either interpreted as a vertical column or as several shapes occluding one another).
- At least for the training trials, this test should automatically advance upon successful completion of the instructed tasks. A 'Done' could be retained for recorded trials.

The Trails Test

Appendix A.10 contains a description of the TRAILS test.

- In order to draw a trail between nodes, clicking on nodes should be required, rather than the current mouse-over action. Currently, only the correct nodes enable mouse-over, such that no errors are possible. Enabling all nodes with a click action would allow for erroneous trails to be created.
- To avoid situations in which digits are mistaken for numbers, and vice versa, such as 'I' and '1', use a serifed font.
- Indications on direction of lines (i.e. arrowheads) could be provided to help direct users with this task. Highlighting the current node (i.e., a different colour or size) may also be necessary.

The Arithmetics Test

Appendix A.11 contains a description of the ARITHMETICS test.

- No recommendations can be made for this test at this time. No outstanding issues with respect to usability were identified for this test. It was generally well-liked by all participants.

The Misplaced Object Search Test

Appendix A.12 contains a description of the MISPLACED OBJECT SEARCH test.

- Include an ‘I don’t know’ or ‘Give up’ option for all trials, with the option of revealing the misplaced object upon giving up.
- A prompt to advance or an automatic advance between trials should be added.

The Misplaced Object Recall Test

Appendix A.13 contains a description of the MISPLACED OBJECT RECALL test.

- Currently, the user must click on the scene where the misplaced object was seen. Alternatively, the user could drag the misplaced object to the scene where it was originally seen. In the second step, the user would then drag the object to a drop zone within the scene, corresponding to the location where the object was seen.

The Sentence Production Test

Appendix A.14 contains a description of the SENTENCE PRODUCTION test.

- The instruction to ‘Use as many words as possible’ should be added. The current instructions do not contain this instruction, which resulted in some short, non-grammatical sentences.
- The sentence drop zone should be made to be larger and more visible (Where to perform the requested action must be made more obvious).
- The ability to occlude words with one another should be disallowed.
- A user should be able to select, move, and delete multiple words already placed in the drop zone, and be able to place words in between words already placed in the drop zone, thus shifting existing words in position to accommodate the new word.

The Square Puzzles Test

Appendix A.15 contains a description of the SQUARE PUZZLES test.

- Add a counter or a way of keeping track how many lines had been moved up until the current time.
- Add on-screen instructions to the training trial(s), as they are currently only provided during test trials.
- Widening the puzzle lines should make them easier to grab and move with the mouse cursor.
- A functional 'Start Over' button was requested. Admittedly, this option would incur a cost to accuracy and completion time.

The Go-Stop Test

Appendix A.16 contains a description of the GO-STOP test.

- The title of this test should be reconsidered. Participants could not form expectations about the test given the current title.
- The possibility of adding a stopwatch to the screen should also be considered. Participants were not aware of the emphasis on response time in this test.

3.6 Summary

The preceding list contained recommendations for improving C-TOC's usability, based on the observations, comments, and suggestions emanating from the clinical interviews and from the cultural advisory panel focus group.

Several suggestions and comments did not factor into this list of recommendations, as they identified aspects of the prototype that were technical limitations of PowerPoint and its limited drag-and-drop functionality. These concerns will be eliminated in the future versions of the prototype, particularly those created programmatically, rather than in PowerPoint.

The results of this study and the subsequent list of design recommendations have since contributed to the many improvements made to the C-TOC prototype. The combination of two perspectives, clinical patients with a range of diagnoses and cross-cultural advisory panelists, was particularly effective in the process of triangulating on usability concerns.

While extensive, we do not believe our findings and recommendations to be exhaustive. Working with more users in subsequent usability evaluation cycles and meetings of the cultural advisory panel have brought us considerably closer to an exhaustive list of usability issues. Many of these issues have now been resolved in the current C-TOC prototype.

As the study documented in this chapter was conducted prior to defining the scope of this thesis and the focus of *Study 2* (Chapter 4), we did not discuss the effects of interruptions occurring in the home and their effect on C-TOC performance with our participants in *Study 1*.

Many of the general recommendations (Section 3.5.1) are also applicable to other applications used by older adults and those with cognitive impairment. This includes providing salient navigation options and easy-to-understand and accessible instructions in consistent locations on the screen. Our recommendations also include the provision of a balance between words and pictures, serving to reduce possible anxiety and distraction. Applications should provide opportunities to practice novel tasks while allowing previous skill transfer. Helpful, consistently-worded feedback should be provided when needed. Finally, applications should allow for a clear differentiation between interactive and non-interactive UI features.

Chapter 4

Study 2: Effects of Interruptions on Older Adults' C-TOC Performance

As users will be accessing C-TOC at home, it is important to consider interruptions that are pervasive in home environments. These include intrusions by other individuals in the household, pets, appliances, phones, doorbells, computer applications, and outdoor events occurring in the vicinity of the home. Interruptions can also be self-initiated, such as trips to the kitchen or washroom. Some of these interruptions place a greater demand on cognitive resources than others. Interruptions may hinder older adults' progress in completing C-TOC tests, potentially affecting the validity of test results. The research goal of this chapter is to understand these effects, which will help inform designs for detecting and mitigating interruptions, both in C-TOC and generally in other applications used by older adults.

We conducted an experiment to determine the effects of interruption on older adults' performance on two primary tasks adapted from C-TOC tests. A group of old adults (70+) was compared against two other age groups (19–54, 55–69).

We expected to find an increased cost of interruption for older adults, with disproportionally worse performance as interruption workload increases.

As in previous work (see Section 2.2.1), we measured COI on primary task performance in terms of task resumption lag, completion time, and accuracy. We manipulated primary task type and interruption workload demand as independent variables. Our experiment maintained fixed levels of interruption duration, frequency, and lag visibility and duration.

4.1 Methodology

4.1.1 Primary Tasks

Two primary tasks were used in this study, adapted from C-TOC test components: a verbal Working Memory (WM) sentence comprehension task (Appendix A.9) and a spatial problem-solving puzzle task (Appendix A.15), hereby referred to as the VERBAL task and the SPATIAL task. These tasks were chosen on the basis of (1) they place demands on different cognitive processes (see Table 3.2) and (2) relative to other C-TOC tests, individual trials of these tasks take longer to complete, providing us with a longer period of time in which a participant can be interrupted. The tasks are similar in that both tasks place a demand on the central executive of WM [3]. The VERBAL task engages the phonological loop of WM; the SPATIAL task engages the visuospatial sketchpad of WM.

The Verbal Task

In a single trial of the VERBAL task, participants arrange geometric figures according to an instruction. A trial constitutes an instruction (Figure 4.1), which is read before advancing to the execution step (Figure 4.2), at which point it is no longer accessible; thus the participant must hold the task instruction in verbal WM. Trial instructions (see Appendix C.6.1) contain several sentence clauses, each requiring the participant to maintain information in WM (i.e., a shape, a colour, an orientation). After executing the instruction, clicking the 'Next' button at the bottom right of the screen completes the trial.

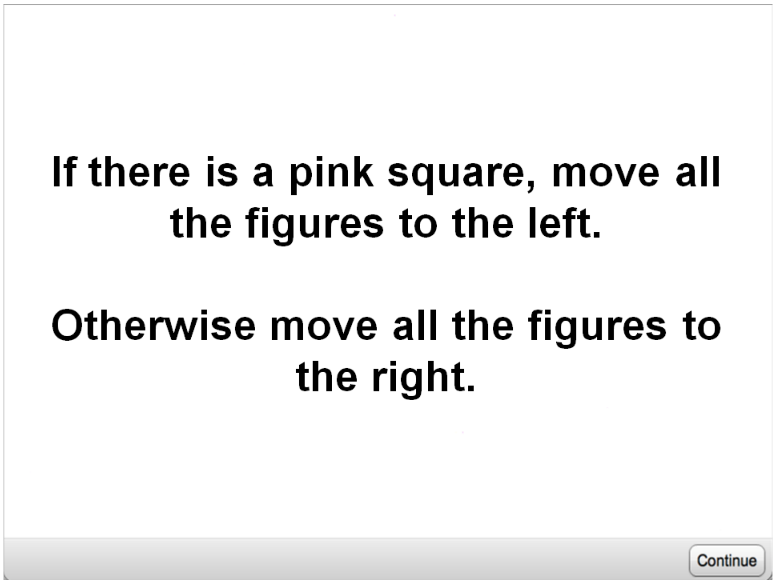


Figure 4.1: The VERBAL task (instruction screen).

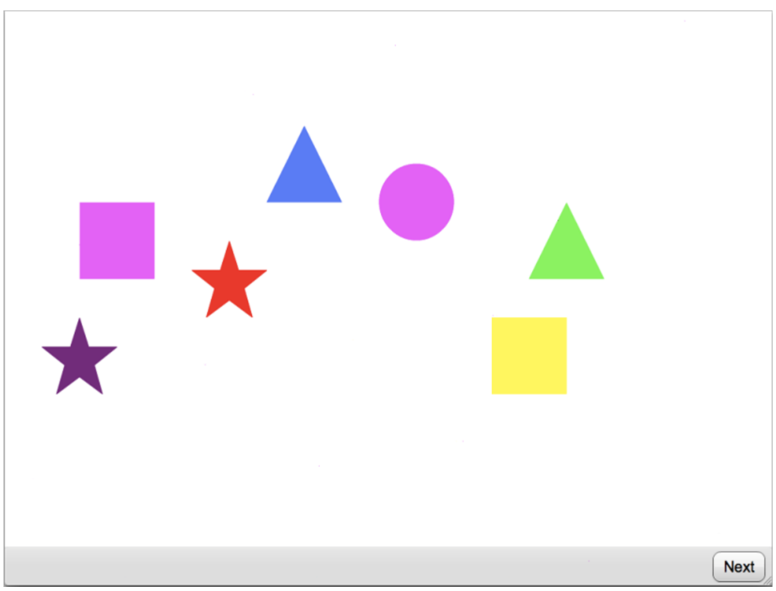


Figure 4.2: The VERBAL task (execution screen prior to any user interaction).

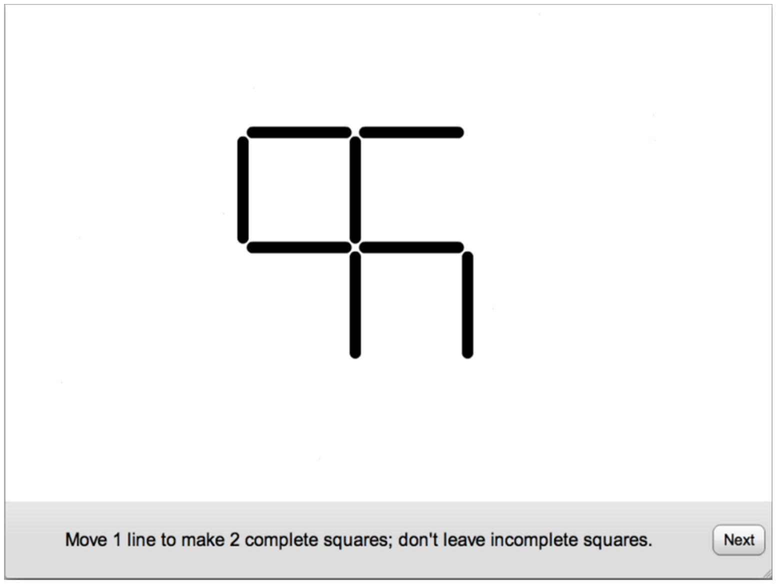


Figure 4.3: The SPATIAL task (initial view).

The Spatial Task

In the SPATIAL task (Figure 4.3), a single trial constitutes one puzzle. The participant is given an array of vertical and horizontal lines, and is instructed to arrange the lines in order to create a number of complete squares in a specified number of moves (Figure 4.4); instructions remain visible throughout a trial. The SPATIAL task also involves WM, albeit in a different modality. Furthermore, less information is maintained in WM than in the VERBAL task. The participant must recall the number of moves she has already made, the number of moves remaining, and/or the original puzzle configuration. After solving the puzzle, clicking the 'Next' button at the bottom right of the screen completes the trial.

4.1.2 Three Interruption Conditions

Three interruption conditions were used in this study. The first of which was an UNINTERRUPTED control condition (no interruptions occurred during these primary task trials). In the two interruption conditions, interrupting

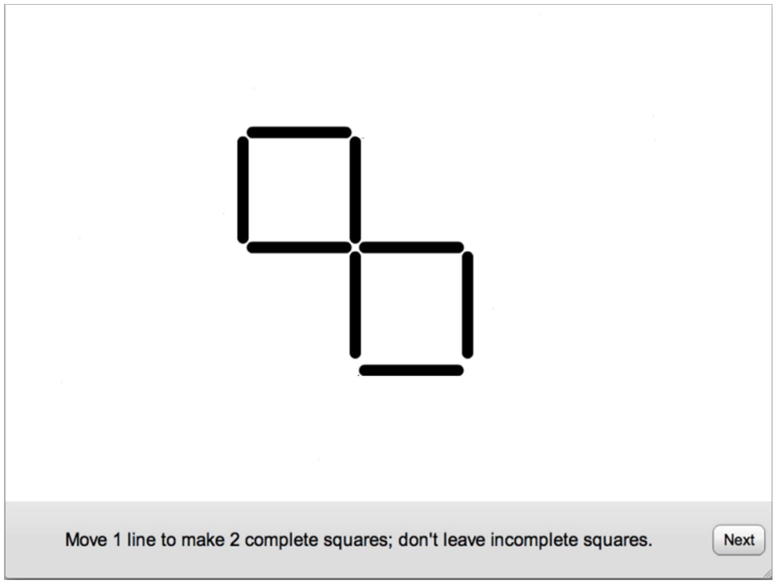


Figure 4.4: The SPATIAL task (same view, completed trial).

tasks were presented to the participant during a subset of primary task trials. Interrupting tasks filled the entire screen, occluding the primary task. In both types of interrupting tasks, an automated randomised sequence of a dozen cartoon images was shown at a rate of 1 image every 1.5s. After these images were shown, the participant was prompted to click in order to dismiss the interruption and return to the interrupted primary task trial. Total interruption time was roughly 20s (18s for the stimuli followed by the time taken by the participant to respond by clicking to dismiss the interruption).

The interrupting tasks are meant to simulate different levels of WM demand posed by interruptions which occur in the home. Some of the interruptions mentioned in the first paragraph of this chapter place greater loads on WM than others. It is not our intention to map our interrupting tasks to specific household interruptions, but rather to represent a range of possible naturalistic interruptions.

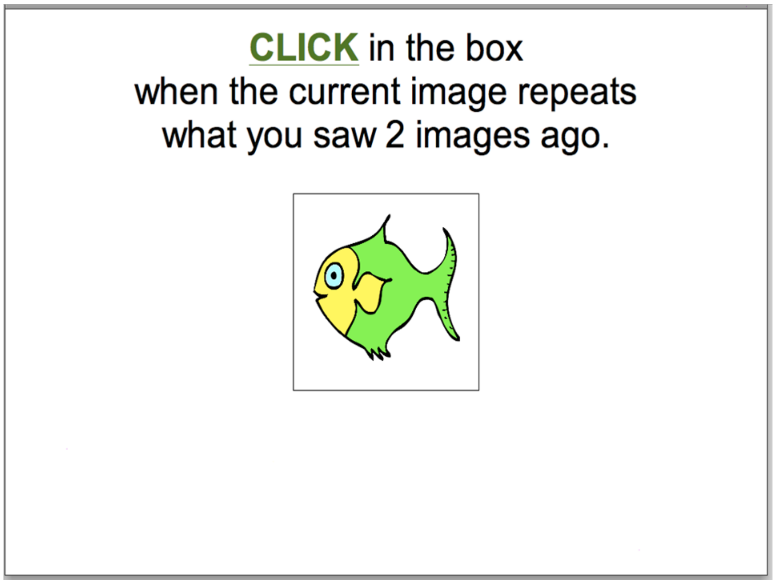


Figure 4.5: The ACTIVE interruption task.

The Active Interruption Task

A high-demand ACTIVE interruption (Figure 4.5) places a demand on WM, intended to simulate an interruption that an individual might experience in the home, one that requires a shift in concentration from the primary task. For this we used an instance of the established 'n-back' WM task [44], in which participants must monitor the series of images and click whenever an image is presented that is the same as the one presented two images prior in the sequence, hence we used the '2-back' variant of the task. In a 2005 meta-analysis of 24 'n-back' studies, Owen et al. [44] found that 23 of them used the '2-back' variant of the task (10 studies used a '1-back' variant, 8 studies used a '3-back' variant). As described in Chapter 2, Monk et al. [39] used a verbal variant of the 'n-back' task as an interrupting task. As the 'n-back' is known to place a demand on WM [44], we believed it would interfere with the WM demands of the primary tasks. In our implementation of the 'n-back' task, feedback is displayed to the participant following correct and incorrect responses.

Figure 4.6: The PASSIVE interruption task. Participants are instructed to watch the sequence of images passively until prompted to dismiss the interruption.

The Passive Interruption Task

In a low-demand PASSIVE interruption (Figure 4.6), participants were instructed to watch the sequence of images passively until prompted to dismiss the interruption. This task was intentionally chosen to be similar to the ACTIVE interruption task (similar visual stimuli, identical duration, both required a click to dismiss after 18s), however it places no demand on WM, nor does it require a motor response (i.e., clicking on the stimuli). The inclusion of this condition allows us to determine if an interruption requires a WM demand to affect primary task performance.

4.1.3 Coordination of Primary and Interrupting Tasks

The two primary tasks each had three blocks of trials, one for each of the three interruption conditions, counterbalanced across participants. Three isomorphic sets of trials were randomly allocated per participant to each block. Task instructions were unique between and within each set of trials, and corresponding trials between sets were isomorphic in terms of instruction complexity, as described in Appendix C.6.

In the PASSIVE and ACTIVE conditions, a subset of trials contained interruptions, as shown in Figure 4.7. This subset was selected at random for each participant; a subset of 4 out of 10 trials were interrupted in the VERBAL task while a subset of 3 out of 8 trials were interrupted in the SPATIAL task. For example, a participant interrupted on trials 2, 5, and 7 in the PASSIVE condition would also be interrupted on trials 2, 5, and 7 in the ACTIVE condition. This was necessary as corresponding trials between blocks had isomorphic task instructions.

Interruption onsets were fixed for each VERBAL trial, and would occur between 1s and 3s into the execution step, typically during or before a first move action is attempted. Therefore, the interruption onset was the same between PASSIVE and ACTIVE conditions for trial n. In SPATIAL trials, interruption onsets would occur 0.5s after the first or second completed move action, such that one or two outstanding moves were required after the interruption. Different interruption onsets for the two primary tasks was necessary due to differences in task structure and task completion time observed in pilot studies.

Following from the findings of related work [2, 52], our interrupting tasks were preceded by an interruption lag lasting 2s. During the interruption lag, the primary task is still visible but interaction is disabled; meanwhile a highly-salient interruption notification appears at the top of the screen (see Figure 4.8). As discussed in related work [2, 52], two seconds provides a sufficient amount of time to encode task goals and form cues for task resumption.

Following the interruption lag, the interrupting tasks occupy the entire screen, the intent being to disrupt the primary task to a greater extent than what a partially-occluding or non-occluding interrupting task could accomplish [31, 51]. In the context of the home, this could represent an interrupting activity occurring somewhere other than on the computer screen, or when an interruption is caused by another computer application, which is full-screen focus.

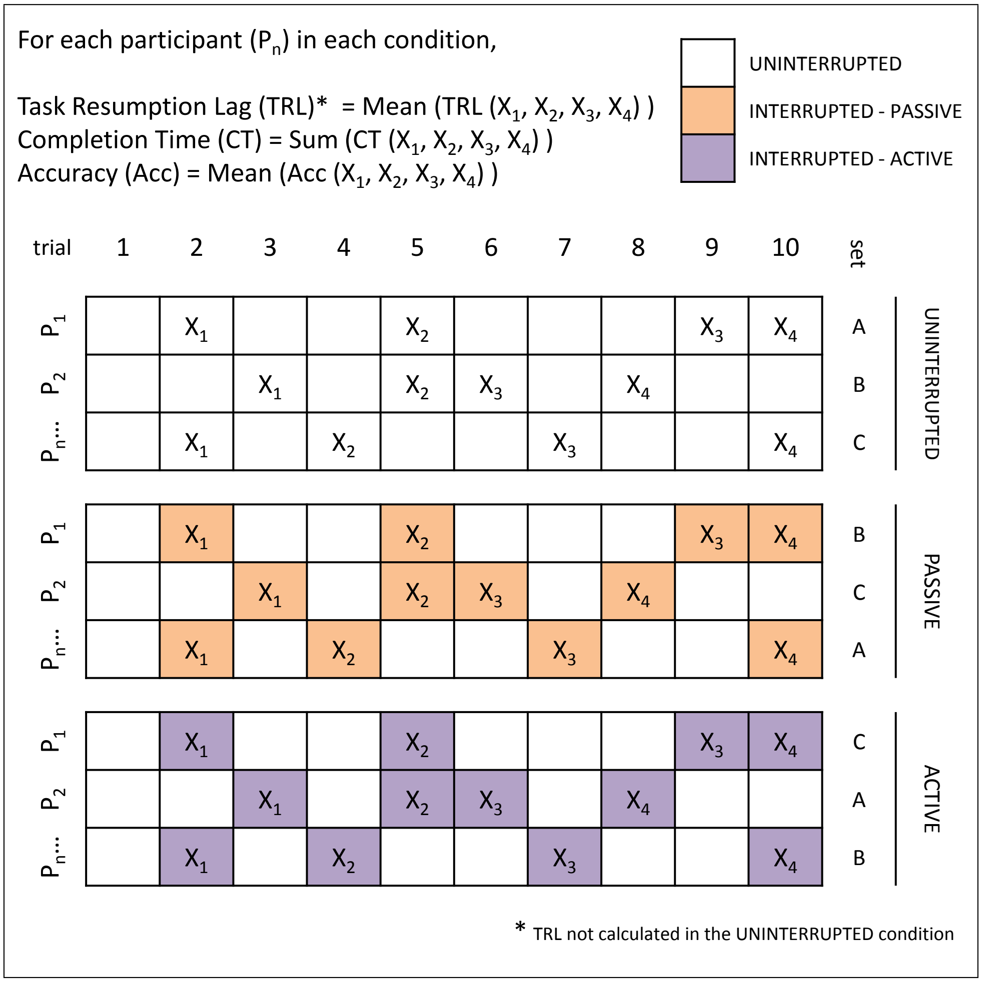


Figure 4.7: An illustrated example of the coordination of interruptions across conditions. For both primary tasks, three isomorphic sets (A, B, C) of trials were randomly allocated per participant to each block (one for UNINTERRUPTED, ACTIVE, & ACTIVE). A subset of trials ($X_{1,2,3,4}$) contained interruptions in the PASSIVE and ACTIVE conditions; this subset was selected at random for each participant. Quantitative measures were calculated for this subset in each condition.

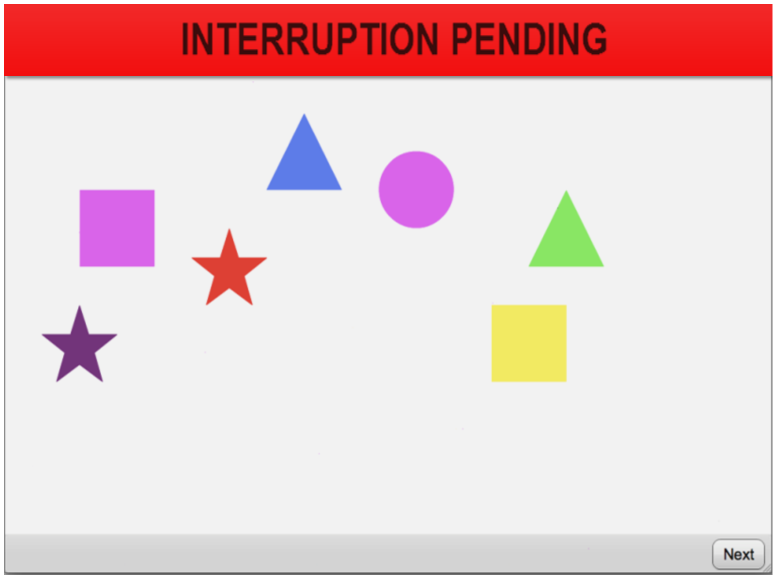


Figure 4.8: The interruption lag: interrupting tasks were preceded by an interruption lag lasting 2s. During the interruption lag, the primary task is still visible but interaction is disabled.

At the end of the interrupting task, the user is prompted to click to dismiss the interruption, returning to the interrupted primary task at the point where it was interrupted.

4.1.4 Design

A 3 x 3 x 2 mixed design was used; age (YOUNG, PRE-OLD, and OLD) was a between-subjects factor, and level of interruption demand (UNINTERRUPTED, PASSIVE, or ACTIVE) and primary task type (VERBAL, SPATIAL) were within-subject factors. All 36 participants experienced all combinations of interruption demand and primary task type.

Order of presentation for the within-subjects factors was fully counterbalanced, such that a participant began with either the Verbal or Spatial task blocks.

4.2 Measures

4.2.1 Quantitative Measures

Three dependent measures were recorded for both primary tasks: Task resumption lag time following an interruption, primary task completion time, and primary task accuracy. We analsyed data from the subset of trials in which interruptions occurred. For completion time and accuracy, we also analysed the corresponding subset of trials from the UNINTERRUPTED condition, as denoted in Figure 4.7.

Task Resumption Lag Time

We measured task resumption lag as the mean time elapsed between returning to the primary task following an interruption and the completion of the first subsequent valid move (i.e., dragging a shape or line to a new position). This was measured in the subset of interrupted trials in the PASSIVE and ACTIVE conditions.

Task Completion Time

Beginning at trial onset, we measured task completion time as the total *uninterrupted* time elapsed completing the subset of interrupted trials in the PASSIVE and ACTIVE conditions, which included task resumption time. The total time to complete the corresponding subset of trials was measured in the UNINTERRUPTED condition.

In the VERBAL task, trial completion time was counted from the start of the execution step to the clicking of the 'Next' button, signalling the completion of the instruction. We did not include the time spent reading trial instructions as part of trial completion time.

Accuracy

We measured accuracy according to a clinical scoring scheme used for C-TOC (see Appendix C.6), based on scoring scheme used in the *Token test* [11]. A mean percentage score was calculated across the subset of interrupted trials in the PASSIVE and ACTIVE conditions, and for the corresponding subset of trials in the UNINTERRUPTED condition. This scheme accounts for the number of moves and correct relative positioning of shapes or lines, allowing for partially correct responses.

Active Interrupting Task Score

Accuracy in the ACTIVE interrupting task was also recorded, measured as the number of correct responses.

4.2.2 Qualitative Measures

Questionnaire

Subjective data concerning task difficulty and demand was collected on a questionnaire following each condition. The questionnaire was adapted from the NASA Task Load Index (NASA-TLX) [23], a standardised instrument for assessing various dimensions of workload. Six questions were posed regarding mental and physical demand, annoyance, perceived performance, and fatigue. Responses were along a 10-point scale. The questionnaire is provided in Appendix C.7.

Interview

At the end of the study, we interviewed participants to probe their perceptions of task difficulty and their task resumption strategies. The interview script is provided in Appendix C.5.1.

4.3 Apparatus

A 2.26 GHZ Core 2 Duo laptop with 1.92 GB RAM and an a 2.0 GHZ Pentium M laptop and 1.5 GB RAM, both with Microsoft Windows XP, were used for the experiment.

The experimental software was written using the Adobe Flex 4.0 SDK (see Appendix C.4). The system recorded all timing and interaction data. We did not make use of the PowerPoint prototype used in *Study 1* (Chapter 3).

Table 4.1: Laboratory study participants.

Group	Age Range	Gender
YOUNG *(19-54)*	19–50, (M = 31.0)	8F / 4M
PRE-OLD *(55-69)*	57–69, (M = 63.4)	9F / 3M
OLD *(70+)*	70–86, (M = 74.8)	6F / 6M

A screen-capture application was also used to record all experimental sessions.

The majority of experimental sessions were conducted in the UBC-CARD, where the laptops were connected to a 17" LCD monitor, with a 1024 x 768 resolution. Six sessions were conducted remotely without an external monitor; in these sessions the laptop display was used, also using a 1024 x 768 resolution.

An optical mouse was used with both laptops; identical mouse gains and tracking speeds were used.

4.4 Participants

Thirty-six healthy participants (no cognitive or motor impairments) were recruited from three age groups (12 each), listed in Table 4.1.

The justification for these age groups rests on age-related changes in cognition that occur around the ages 55 and 70 [9]: higher cognitive function remains relatively stable up to about age 55, after which there is a small decline, followed by a much steeper decline after 70.

We did not control for levels of formal education and computer literacy. We acknowledged that our YOUNG participants generally had higher levels of formal education and computer literacy.

Participants were recruited through advertisements placed throughout the community (see Appendix C.1), and received $5 for each half hour of participation.

Two participants were removed from the analysis. One participant's behaviour clearly showed that she was not following our instructions, and her comments at the study indicated that she was confused about the two types of interrupting tasks. We used a screen-capture application to record the on-screen activity of all participants; one participant's video capture file was corrupted and thus scoring could not be completed for him.

4.5 Procedure

The experiment was designed to fit into a single 90 minute session.

We first administered the MOCA [40], a screening measure to help ensure that participants had no existing undiagnosed cognitive impairment (see Appendix C.3.1). Additionally, we administered the North American Adult Reading Test (NAART) [53] to help ensure participants had sufficient English fluency to follow our instructions (see Appendix C.3.2). Cutoff criteria were used for both tests: participants required a score of 26 or higher (out of 30) on the MOCA and were required to read at least 25% of words used in the NAART correctly. Based on these criteria, we excluded five participants (not included in the 36 above). They were allowed to finish the study, but their data were not included in our analysis. Participants who scored less than 26 on the MOCA were later contacted by clinicians in our research group to arrange further consultation.

Participants were given examples of the interrupting tasks and asked to practice the ACTIVE interrupting task until they were familiarised with it.

Participants then completed 4 blocks of trials for both primary tasks. The first block in each task was a short 4-trial practice block containing UNINTERRUPTED and interrupted trials. The remaining three test blocks in the VERBAL task contained 10 trials, while test blocks in the SPATIAL task contained 8 trials. The number of trials in a test block were representative of the number of trials appearing in the corresponding C-TOC tests. All trial instructions are provided in Appendix C.6.

Participants were asked to complete each trial as quickly and as accurately as possible, as we desired to reduce the effect of any speed or accuracy bias, as described by Zhai et al. [62]. After each test block, participants filled out a copy of the questionnaire.

Once all blocks were completed, participants were interviewed.

4.6 Hypotheses

H1. Age & Interruption Demand

1. *Overall,* YOUNG *adults will perform better than older (*PRE-OLD, OLD*) adults on the primary tasks.*

2. *Older (*PRE-OLD, OLD*) adults will incur a disproportionately larger* COI *when interruption demand increases.*

H2. Age, Task & Interruption Demand

1. *Given that the* VERBAL *task places a greater load on* WM*, increased interruption demand will incur a disproportionately greater* COI *on the* VERBAL *task than on the* SPATIAL *task.*

2. *This difference in* COI *will be greater for older (*PRE-OLD, OLD*) adults.*

4.7 Results

Task resumption lag and task completion time results were log-transformed, correcting for positive skews. We performed a 2 x 3 (level of interruption demand x age) Analysis of Variance (ANOVA) on the task resumption lag data and a 3 x 3 (level interruption demand x age) ANOVA on the completion time data. The accuracy data for both tasks was negatively skewed, so we performed nonparametric factorial 3x3 ANOVAs using the Aligned Rank Transform (ART) [59], a method that can accommodate repeated measures designs and examine interaction effects. All pairwise comparisons were protected against Type I error using a Bonferroni adjustment. We report on measures that were significant ($p < .05$) or represent a possible trend ($p < .10$). We do not report non-significant or unhypothesised findings. Along with statistical significance, we report partial eta-squared (η^2), a measure of effect size. To interpret this value, 0.01 is a small effect size, 0.06 is medium, and 0.14 is large [7]. We report on data from 36 participants.

4.7.1 Verbal Task

Task Resumption Lag Time

Task resumption lag increases with age and interruption demand. With ACTIVE *interruptions,* OLD *adults appear to be disproportionally slower than* YOUNG *adults to resume the task* (Figure 4.9).

The main effect of age was significant ($F_{2,33} = 11.89$, $p < .001$, $\eta^2 = .419$); pairwise comparisons showed that OLD and PRE-OLD adults were slower to resume the task than YOUNG adults ($p < .001$, $p = .002$, respectively). The main effect of interruption demand is also significant ($F_{1,33} = 35.99$, $p < .001$, $\eta^2 = .522$); participants were slower to resume the task following ACTIVE interruptions than PASSIVE ones ($p < .001$). A trend suggests, as hypothesised, an interaction between age and interruption demand ($F_{2,33} = 3.22$, $p = .053$, $\eta^2 = .163$). Pairwise comparison on the interaction effect showed that OLD and PRE-OLD adults were slower to resume the task following ACTIVE interruptions than PASSIVE ones ($p < .001$, $p = .005$, respectively).

Task Completion Time

Completion time increases with age and interruption demand (Figure 4.10).

Closely mirroring the resumption lag results, there was a main effect of age on completion time ($F_{2,33} = 12.00$, $p < .001$, $\eta^2 = .421$). Pairwise comparisons showed that OLD adults were slower than YOUNG adults ($p < .001$). There was also a significant main effect of interruption demand ($F_{2,66} = 12.47$, $p < .001$, $\eta^2 = .274$). Pairwise comparisons showed that completion times were longer in the ACTIVE than in the PASSIVE and UNINTERRUPTED conditions ($p < .001$, $p = .006$, respectively). Unlike the resumption lag results, however, there was no interaction of age and interruption demand.

Accuracy

OLD *are less accurate than* YOUNG (Figure 4.11).

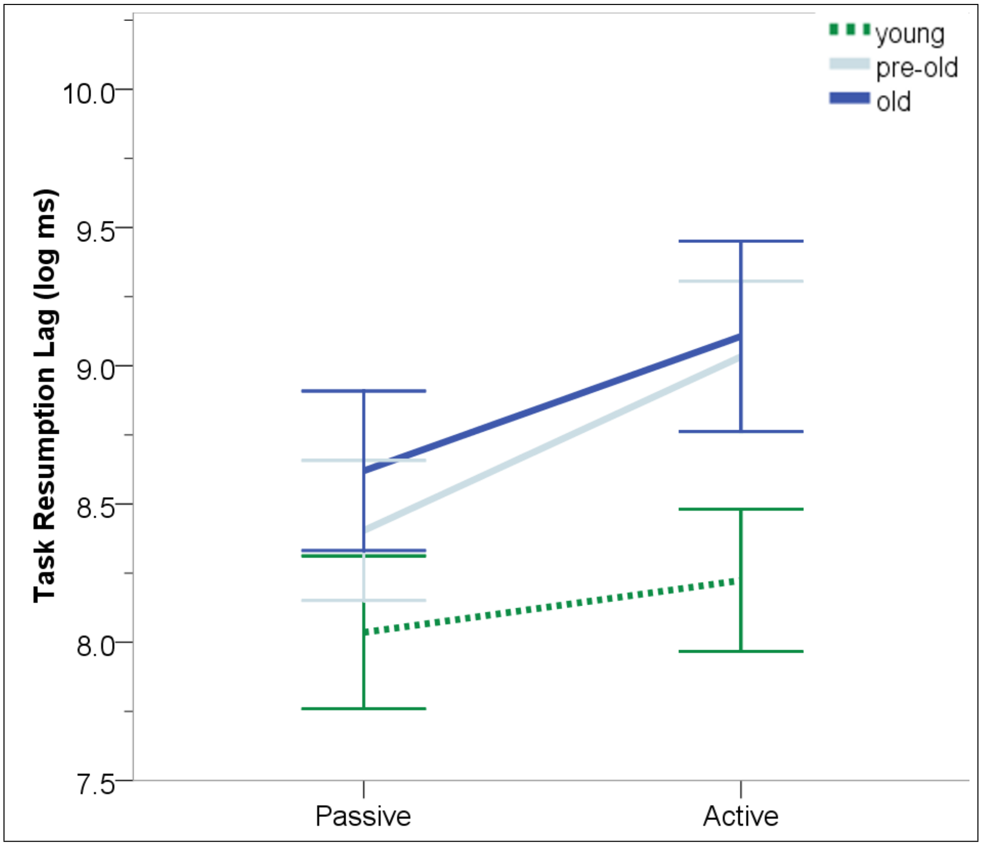


Figure 4.9: VERBAL task resumption lag time (log ms).

The main effect of age was significant ($F_{2,33} = 10.46$, $p < .001$, $\eta^2 = .388$), where OLD adults were less accurate than YOUNG adults ($p = .001$). However, the latter performed at ceiling levels, so the observed effect of age is likely to be smaller than it would have been in the absence of a ceiling effect. Given this result, we are unable to predict if an interaction between age and interruption demand exists.

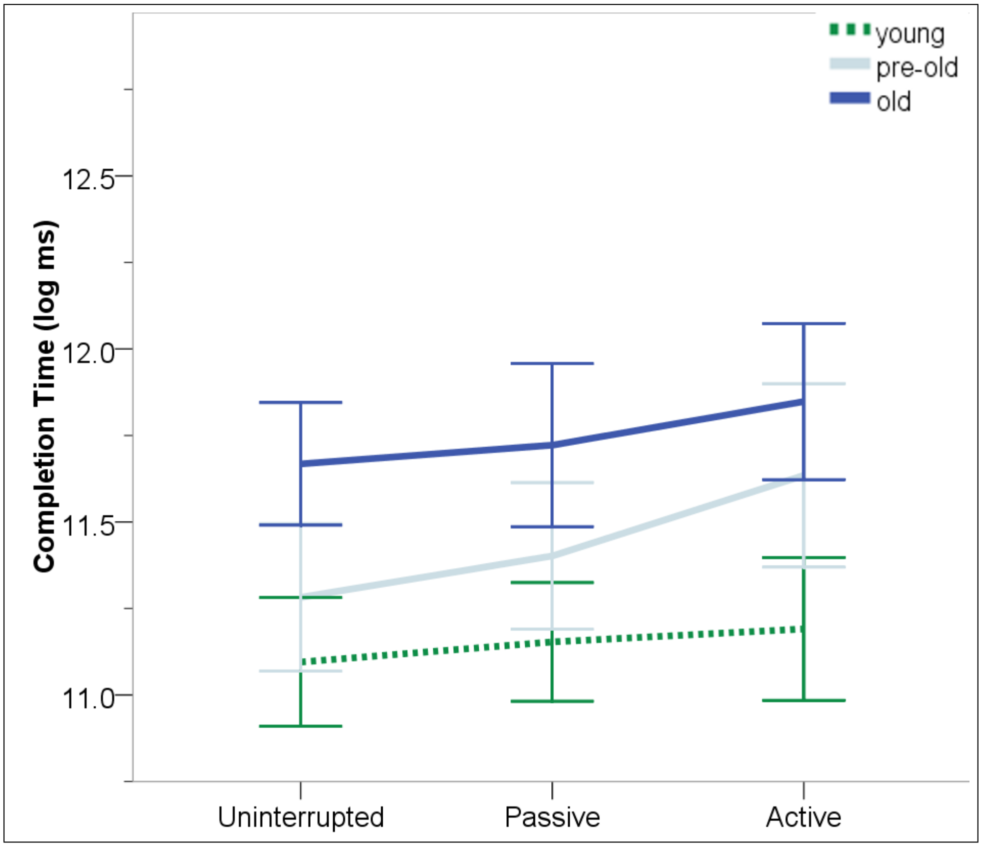


Figure 4.10: VERBAL task completion time (log ms).

4.7.2 Spatial Task

Task Resumption Lag Time

Task resumption lag increases with age, however YOUNG *resume the task faster with* ACTIVE *interruptions, whereas* OLD *do not* (Figure 4.12).

The main effect of age was significant ($F_{2,33} = 3.40$, $p = .046$, $\eta^2 = .171$). The effect of interruption demand was not significant, however the interaction between age and interruption demand was significant ($F_{2,33} = 5.60$, $p = .008$, $\eta^2 = .253$). Pairwise comparison on the interaction effect showed that YOUNG adults resumed the task faster in the ACTIVE condition than in the PASSIVE condition ($p = .019$).

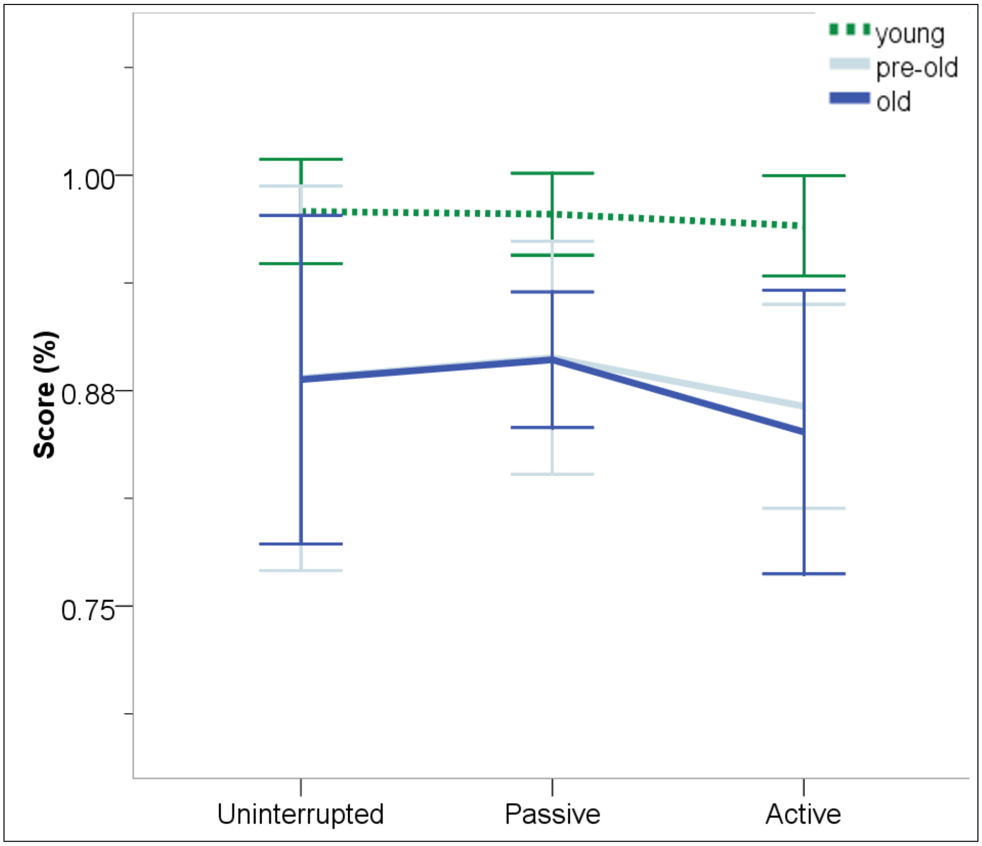


Figure 4.11: VERBAL task accuracy (%).

Task Completion Time

Completion time increases with age, OLD *are slower than* YOUNG *except in the* PASSIVE *condition* (Figure 4.13).

There was a main effect of age ($F_{2,33} = 4.09$, $p = .026$, $\eta^2 = .199$). Pairwise comparisons showed that OLD adults were slower than YOUNG adults ($p = .022$). There was no significant effect of interruption demand. However, different levels of interruption demand impacted the age groups differently (interaction effect: $F_{4,66} = 3.28$, $p = .016$, $\eta^2 = .166$). Pairwise comparison on the interaction effect showed that OLD adults are slower than YOUNG adults in the UNINTERRUPTED condition ($p = .010$). Surprisingly, there were no significant differences between groups in the PASSIVE condition.

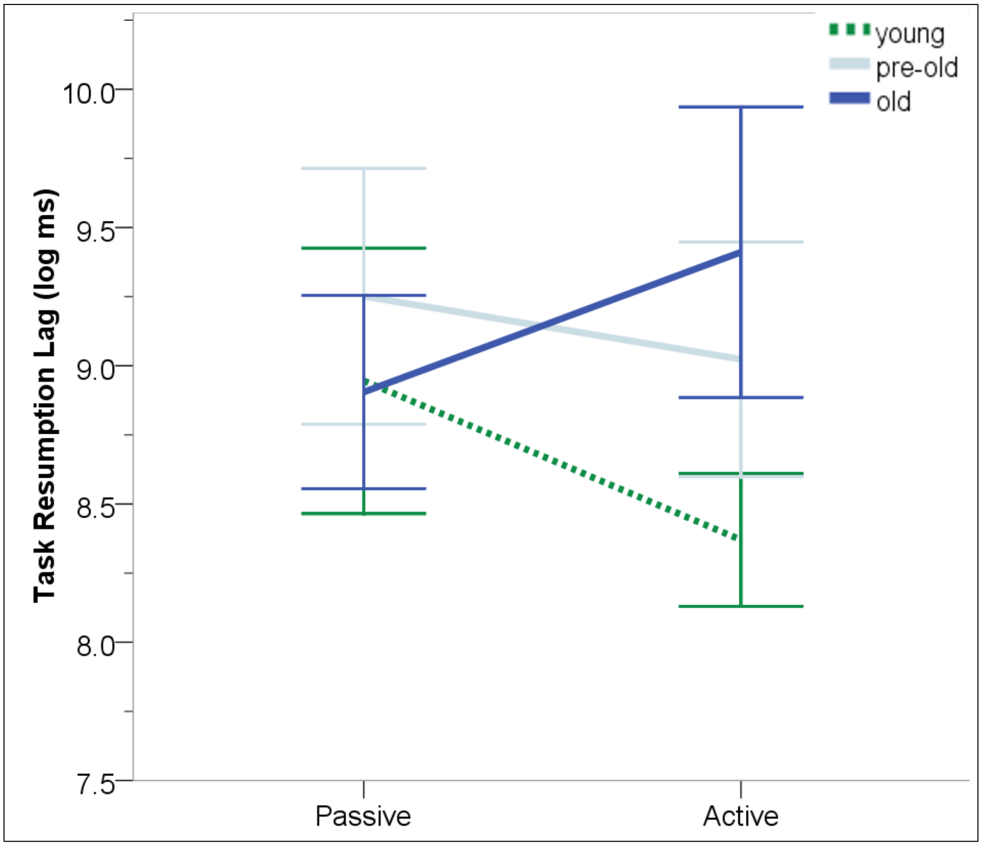


Figure 4.12: SPATIAL task resumption lag time (log ms).

In the ACTIVE condition, OLD adults are again slower than YOUNG adults ($p = .004$).

Accuracy

The main effect of age was at a trend level ($p = .078$) (Figure 4.14). YOUNG adults performed at ceiling levels. As in the VERBAL task, the observed effect of age is likely to be smaller than it would have been in the absence of a ceiling effect. Given this result, we are unable to predict if an interaction between age and interruption demand exists.

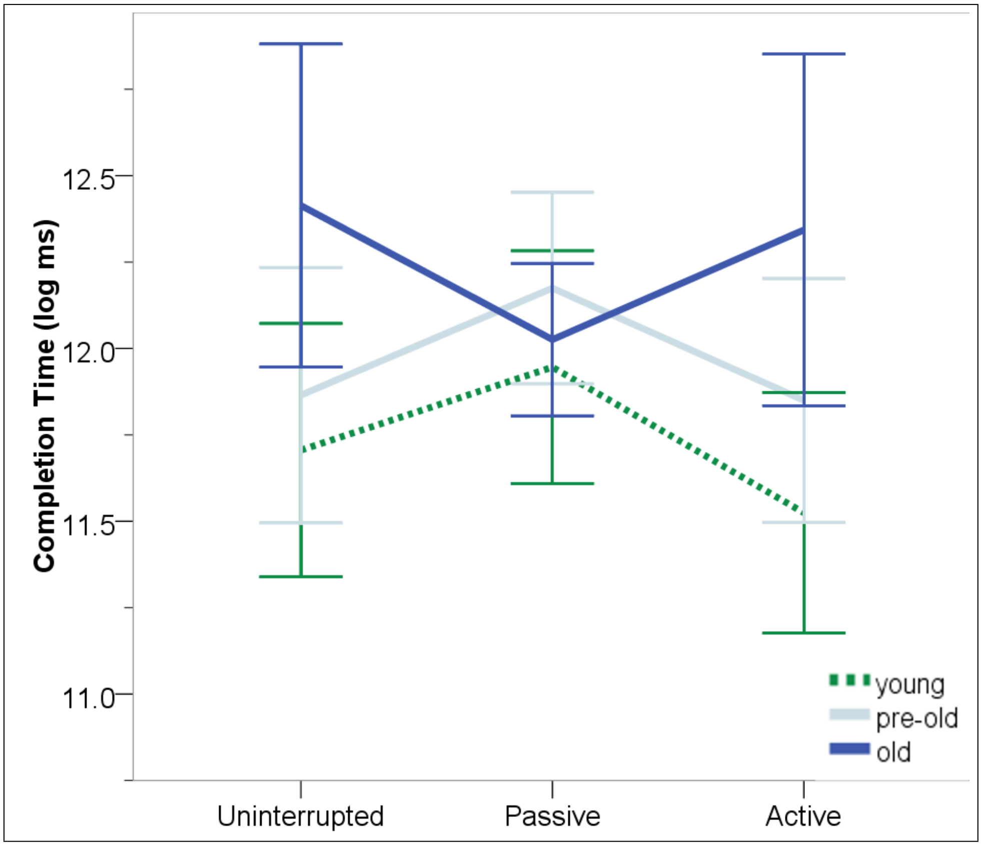


Figure 4.13: SPATIAL task completion time (log ms).

4.7.3 Between-Tasks Analysis

Differences in performance between the two primary tasks were expected. Results of omnibus 3 (age) x 3 (interruption demand) x 2 (primary task) ANOVAs with task as a within-subjects factor were thus not surprising: less time was taken to resume and complete the VERBAL task than the SPATIAL task (both $p < .001$). As expected, interactions between age, task, and interruption demand were also found. Accuracy did not differ as a factor of task or interruption demand.

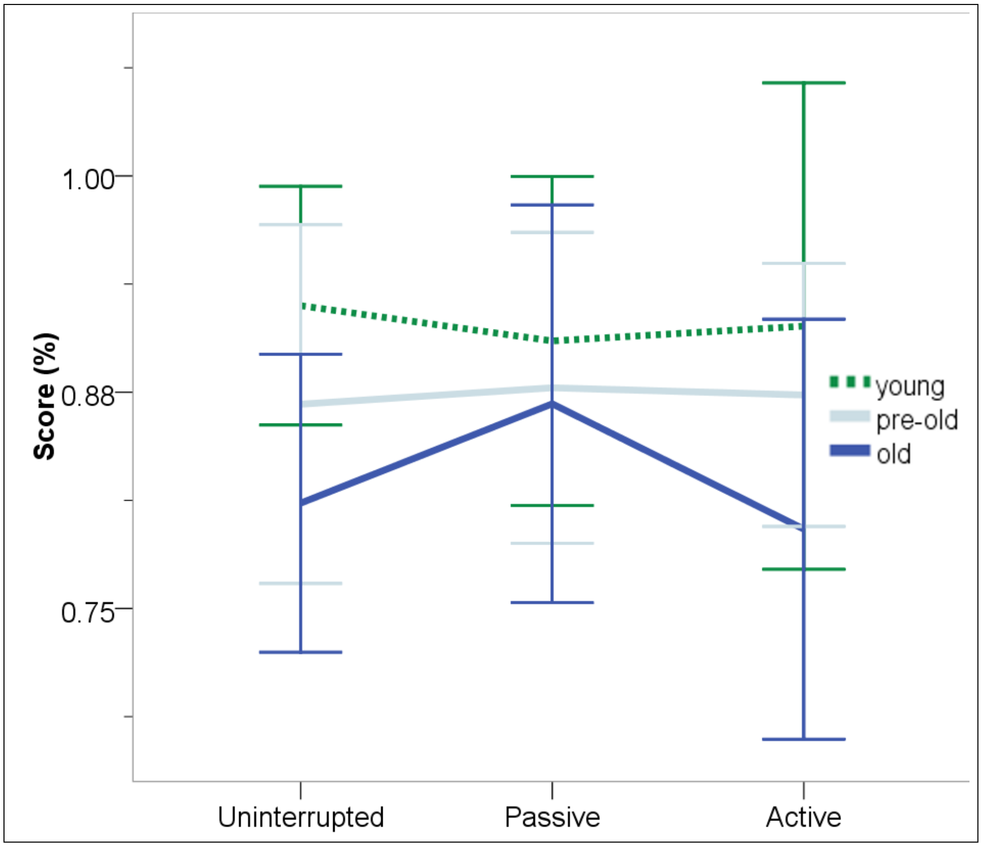


Figure 4.14: SPATIAL task accuracy (%).

4.7.4 Active Interruption Task

Mean scores on the ACTIVE task were 9.5 (YOUNG), 8.8 (PRE-OLD), and 8.2 (OLD) out of 10, indicating that participants attended to the task, as they were instructed. Nevertheless, we wanted to examine any possible main or interaction effects. Accuracy scores on the ACTIVE interruption task were negatively skewed, so we conducted a non-parametric factorial 3x2 (age x task) ANOVA using the ART technique [59] (see Section 4.7). There was a significant main effect of age on task score ($F_{2,33} = 12.08$, $p < .001$, $\eta^2 = .437$). Pairwise comparisons showed that OLD and PRE-OLD adults were less accurate than YOUNG adults ($p < .001$, $p = .008$, respectively). The latter age group performed at ceiling levels. Therefore the observed effect of age is likely to be smaller than it would have been in the absence of a ceiling effect. Given this result, we are unable to predict if an interaction between age and interruption demand exists.

The effect of task (VERBAL vs. SPATIAL) was not significant, nor was there an interaction of age and task.

4.7.5 Subjective Findings: Questionnaire Responses

Responses to questionnaire scale questions regarding primary and interrupting task difficulty and demand were analysed separately for both primary tasks. To examine repeated measures and interaction effects, we conducted non-parametric factorial 3x3 (age x interruption demand) ANOVAs, again using the ART technique [59] (see Section 4.7).

Verbal Task

OLD *adults report higher mental demand and greater annoyance.*

OLD adults reported higher levels of mental demand than YOUNG ($p =$ .050) and PRE-OLD adults ($p = .030$). Not surprisingly, the UNINTERRUPTED condition was reported to be less annoying than the PASSIVE ($p = .003$) and ACTIVE ($p < .001$) conditions. Surprisingly, the PASSIVE condition was not less annoying than the ACTIVE condition. Perceived performance did not differ between interruption conditions.

Spatial Task

Unlike the VERBAL task, mental demand and annoyance did not differ between interruption conditions in the SPATIAL task. However, OLD *adults report highest performance in the* PASSIVE *condition;* YOUNG *adults report lowest performance in the* PASSIVE *condition* (Figure 4.15).

OLD adults reported lower performance than YOUNG adults ($p = .012$). Interestingly, an interaction between age and interruption demand was at a trend level ($p = .061$).

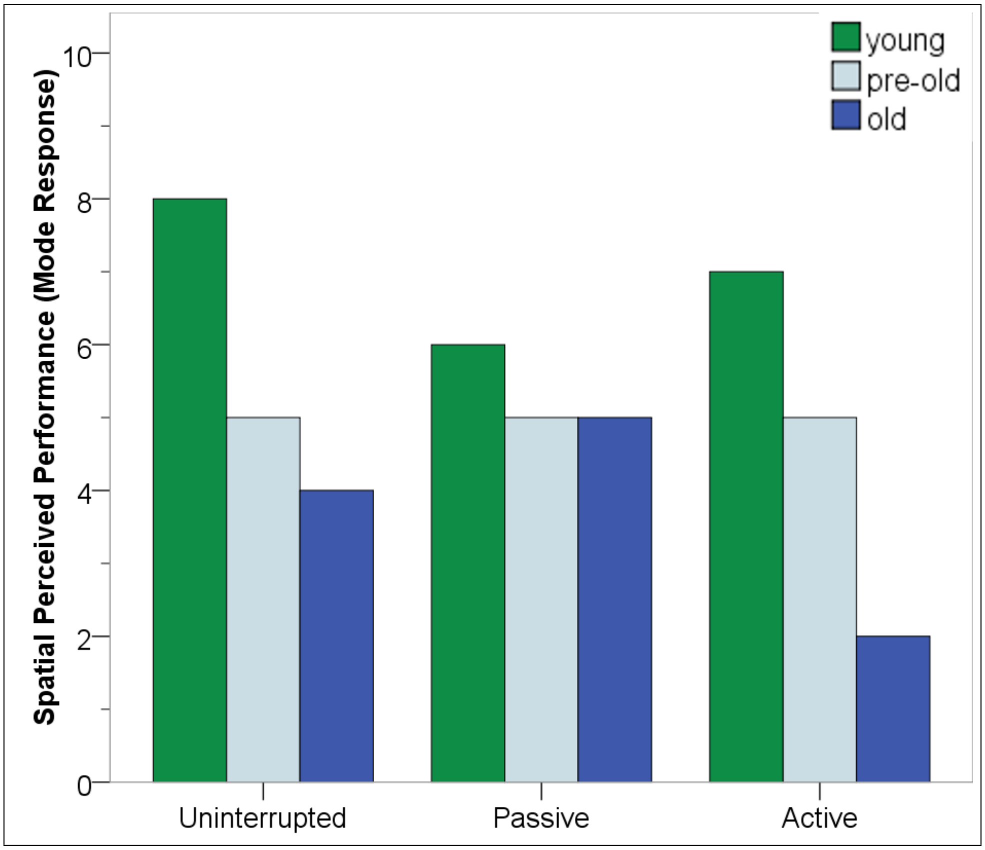


Figure 4.15: Self-reported performance on the SPATIAL task (mode response).

4.7.6 Subjective Findings: Interview Comments

The quantitative results include several divergent and trend levels results. This suggests that interruptions may be experienced differently dependent on age and primary task. We reviewed participants interview comments regarding these differences.

When asked about the relative difficulty of the two primary tasks, regardless of interruption condition, the majority of participants (7 YOUNG, 8 PRE-OLD, 8 OLD) said that the SPATIAL task was more difficult. However, a majority (9 YOUNG, 6 PRE-OLD, 12 OLD) said that the VERBAL task was disrupted to a greater extent by interruptions.

When asked about task resumption strategies, there were no clear dominant strategy across age groups. A minority of participants (5 YOUNG, 2 PRE-OLD, and 4 OLD) reported using a conscious strategy for task resumption during the 2s interruption lag. This strategy involved rehearsing key bits of information from the instruction in the VERBAL task (i.e., colours, shapes, positions, orientations), or the sequence of remaining moves in the SPATIAL task. It was unclear if these strategies changed over the course of several interruptions. Roughly half of the participants in each group reported a conscious rehearsal strategy during the PASSIVE interruption, as stated by this participant:

> *"I watched the images but I was also focusing on what I should be doing when I get back* [to the primary task], *every couple seconds reminding myself of what I should be doing when I go back"*.(OLD P4)

All participants agreed that the ACTIVE interruption was too demanding to allow continued thinking of the primary task.

Many participants, particularly those in the OLD group, did not form any task resumption strategy for the SPATIAL task. They claimed that interruptions had little or no effect on their performance, which is at odds with empirical findings.

Of those who didn't rehearse task resumption cues, 4 YOUNG participants admitted to attentively watching the images displayed during the PASSIVE interruption, and in some cases mentally practicing the ACTIVE interruption, as this was not explicitly discouraged. The remaining YOUNG participants claimed to take a break during the PASSIVE interruption, allowing their mind to wander, which may have factored into their performance, as indicated by the following:

> *"I zoned out, that was the problem. I thought not having to do anything was harder. With the* ACTIVE *interruption I was focused, I was still doing stuff. The information I had stored was still there, whereas when you zone out, things get lost"*. (YOUNG P5)

OLD participants who did not rehearse task resumption cues during the PASSIVE interruption reportedly took a break, and did not admit to practicing the ACTIVE interruption task. This lack of activity may have resulted in a feeling of impatience, as suggested by this participant:

> *"I found the* PASSIVE *interruption annoying, you were just sitting there, but it was easier. I didn't rehearse, I had done that already during the* [interruption lag]. *I am an action person so I didn't mind shifting to a new activity whereas just sitting there was annoying"*. (OLD P11)

Altogether, these comments do not decisively explain why ACTIVE and PASSIVE interruptions give rise to different performance effects for YOUNG and OLD participants, particularly in the SPATIAL task. However, we noted different reported behaviour among participants who did not rehearse primary task cues during the PASSIVE interruption. YOUNG participants were divided, some reported attending to the PASSIVE interruption while others let their mind wander. OLD participants reported ignoring it and feeling impatient.

4.7.7 Summary

We summarise our results according to our hypotheses:

H1. Age & Interruption Demand

1. *Overall,* YOUNG *adults will perform better than older (*PRE-OLD, OLD*) adults.*

   Supported. OLD adults took longer to resume and complete a primary task, and were less accurate than YOUNG adults.

2. *Older (*PRE-OLD, OLD*) adults will incur a disproportionately larger* COI *when interruption demand increases.*

   Partially supported. OLD adults took disproportionally longer than YOUNG adults to resume the VERBAL task in the ACTIVE condition,

Table 4.2: Summary of quantitative findings from the experiment; the effect of age is significant unless otherwise noted.

Task	Dependent Variable	Result
VERBAL	Resumption lag time:	OLD disproportionally. slower in ACTIVE cond.
	Completion time:	All groups slower in ACTIVE cond.
SPATIAL	Resumption lag time:	YOUNG faster in ACTIVE cond.
	Completion time:	No age difference in PASSIVE cond.
	Accuracy:	Age effect not sig.

when compared to the PASSIVE condition. Not supported by completion time or accuracy results, nor by SPATIAL task results.

H2. Age, Task & Interruption Demand

1. *Given that the* VERBAL *task places a greater load on* WM*, increased interruption demand will incur a disproportionately greater* COI *on the* VERBAL *task than on the* SPATIAL *task.*

   Partially supported. The effect of increased interruption demand was significant for the VERBAL task but not for the SPATIAL task, in terms of completion time and resumption lag time, but not accuracy.

2. *This difference in* COI *will be greater for older (*PRE-OLD*,* OLD*) adults.*

   Partially supported. OLD adults have disproportionately longer task resumption lags than YOUNG adults following an ACTIVE interruption in the VERBAL task, however completion times were not disproportionally longer.

   In the SPATIAL task, OLD and YOUNG experience the PASSIVE and ACTIVE interruptions differently. Subjective responses fail to explain this finding. Despite this, there is no significant COI as a factor of interruption demand.

4.8 Discussion

In this section, we comment on the key quantitative findings listed in Table 4.2. Qualitative findings from questionnaires and interviews are also considered.

4.8.1 Age, Interruptions, & Compensatory Behaviour

Old adults compensate for slower task resumption.

We reported a COI on a verbal memory task, where increased interruption demand incurred disproportionally longer task resumption times among old adults. This finding is also supported by cognitive ageing literature, where task switching response times have been shown to be greater for older adults [56]. Unexpectedly, however, our OLD participants were not disproportionately slower to complete the task. Therefore, after being initially slow to resume the task, old adults compensated by increasing their rate of activity (i.e., shorter time between valid actions, such as moving objects), relative to their rate of activity in the uninterrupted condition. We speculate that this behaviour is the result of an age-specific *Zeigarnik effect* [61] for older adults: a motivated effort to work with heightened efficiency after being interrupted, making up for lost time incurred by the initial COI. Alternatively, longer resumption lags may have allowed for the formulation of more efficient strategies for completing the primary task. However, this was not confirmed by interview responses.

4.8.2 Age & Primary Task Performance

Old adults have lower primary task performance relative to young adults.

The cognitive ageing literature offers an explanation. Lower processing speed [46], reduced activation in working memory [8], and compromised attentional modulation [24] may have each factored into our OLD participants performance.

4.8.3 Primary Task Accuracy & Interruptions

Primary task accuracy was not affected by interruptions, regardless of age.

This indicates that task goals, such as instructions in trials of the verbal memory task, were successfully encoded into long-term working memory before an interruption took place [43]. Thus, task goals were largely resistant to any interference caused by a demanding interruption, despite OLD adults taking longer to reactivate the suspended goal after a demanding interruption.

The absence of an effect of interruption demand on task accuracy in the spatial reasoning task was also observed. A ceiling effect was observed. While task instructions were visible throughout a trial, participants still had to remember a non-trivial amount of information: the original puzzle configuration, the number of moves completed, and the number of moves left remaining. Results indicate that this information survived interruptions.

4.8.4 Methodological Implications

Low-Demand Interruption Task

Low-demand interruptions affect age groups differently.

While there was no COI in the SPATIAL task, we observed that different age groups experienced low- (PASSIVE) and high-demand (ACTIVE) interruptions differently. In particular, young adults were faster than old adults to resume and complete the task in the active condition, but the groups did not differ in the passive condition. In the PASSIVE condition, mind-wandering may have caused young adults' performance to slip, as suggested by some interview comments. This mind-wandering afforded by the low-demand interruption may have actually had a greater negative effect on primary task performance than the high demand interruption.

It is also possible that older adults were more conscientious than young adults, and resisted mind-wandering. Cognitive testing is a sensitive topic for older adults [12], and thus a fear of poor performance may have resulted in increased conscientiousness.

We deliberately designed our low-demand interruption task to require no action, while maintaining a high visual similarity to the high-demand interruption task. We expect that a low-demand interruption task that requires a simple action (such as clicking on every image) would reduce the likelihood of mind-wandering.

Regardless of age, low-demand PASSIVE interruptions give participants a choice: they can allow their mind to wander, or they can use the opportunity to rehearse primary task cues. Interviews with participants revealed no age-specific trends with regards to their cognitive processes occurring during PASSIVE interruptions. An aim of future research will be to identify any age-related differences in strategy.

High-Demand Interruption Task

High-demand interruptions may not have been difficult enough for our young adults.

Among our YOUNG participants, there were no significant differences in task performance between the interruption conditions, with the exception that YOUNG participants task resumption lag times in the SPATIAL task were significantly shorter with high-demand ACTIVE interruptions than with low-demand PASSIVE interruptions. By contrast, Monk et al. [39] showed that increased interruption demand incurs a COI in terms of longer task resumption lag times. This suggests that our ACTIVE interruption task may not have been sufficiently demanding for our YOUNG participants.

We also observed a ceiling effect for YOUNG participants in terms of score on the ACTIVE interruption task; our OLD participants performed worse, but were still far from floor levels, and thus could have endured a more challenging task. Future research could increase n in the 'n-back' task, beyond what we used, the most-often studied '2-back' variant [44], until ceiling and floor effects are avoided for all participants.

An alternate explanation is that *the combination* of primary and interrupting tasks was not sufficiently difficult for our young adults. This would not be altogether surprising, as the C-TOC tasks are intended for older adults, and should be considerably easier for young adults.

4.8.5 Design Implications

Our primary tasks were adapted from C-TOC, a self-administered computerised cognitive assessment. The reader should be reminded that we do not expect C-TOC to be as accurate as exhaustive clinical assessments, but that it should be sufficiently accurate for the purposes of triaging and prioritising patients. As such, the current investigation was motivated by our intent to preserve the validity of C-TOC performance test results in the face of interruptions in the home. In light of our results, we address implications for preventing, detecting, and mitigating interruptions in terms of UI design.

Preventing Interruptions

Prevent interruptions with prompts tailored to each test.

A prompt to prevent external interruptions and distractions currently appears in the preparation screen displayed once at the beginning of C-TOC (Figure 4.16). Since we found divergent effects of interrupting tasks on different C-TOC tasks, it may be necessary to repeat this prompt at the outset of tasks that are particularly sensitive to the effects of interruptions, such as the VERBAL task.

Performance is currently weighted differently between completion time and accuracy for each test. These weights could be made explicit to test-takers, to increase awareness of how each test is scored and how an interruption may affect their score.

Detecting & Mitigating Interruptions

Inactivity Thresholds

Detect interruptions by requiring user response. Mitigate interruptions with trial replacement and test restarts.

Periods of inactivity and unusual variation in test performance cannot always be assumed to be caused by an interruption: an older individual may be challenged by the test and a lack of activity may represent genuine performance.

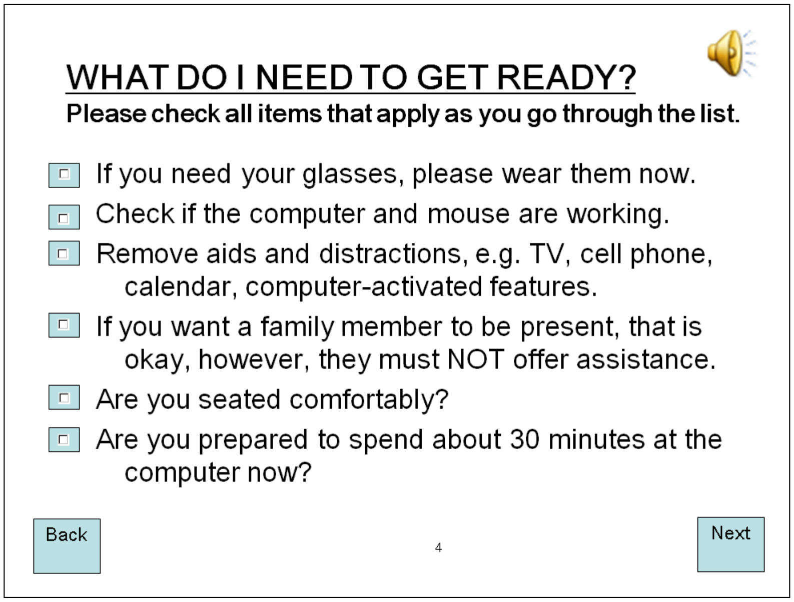


Figure 4.16: Preparation screen displayed before beginning C-TOC, which includes a prompt to prevent external interruptions and distractions.

A threshold amount of inactive time should be determined for each C-TOC test. Once this threshold is reached, a highly salient prompt should appear, querying the test-taker to determine if the current inactivity is due to an interruption; mouse movement or a key press would dismiss the prompt, allowing immediate continuation of the primary task trial. If the prompt is not quickly dismissed, C-TOC could resolve that an interruption has occurred. In this case, the interrupted primary task trial should be discarded and replaced with an isomorphic trial upon task resumption.

In cases of prolonged interruptions, a global inactive time threshold should also be determined; once passed, the test-taker would be required to restart the current test, or, if need be, the entire C-TOC battery.

Examining Rate of Activity

Detect interruptions by examining variation in rate of activity.

If the task is one in which older adults are known to compensate following task resumption (such as our VERBAL task), the rate of activity, or average time between valid actions in a task (i.e., moving objects), before and after the period of inactivity should always be compared, once the trial is completed.

When it appears likely that an interruption occurred, the performance would be classified as invalid and the user would be required to complete an isomorphic replacement trial.

Interruptions & Performance Validity

The user was interrupted. Is their performance invalid?

Effects of interruptions on primary task performance may not always incur a cost to performance, as was observed in the SPATIAL task. In these cases, trial completion time (minus inactive time) can be retained for assessment purposes.

In cases such as the VERBAL task, completion time results may no longer be valid, however accuracy results will still be reliable. The decision to retain performance data despite the occurrence of interruptions will vary from test to test and will depend greatly on how the test is scored. Alternative scoring schemes may need to be developed for interrupted tests.

Accuracy ultimately remains the most important performance criteria, in C-TOC and in existing clinical and NPT testing. Completion time is a secondary measure of performance. Given our result that accuracy remains unaffected by interruptions, C-TOC test results remain largely valid even if the user was interrupted.

General Implications

In general, segment tasks and determine inactivity thresholds.

Our findings are relevant to the design of all applications used by older adults in contexts where interruptions and distractions might occur and have potentially detrimental effects, such as online banking or booking a travel itinerary. Segmenting longer tasks into smaller sub-tasks and setting inactive time thresholds based on the task structure can limit the effects of interruptions.

Most applications will require simultaneous contribution from several cognitive processes. It is therefore important to realise that older adults will exhibit a range of behaviour when interrupted.

Chapter 5

Conclusion & Future Work

The primary goal of the work presented in this thesis was to make contributions to the design of C-TOC, a self-administered computerised cognitive assessment test. In the introductory Chapter 1, (Section 1.1), three concerns were identified pertaining to the viability of C-TOC. The research presented in this thesis addresses two of these concerns, namely whether C-TOC is usable by older adults and whether it will work in the context of an older adult's home, where interruptions and distractions are likely to occur.

In the first evaluation cycle of C-TOC, we conducted a clinical usability observation + interview study with patients of the UBC-CARD, several of whom having diagnoses of MCI or CIND. We simultaneously participated in a usability observation + focus group session with a panel of health workers who work regularly with older adults from various ethnocultural backgrounds. These investigations resulted in a list of recommendations (see Section 3.5) for improving the usability in subsequent versions of C-TOC.

We acknowledged the potential of interruptions and distractions having detrimental effects on older adults' C-TOC performance in a home context. We conducted a controlled experiment (*Study 2*, Chapter 4) to examine the effect of increased interruption demand on older adults' C-TOC test performance. Our results indicate that increased interruption demand can incur a cost to performance for older adults; however, these effects are dependent on the cognitive processes required by the primary task. Increased interruption demand did not affect task accuracy. Demanding interruptions caused longer task completion times in a verbal memory task, but this was true of both young and old participants. Overall, the results suggest that individuals are fairly robust to interruptions, even when they are demanding. This is a reassuring result with respect to the viability of C-TOC in home environments. As individuals with memory impairments are expected to struggle following interruptions, lower task accuracy will help identify them.

In summary, the contributions of this research include the significant finding of divergent effects of increased interruption demand on older adults primary task performance. Increased interruption demand can incur a cost to performance for older adults; however, these effects are dependent on the cognitive processes required by the primary task. The contributions also include implications and recommendations for improving the design of C-TOC; many of these implications are also promising for the design of other applications used by older adults.

5.1 Limitations

5.1.1 Study 1

The evaluation of C-TOC.v1 (Chapter 3) was carried out in clinical and focus group settings that are quite unlike the anticipated setting: an older adult's home.

The presence and continuous observation of experimenters in the clinical interviews, as well as the use of video recording, may have been distracting and likely contributed to a higher level of anxiety. Alternatively, continuous observation might have been helpful in keeping participants on task.

In the cross-cultural advisory panel session, the computer lab setting with more than ten other participants concurrently completing the C-TOC test battery may have induced some anxiety, and any chatter may have caused considerable distraction.

An insufficient amount of time was dedicated to discussion with the cross-cultural advisory panel after using the C-TOC prototype. In particular, there was no opportunity to discuss the usability of the prototype at length. In future meetings with the advisory panel, this must be an explicit agenda item. To facilitate this, subsequent versions of C-TOC could be web-based, and accessible from the homes or workplaces of the panel members. This will allow panel members to use and reflect upon the content and usability of the prototype in its intended setting before convening with the rest of the panel in a focus group setting. However, this would require a very robust interactive prototype.

The first evaluation cycle included only 7 usability interviewees and 5 cross-cultural advisory panel members. A greater diversity in terms of healthy and impaired older users, as well as a wider cross-cultural sampling in future evaluation cycles will add more perspectives, illuminating usability concerns that may have been overlooked in the first cycle.

The prototype was also limited, as functionality was restricted to those provided by PowerPoint. Virtual memory issues and broken hyperlinks within the PowerPoint prototype also caused some delays in completing the test battery for all participants. A robust, web-based interactive prototype will allow participants to look past low-level usability issues.

5.1.2 Study 2

As with many experiments, our examination of interruptions on older adults' C-TOC performance in Chapter 4 balanced realism and generalisability for increased precision.

This is particularly true of the interrupting tasks selected for the experiment, which will differ from naturalistic interruptions experienced in the homes of older adults, in terms of the amount of mind-wandering afforded and working memory demand imposed. These interrupting tasks were chosen on the basis that they represent a range of possible naturalistic interruptions in terms of the WM demands they impose. We also acknowledge that only our high-demand interrupting task required a motor response, a mouse click. We could have included a simple motor response in the low-demand interrupting task, or we could have designed the high-demand interrupting task to require a non-motor response (i.e., a verbal response). Future work should assess these variants of interrupting tasks through extensive piloting.

Our younger participants performed at ceiling levels, in terms of accuracy, on both primary tasks and on the high-demand ACTIVE interrupting task. Therefore we are unable to predict the true magnitude of the age effects we reported. In the case of both primary tasks, we are unable to predict if an interaction effect between age and interruption demand exists. Additional piloting to calibrate the difficulty of these tasks should be conducted in future work in order to avoid ceiling effects.

We acknowledge the differences in level of education and computer literacy between the age groups, which may have contributed to our results. We also recognise that some of our older participants were recruited using online classified advertisements and email, potentially indicative of a higher level of computer literacy than many of their peers.

The experiment involved 2 C-TOC tests and did not explicitly address the effect of interruptions on the remaining 13 tests. Due to the diversity of these tests we cannot reliably generalise our current findings to all of these tests.

Finally, we have yet to manipulate other factors found to be predictive of the COI (see Section 2.2.1). These factors include similarity between interrupting and primary tasks, interruption duration and frequency, the length and visibility of the interruption lag.

Given these limitations, this research should be regarded as a first step towards understanding the interactions of age, task, and interruption workload demand.

5.2 Future Work

5.2.1 C-TOC Development & Evaluation

There is incentive for future versions of C-TOC to be accessible over the web, allowing the cultural advisory panel and representative clinical users to interact with the prototype from their homes or workplaces. This will allow more time for discussion at future meetings with the panel and representative users.

While the evaluation of C-TOC.v1 was conducted with representative users in clinical settings, the involvement of focus groups of representative older users and their family caregivers should be considered in future evaluations, taking into account design and evaluation considerations put forward by prior research [14, 15, 34, 36], described in Section 2.1.3.

This first evaluation cycle also did not include a panel of HCI experts to evaluate the prototype; future evaluation should enlist such a panel to triangulate on a consistent and satisfactory design.

More broadly, future development of the C-TOC test battery will incorporate several related interaction design issues:

- Presenting task instructions effectively.
- Keeping the the test-taker engaged and motivated.
- Preventing and detecting cheating.
- Addressing privacy concerns.
- Designing for different levels of computer literacy.
- Designing for different cultural backgrounds.
- Given what was learned in our laboratory experiment (Chapter 4), preventing and detecting task interruption, mitigating the effects of interruptions and distractions, and promoting effective task resumption are logical directions for future development.
- Effectively prompting and offering assistance when a user 'gets stuck'. Related work has investigated intelligent task prompting and context-aware assistance for users with cognitive impairments, especially for activities of daily living, such as feeding and hygiene; Mihailidis et al. [38] have developed one such system to help those with moderate dementia wash their hands. To our knowledge, intelligent prompting

either for healthy older adults and mildly-impaired individuals, or for the domain of computerised testing, has yet to be established.

5.2.2 C-TOC, Interruptions & Older Adults

Our future work includes further controlled experiments to investigate interactions of age and interruption demand on other C-TOC tests (Table 3.1) that engage cognitive processes other than language comprehension / verbal WM and problem solving / spatial WM.

Additional factors known to be predictive of the COI will also be examined, including interruption frequency and modality.

Determining appropriately demanding and externally valid levels of interruption demand will require further exploration and piloting.

Follow-up studies will also include an examination of the effects of interruption on older adults experiencing cognitive decline (i.e., those diagnosed as having MCI/CIND), compared to healthy older adults. Effects of different levels of computer literacy should also be explicitly analysed.

We also seek a deeper qualitative understanding of age differences in strategy, with respect to different primary tasks and levels of interruption demand.

Finally, we will evaluate methods for preventing, detecting, and mitigating interruptions, as well as the effects of these methods on test validity.

5.3 Concluding Remarks

C-TOC represents the first step in the process of making diagnostic services more readily available, providing clinicians with initial screening results for prioritising patients, so that individuals exhibiting pathological cognitive decline can be diagnosed and treated sooner.

It is our hope that the research documented in this thesis will contribute to the development of a usable and valid cognitive assessment tool. This includes the early identification of usability concerns, from low-level interaction design issues to preventing and detecting interruptions occurring in the home, and subsequently mitigating their effects on performance to retain test validity.

Future work will address other issues surrounding C-TOC in home settings. These include varying levels of computer literacy among users, methods for motivating test-takers, privacy concerns, and the possibility of cheating.

In the years to come, more older adults will experience cognitive decline and present concerns regarding cognitive health. Tools such as C-TOC will make a difference for these individuals, and will aid in the process of prioritising care for those who need it most.

Bibliography

[1] P. D. Adamczyk and B. P. Bailey. If not now, when?: The effects of interruption at different moments within task execution. In *Proc. CHI '04*, pages 271–278, 2004. → pages 11

[2] E. M. Altmann and J. G. Trafton. Task interruption: Resumption lag and the role of cues. In *Proc. 26th Annual Conference of the Cognitive Science Society*, pages 43–48, 2004. → pages 4, 11, 14, 18, 19, 49

[3] A. Baddeley and G. Hitch. Working Memory. *The Psychology of Learning and Motivation: Advances in Research and Theory*, 8:47–89, 1974. → pages 43

[4] B. P. Bailey, J. A. Konstan, and J. V. Carlis. Measuring the effects of interruptions on task performance in the user interface. In *Proc. Systems, Man and Cybernetics '00*, pages 757–762, 2000. → pages 4, 12, 13, 15, 19

[5] R. Cabeza. Task-independent and task-specific age effects on brain activity during working memory, visual attention and episodic retrieval. *Cerebral Cortex*, 14(4):364–375, 2004. → pages 16, 17

[6] W. C. Clapp and A. Gazzaley. Distinct mechanisms for the impact of distraction and interruption on working memory in aging. *Neurobiology of Aging*, pages 1–15, 2010. → pages 4, 16, 17

[7] J. Cohen. Eta-squared and partial eta-squared in communication science. *Human Communication Research*, 28:473–490, 1973. → pages 56

[8] F. I. M. Craik and M. Byrd. *Aging and cognitive deficits: The role of attentional resources*, pages 191–211. In Craik, F. I. M. and Trehub, S. E. (eds.), *Aging and Cognitive Processes.* Plenum Press, 1982. → pages 16, 17, 69

[9] F. I. M. Craik and T. A. Salthouse, editors. *The Handbook of Human Aging and Cognition.* Erlbaum, Hillsdale, NJ, 2nd edition, 1992. → pages 54

[10] E. Cutrell, M. Czerwinski, and E. Horvitz. Notification, disruption, and memory: Effects of messaging interruptions on memory and performance. *Human Factors*, pages 263–269, 2001. → pages 15

[11] E. De Renzi and L. A. Vignolo. The token test: A sensitive test to detect receptive disturbances in aphasics. *Brain : A Journal of Neurology*, 85:665–678, 1962. → pages 52

[12] A. Dickinson, J. Arnott, and S. Prior. Methods for human - computer interaction research with older people. *Behaviour & Information Technology*, 26(4):343–352, 2007. → pages 7, 70

[13] B. Dubois, H. H. Feldman, C. Jacova, S. T. DeKosky, P. Barberger-Gateau, J. Cummings, A. Delacourte, D. Galasko, S. Gauthier, and G. Jicha. Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *The Lancet Neurology*, 6(8):734–746, 2007. → pages 8, 23

[14] R. Eisma, A. Dickinson, J. Goodman, O. Mival, A. Syme, and L. Tiwari. Mutual inspiration in the development of new technology for older people. In *Proc. of Include '03*, 2003. → pages 9, 80

[15] R. Eisma, A. Dickinson, J. Goodman, A. Syme, L. Tiwari, and A. F. Newell. Early user involvement in the development of information technology-related products for older people. *Universal Access in the Information Society*, 3(2):131–140, 2004. → pages 9, 80

[16] S. Farrimond, R. G. Knight, and N. Titov. The effects of aging on remembering intentions: performance on a simulated shopping task. *Applied Cognitive Psychology*, 20(4):533–555, 2006. → pages 4, 16, 17, 18, 19

[17] H. H. Feldman and C. Jacova. Mild cognitive impairment. *American Journal of Geriatric Psychiatry*, 13(8):645–655, 2005. → pages 8, 23

[18] H. H. Feldman, C. Jacova, A. Robillard, A. Gracia, T. Chow, M. Borrie, H. M. Schipper, M. Blair, A. Kertesz, and H. Chertkow. Diagnosis and treatment of dementia: 2. Diagnosis. *Canadian Medical Association Journal*, 178(7):825–836, 2008. → pages 8, 23

[19] J. Fogarty, S. E. Hudson, C. G. Atkeson, D. Avrahami, J. Forlizzi, S. Kiesler, J. C. Lee, and J. Yang. Predicting human interruptibility with sensors. In *Proc. CHI '05*, pages 257–264, 2005. → pages 16

[20] M. F. Folstein, S. E. Folstein, and P. R. McHugh. Mini-Mental State: a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12:189–198, 1975. → pages xiv, 1

[21] T. Gillie and D. Broadbent. What makes interruptions disruptive? A study of length, similarity, and complexity. *Psychological Research*, 50 (4):243–250, 1989. → pages 4, 11, 13, 18

[22] J. Goodman, A. Dickinson, and A. Syme. Gathering requirements for mobile devices using focus groups with older people. In *Designing a More Inclusive World, Proc. 2nd Cambridge Workshop on Universal Access and Assistive Technology*, pages 1–10, 2004. → pages 24, 109

[23] S. Hart and L. Staveland. *Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research*, pages 139–183. In Hancock, P. and Meshkati, N. (eds.). *Advances in Human Psychology: Human Mental Workload*. Elsevier Science, 1988. → pages xv, 53, 161

[24] L. Hasher and R. T. Zacks. Working memory, comprehension, and aging: A review and a new view. *Psychology of Learning and Motivation*, 22:193–225, 1988. → pages 4, 16, 17, 69

[25] K. Hawkey, K. M. Inkpen, K. Rockwood, M. McAllister, and J. Slonim. Requirements gathering with Alzheimer's patients and caregivers. *Proc. Assets '05*, pages 142–149, 2005. → pages 9

[26] H. M. Hodgetts and D. M. Jones. Contextual cues aid recovery from interruption: the role of associative activation. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 32(5): 1120–1132, 2006. → pages 15, 18

[27] E. Horvitz and J. Apacible. Learning and reasoning about interruption. *Proc. ICMI '03*, pages 20–27, 2003. → pages 16

[28] S. E. Hudson, J. Fogarty, and C. Atkeson. Predicting human interruptibility with sensors: a Wizard of Oz feasibility study. In *Proc. CHI '03*, pages 257–264, 2003. → pages 16

[29] K. M. Inkpen. Drag-and-drop versus point-and-click mouse interaction styles for children. *ACM Transactions on Computer-Human Interaction*, 8(1):1–33, 2001. → pages 90

[30] S. T. Iqbal and B. P. Bailey. Leveraging characteristics of task structure to predict the cost of interruption. *Proc. CHI '06*, pages 741–750, 2006. → pages 15, 18, 19

[31] S. T. Iqbal and E. Horvitz. Disruption and recovery of computing tasks: Field study, analysis, and directions. *Proc. CHI '07*, pages 677–686, 2007. → pages 11, 16, 49

[32] C. Jacova. Exploring and validating the contributions of C-TOC (Cognitive Testing on Computer): Use in clinical assessment. Research Proposal, 2011. → pages 1, 2, 3, 6, 7

[33] C. Jacova, H. S. Lee, S. Le Huray, J. McGrenere, B. L. Beattie, H. Feldman, and G.-Y. Robin Hsiung. Cognitive Testing on Computer (C-TOC): Designing a computerized test battery for evaluation of cognitive impairment with user and community health professional input. *Alzheimer's and Dementia*, 6(4):S317–S318, 2010. → pages xiv, 1

[34] S. H. Kurniawan and P. Zaphiris. Research-derived web design guidelines for older people. In *Proc. Assets '05*, pages 129–135, 2005. → pages 10, 80

[35] M. L. Lee and A. K. Dey. Providing good memory cues for people with episodic memory impairment. *Proc. Assets '07*, pages 131–138, 2007. → pages 9

[36] M. Massimi, R. M. Baecker, and M. Wu. Using participatory activities with seniors to critique, build, and evaluate mobile phones. In *Proc. Assets '07*, pages 155–162, 2007. → pages 9, 80

[37] D. McFarlane and K. Latorella. The ccope and importance of human interruption in human-computer interaction design. *Human-Computer Interaction*, 17:1–61, 2002. → pages 18

[38] A. Mihailidis, J. Boger, M. Canido, and J. Hoey. The use of an intelligent prompting system for people with dementia. *Interactions*, pages 34–37, 2007. → pages 80

[39] C. A. Monk, J. G. Trafton, and D. A. Boehm-Davis. The effect of interruption duration and demand on resuming suspended goals. *Journal of Experimental Psychology. Applied*, 14(4):299–313, 2008. → pages 11, 14, 18, 19, 47, 71

[40] Z. S. Nasreddine, N. A. Phillips, V. Bédirian, S. Charbonneau, V. Whitehead, I. Collin, J. L. Cummings, and H. Chertkow. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53 (4):695–699, 2005. → pages xv, 1, 7, 55, 129

[41] B. O'Conaill and D. Frohlich. Timespace in the workplace: Dealing with interruptions. In *Proc. CHI '95*, pages 262–263, 1995. → pages 4, 10

[42] Ofcom. Media literacy audit - digital lifestyles: Adults aged 60 and over. Technical report, Ofcom, London, UK, 2009. → pages 6

[43] A. Oulasvirta and P. Saariluoma. Surviving task interruptions: Investigating the implications of long-term working memory theory. *International Journal of Human-Computer Studies*, 64(10):941–961, 2006. → pages 11, 12, 14, 19, 70

[44] A. M. Owen, K. M. McMillan, A. R. Laird, and E. Bullmore. N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, 25(1):46–59, 2005. → pages 12, 47, 71

[45] C. Parnin and R. DeLine. Evaluating cues for resuming interrupted programming tasks. *Proc. CHI '10*, pages 93–102, 2010. → pages 15, 18

[46] T. A. Salthouse. The processing-speed theory of adult age differences in cognition. *Psychological Review*, 103(3):403–28, 1996. → pages 16, 17, 69

[47] D. D. Salvucci. On reconstruction of task context after interruption. *Proc. CHI '10*, pages 89–92, 2010. → pages 11, 13, 14

[48] N. Selwyn, S. Gorard, J. Furlong, and L. Madden. Older adults' use of information and communications technology in everyday life. *Ageing and Society*, 23(5):561–582, 2003. → pages 6

[49] J. Sevilla, G. Herrera, B. Martínez, and F. Alcantud. Web accessibility for individuals with cognitive deficits. *ACM Transactions on Computer-Human Interaction*, 14(3):12.1–25, 2007. → pages 8

[50] C. Speier, I. Vessey, and J. S. Valacich. The Effects of Interruptions, Task Complexity, and Information Presentation on Computer-Supported Decision-Making Performance. *Decision Sciences*, 34(4):771–797, 2003. → pages 4, 11, 13, 18, 19

[51] N. A. Storch. Does the user interface make interruptions disruptive? A study of interface style and form of interruption. Technical report, Lawrence Livermore National Laboratory, Springfield, 1992. → pages 11, 16, 49

[52] G. J. Trafton. Preparing to resume an interrupted task: effects of prospective goal encoding and retrospective rehearsal. *International Journal of Human-Computer Studies*, 58(5):583–603, 2003. → pages 11, 14, 18, 19, 49

[53] B. Uttl. North American Adult Reading Test: age norms, reliability, and validity. *Journal of Clinical and Experimental Neuropsychology*, 24(8):1123–1137, 2002. → pages xv, 55, 129, 135

[54] B. Uttl. Transparent meta-analysis of prospective memory and aging. *PloS ONE*, 3(2):e1568 1–31, 2008. → pages 16, 17

[55] G. Ward and A. Allport. Planning and problem-solving using the five-disc tower of london task. *Quartlery Journal of Experimental Psychology*, 50A:49–78, 1997. → pages 15, 18

[56] C. Wasylyshyn, P. Verhaeghen, and M. J. Sliwinski. Aging and task switching: a meta-analysis. *Psychology and Aging*, 26(1):15–20, 2011. → pages 4, 16, 17, 69

[57] R. L. West. An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, 120(2):272–292, 1996. → pages 16, 17

[58] J. Wherton and A. Monk. Technological opportunities for supporting people with dementia who are living at home. *International Journal of Human-Computer Studies*, 66(8):571–586, 2008. → pages 8

[59] J. O. Wobbrock, L. Findlater, D. Gergle, and J. J. Higgins. The Aligned Rank Transform for nonparametric factorial analyses using only ANOVA procedures. In *Proc. CHI '11*, pages 143–146, 2011. → pages xiv, 56, 63, 64

[60] M. Wu, J. Birnholtz, B. Richards, R. M. Baecker, and M. Massimi. Collaborating to remember: A distributed cognition account of families coping with memory impairments. *Memory*, pages 825–834, 2008. → pages 9

[61] B. Zeigarnik. Das behalten erledigter und unerledigter handlungen. *Psychologische Forschung*, 9:1–85, 1927. → pages 12, 69

[62] S. Zhai, J. Kong, and X. Ren. Speed-accuracy trade-off in Fitts law tasks On the equivalency of actual and nominal pointing precision. *International Journal of Human-Computer Studies*, 61(6):823–856, 2004. → pages 55

[63] F. R. H. Zijlstra, R. A. Roe, A. B. Leonora, and I. Krediet. Temporal factors in mental work: Effects of interrupted activities. *Journal of Occupational and Organizational Psychology*, 72(2):163–185, 1999. → pages 10, 12, 14, 18

Appendix A

An Overview of C-TOC.v1

This appendix contains a description of C-TOC.v1[1].

C-TOC.v1 is a pilot version of the test battery implemented in a series of interactive PowerPoint files. It can be run from a PC running Microsoft Windows with Microsoft Office 2007 installed.

Complete functionality of the interactive elements of the C-TOC prototype require that PowerPoint macros are enabled (administrator privileges are required to change PowerPoint security settings and enable all macros). A single macro has been used throughout the prototype to facilitate the movement of interactive objects in full-screen slide mode, analogous to traditional drag-and-drop functionality in many desktop operating systems, however without requiring continuous depression of the left mouse button. Instead, all objects are moved by clicking once on an object, moving the object along with the cursor, and clicking again to leave the object in place. This point-and-click interaction style has not been compared to drag-and-drop among older adults, however a study by Inkpen [29] with children revealed that this point-and-click style of interaction was more effective in terms of speed and error rate.

[1] C-TOC.v1, evaluated in Chapter 3, was developed in Microsoft PowerPoint 2007 by Hyunsoo (Steve) Lee under the direction of Dr. Claudia Jacova. For copyright information, see C. Jacova, G.Y. Hsiung, H. Lee, J. & McGrenere. Cognitive Testing on Computer (C-TOC). Invention Disclosure UBC University Industry Liaison Office file no. 11–123, 2011–07–19

Each test is preceded by one or two practice trials, in which feedback is given to the user after each trial. Should the user respond to a practice item incorrectly, the practice trial is repeated until a correct response is made. Aside from the first test (PICTURE-WORD PAIRS, Section A.2), no feedback is given to a user after any other non-practice trial response.

The following subsections describe each component of the C-TOC-V.1 test battery in detail.

A.1 C-TOC Introduction

C-TOC begins by preparing the user to undergo the test battery. This process includes an explanation of the features of a computer mouse (Figure A.2), a statement explaining what is expected from the test-taker (Figure A.3), a preparation checklist serving to eliminate potential distractions (Figure A.5), an introduction to the C-TOC help menu (Figure A.6). Then personal information is gathered (Figure A.4), requiring the user to make selections from buttons and drop-down menus. Finally, practice using the mouse is performed; users must click on target objects, and in one case, move a circle shape into a square target area (Figure A.7).

A.2 The Picture-Word Pairs Test

PICTURE-WORD PAIRS is a memory encoding test, and presents users with 4 images and an instruction to click on one of the objects given a category name, as shown in Figure A.10. Once an item is clicked, its specific name is shown. After clicking again, feedback is presented to the user (correct Figure A.8 or incorrect Figure A.9). If the user has made an incorrect solution, the trial is repeated until the correct selection is made; this is necessary as the correct response is required in the recall test (Section A.3).

A.3 The Word Recognition Test

The WORD RECOGNITION is the recall counterpart of the PICTURE-WORD PAIRS encoding test (Section A.2). Each trial asks the user about the specific

Figure A.1: The opening screen of C-TOC.

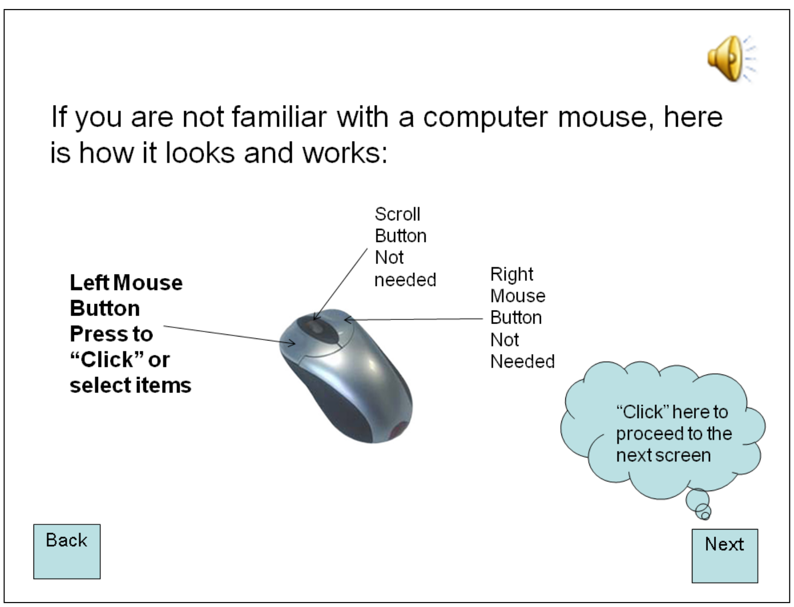


Figure A.2: The C-TOC Introduction: instructions for using the mouse.

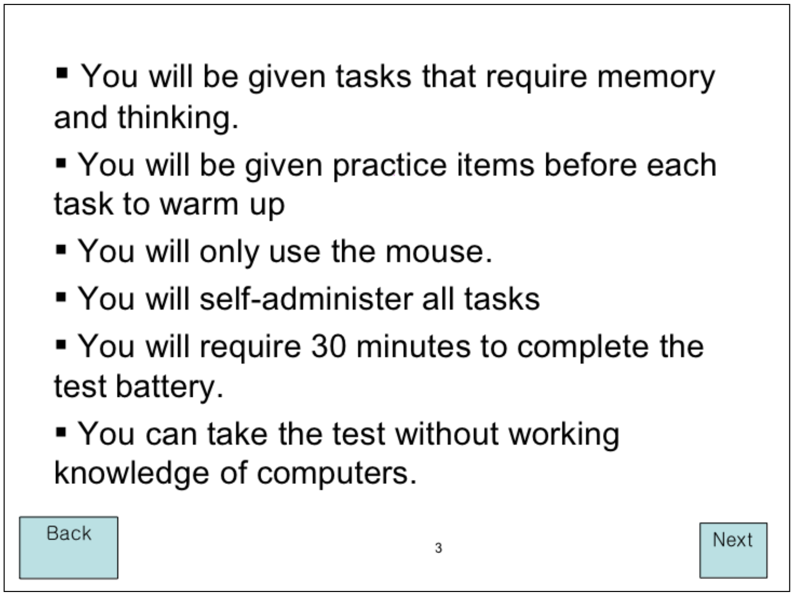


Figure A.3: The C-TOC Introduction: a briefing of what to expect from C-TOC.

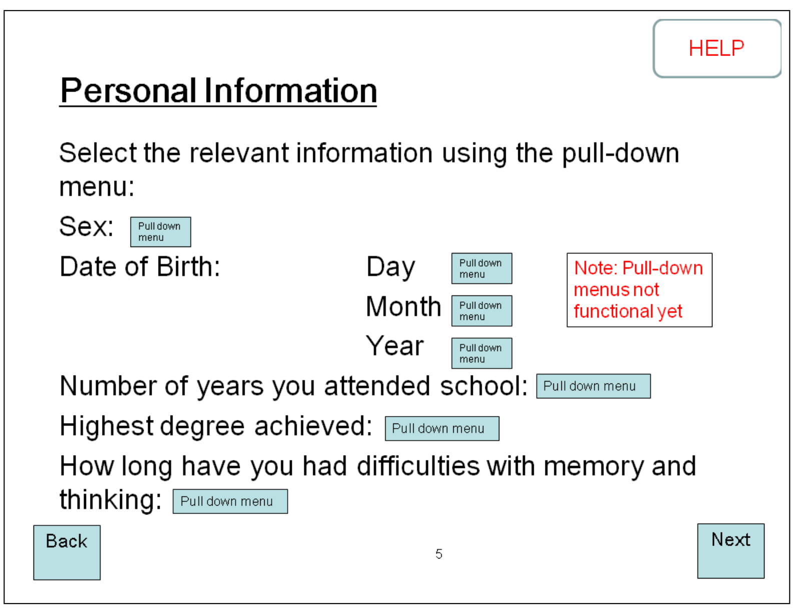


Figure A.4: Collecting background information from the user.

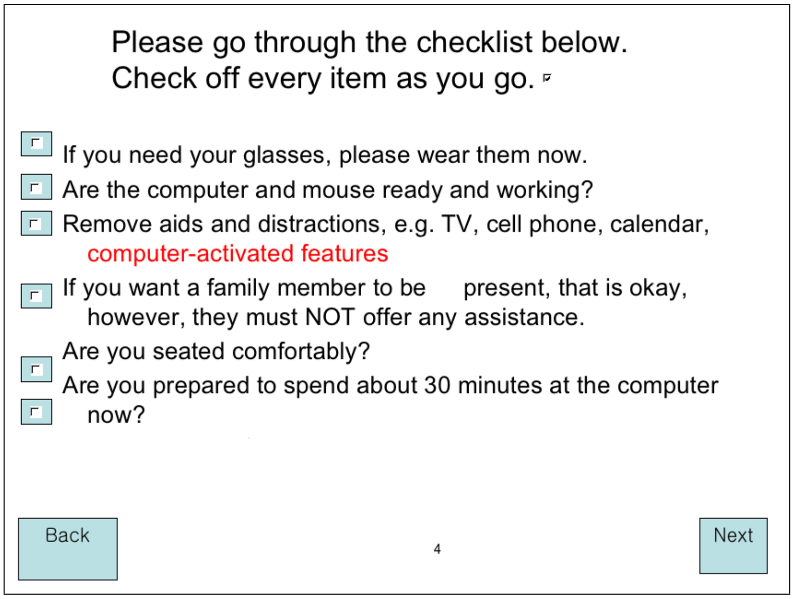


Figure A.5: Preparation screen displayed before beginning C-TOC, which includes a prompt to prevent external interruptions and distractions

items seen during the PICTURE-WORD PAIRS test. Each trial presents the user with a prompt to select a specific item name from a list of four choices given a category name, as shown in Figure A.11. A clickable button is placed beside each choice. Feedback is not given between trials , and each trial may only be attempted once.

A.4 The Temporal Orientation Test

The TEMPORAL ORIENTATION test requires users to report the current date by selections from lists of seasons, months, years, and days of the month, as shown in Figure A.12.

A.5 The Symbol-Digit Matching Test

In the SYMBOL-DIGIT MATCHING test, the user must match the target symbol at the centre of the screen to one of the symbols at the top of the screen

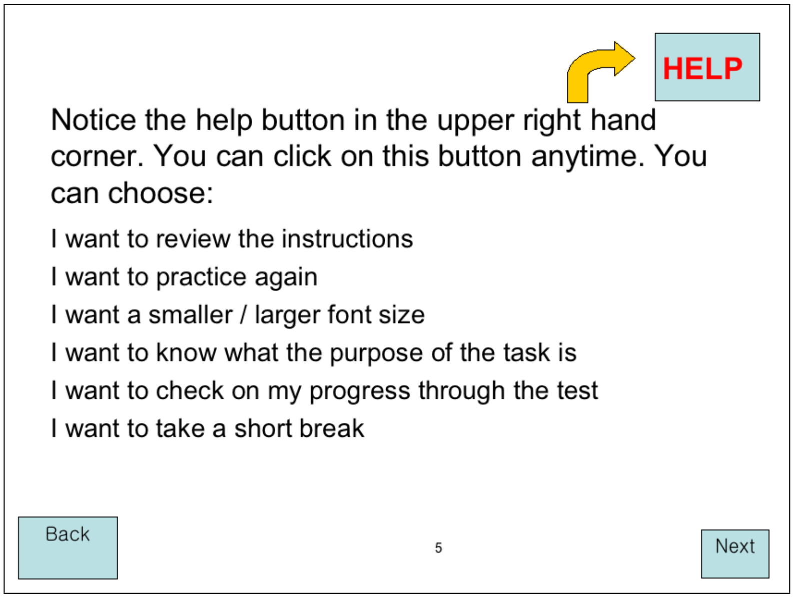


Figure A.6: Contents of the C-TOC help menu. In future versions of C-TOC, the help menu will be accessible from any location in the test battery, and will be contextually aware with regard to the current test.

by clicking on the corresponding number at the bottom of the screen (as in Figure A.13).

A.6 The Similarities Test

In the SIMILARITIES test, the user is verbally prompted with two items; this is accompanied by a list of three ways in which these two items are similar, ranging from a general high-level similarity to a precise similarity; the user must select the answer that best captures their similarity by clicking on the labeled button beside each list item (as in Figure A.14).

A.7 The Pattern Construction Test

In the PATTERN CONSTRUCTION test, the user must reconstruct a target pattern with a set of movable shapes, using click and drag (as in Figure A.15).

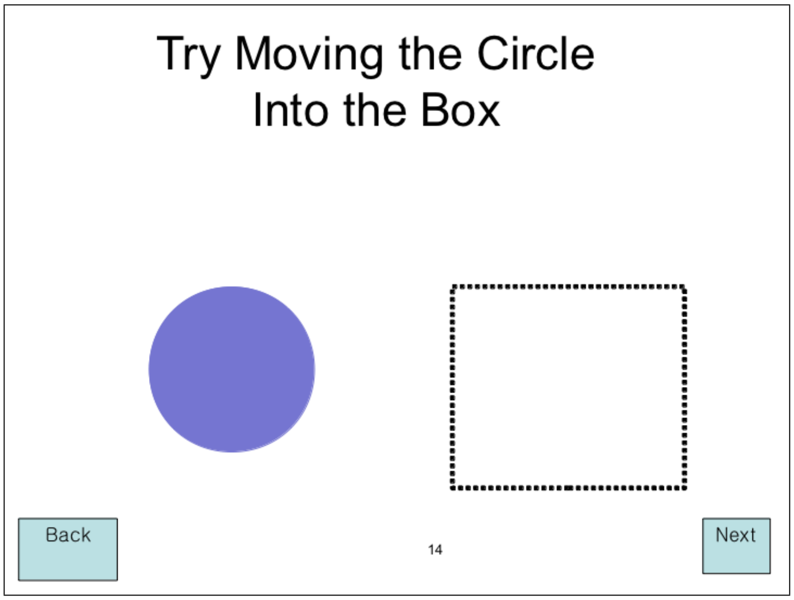


Figure A.7: Practice using the mouse (i.e., selecting and moving objects)

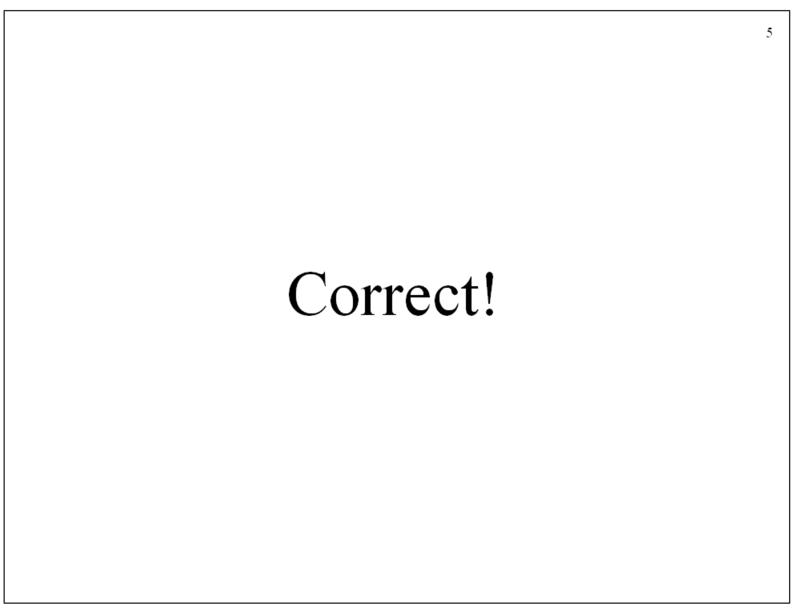


Figure A.8: Correct feedback in C-TOC.

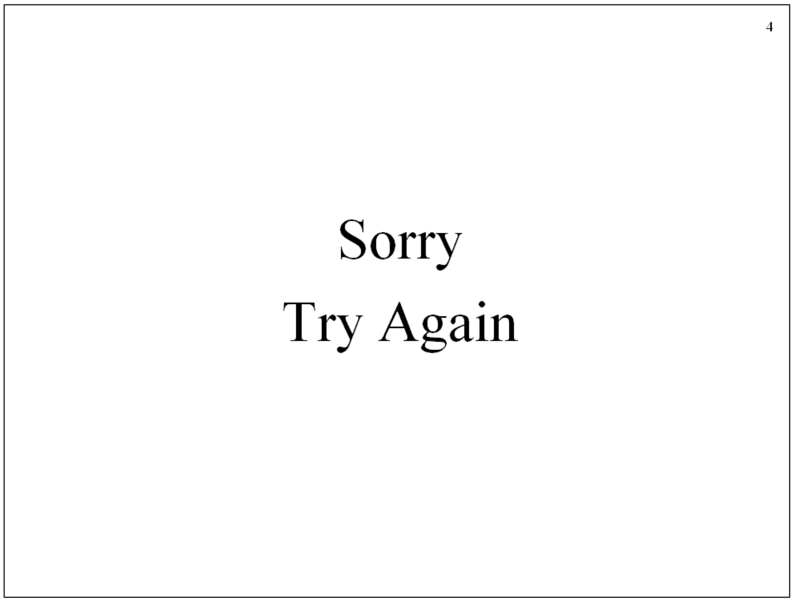


Figure A.9: Incorrect feedback in C-TOC.

Figure A.10: The PICTURE-WORD PAIRS test.

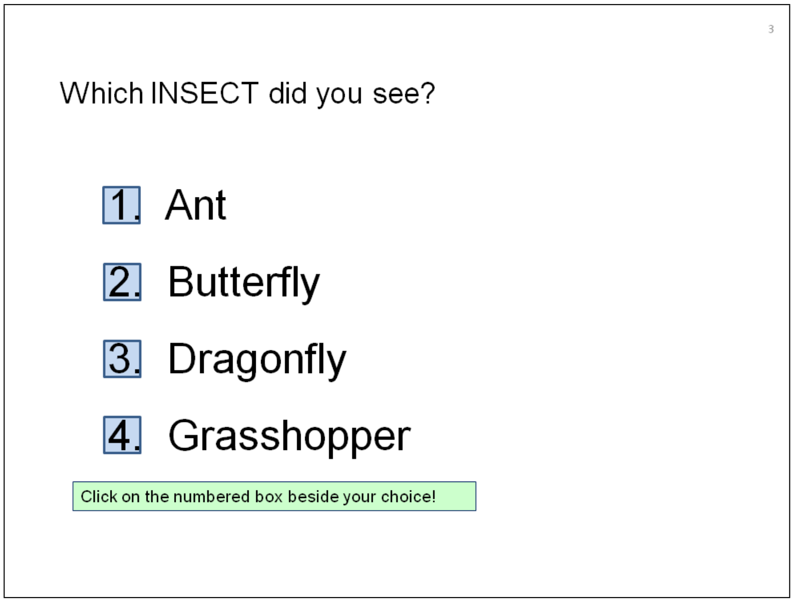


Figure A.11: The WORD RECOGNITION test.

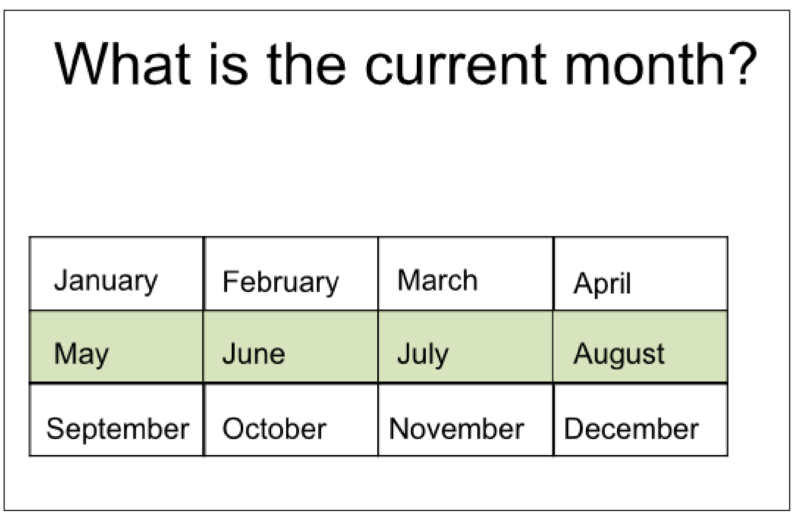


Figure A.12: The TEMPORAL ORIENTATION test.

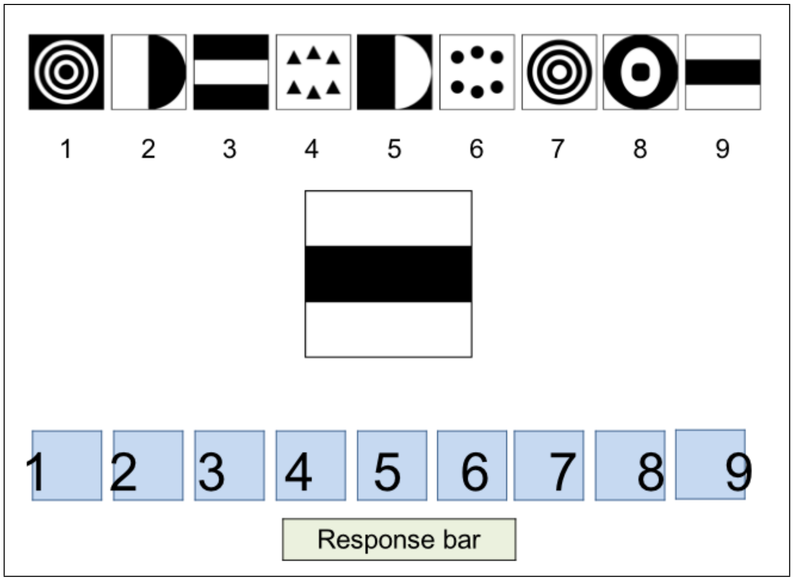


Figure A.13: The SYMBOL-DIGIT MATCHING test.

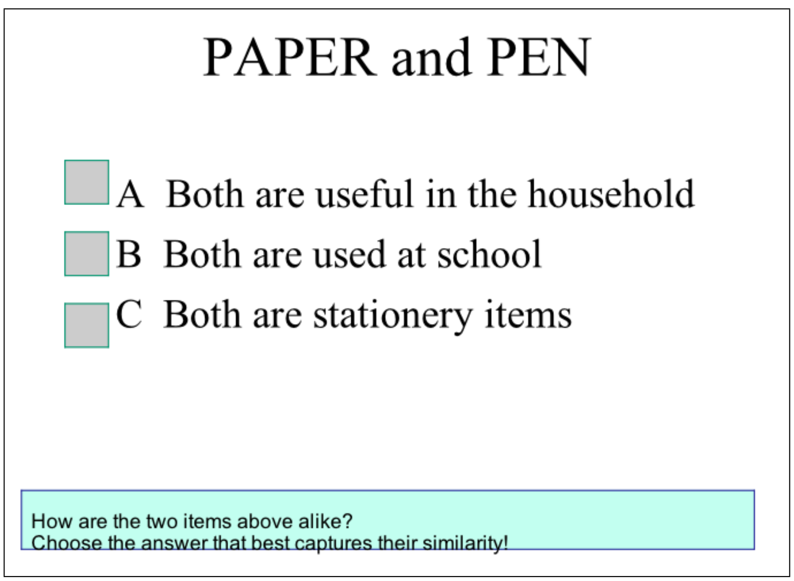


Figure A.14: The SIMILARITIES test.

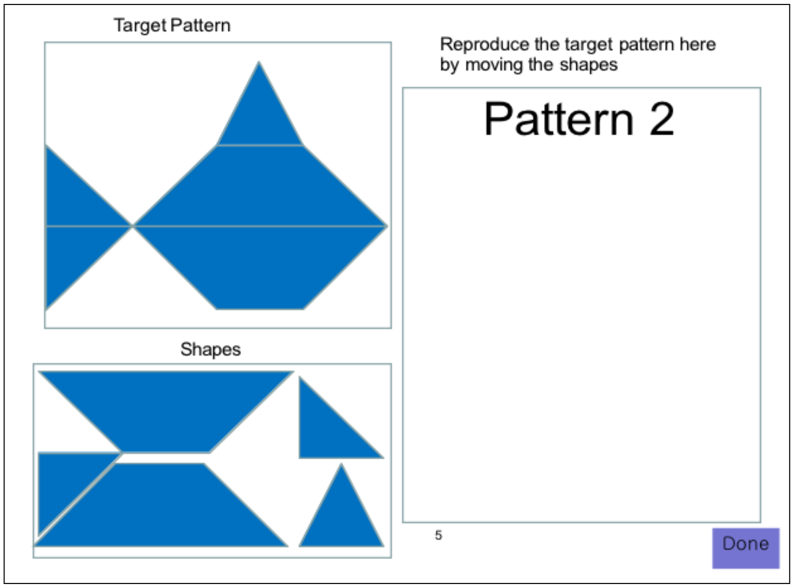


Figure A.15: The PATTERN CONSTRUCTION test.

These shapes can be moved but not rotated.

A.8 The Pattern Recall Test

PATTERN RECALL is the recall counterpart of the PATTERN CONSTRUCTION test (Section A.7). For each trial encountered in the PATTERN CONSTRUCTION test, the user must reconstruct each target pattern from memory with a set of movable shapes. As in the PATTERN CONSTRUCTION test, these shapes can be moved but not rotated. The screen layout for this test is similar to that of the PATTERN CONSTRUCTION test, however the target pattern area is replaced with a '?' (as in Figure A.16).

A.9 The Sentence Comprehension Test

Each trial of the SENTENCE COMPREHENSION test involves two stages: first, an instruction is given (as in Figure A.17). Upon proceeding, the user must keep the instruction in memory and carry it out by moving shapes (using click and drag) on the screen (as in Figure A.18); all shapes can be moved

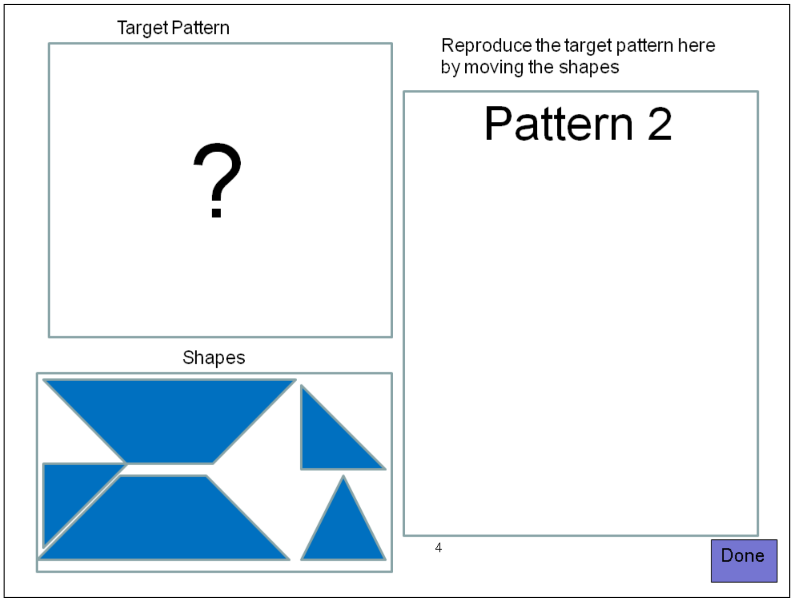


Figure A.16: The PATTERN RECALL test.

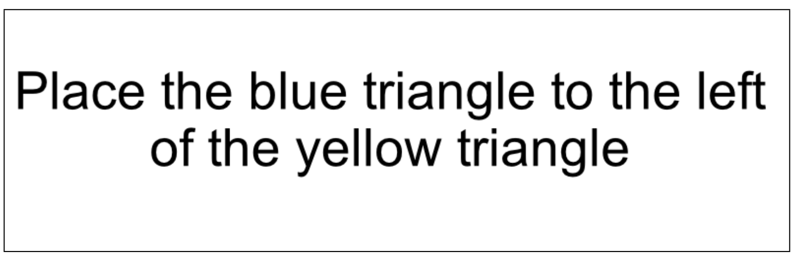


Figure A.17: The SENTENCE COMPREHENSION test (instruction screen).

but not rotated.

A.10 The Trails Test

In the TRAILS test, the user must move the mouse from the centre of the screen (Node 1, labeled 'Begin', is positioned there) to the next node in the series (1 - A - 2 - B - 3 - C - . . .), as shown in Figure A.19). Once the mouse cursor moves above the next correct node in the series, a line is constructed

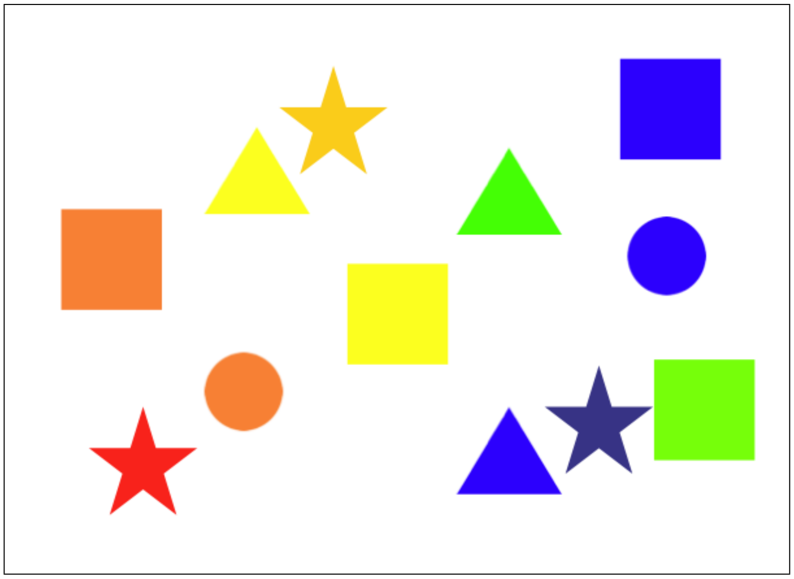


Figure A.18: The SENTENCE COMPREHENSION test (execution screen prior to any user interaction).

between it and the last completed node in the series. Initially, all nodes are unconnected. A completed puzzle is one in which all nodes are connected.

A.11 The Arithmetics Test

Each trial of the ARITHMETICS test poses a simple arithmetic problem using the 4 basic operators. A grid of clickable buttons corresponding to each number from 1 to 100 is provided; the user must click on the correct answer to the problem (see Figure A.20).

A.12 The Misplaced Object Search Test

Each trial in the MISPLACED OBJECT SEARCH test presents the user with a picture of a scene. Within the scene, an object that is not contextually appropriate can be found (i.e., a basketball in an office, a lawnmower in a bedroom, etc.). Locating and clicking on this object is the objective of this test. Clicking on the object once enlarges the object; clicking again proceeds

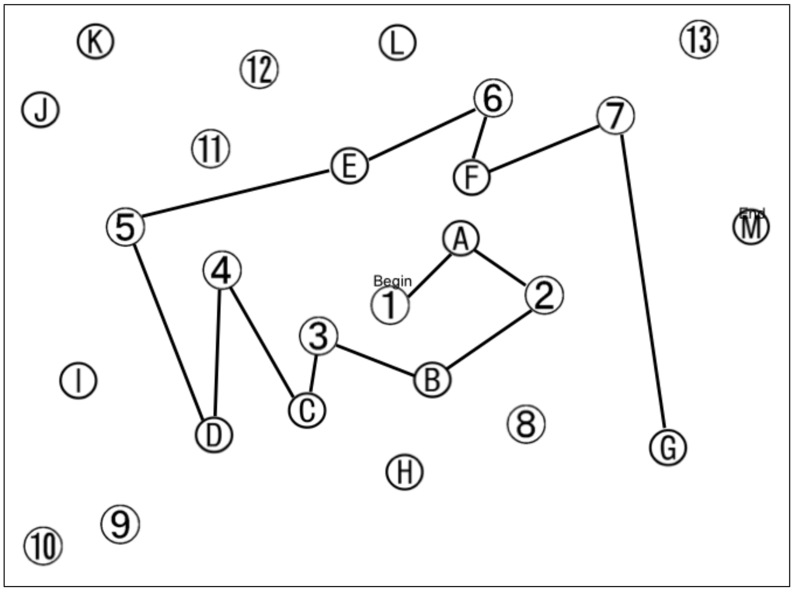
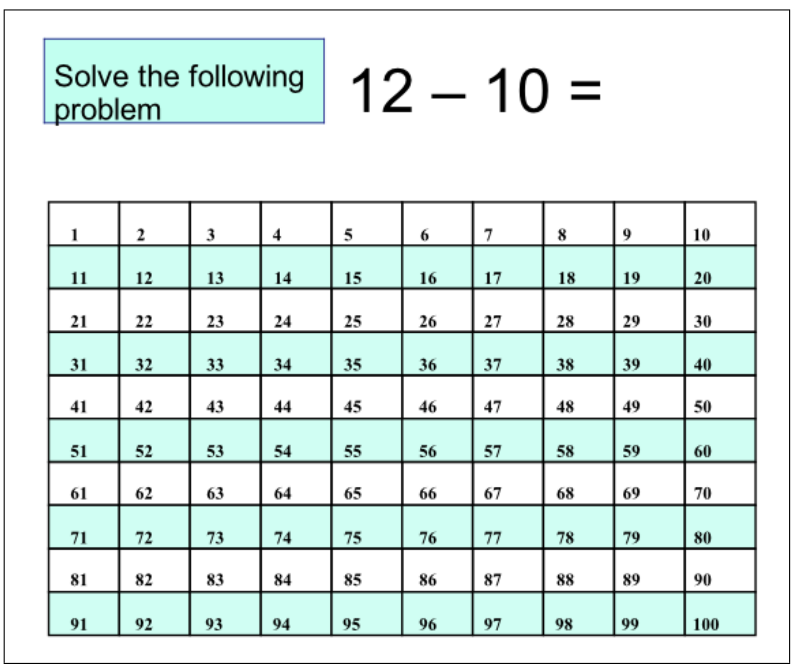


Figure A.19: The TRAILS test.



1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50
51	52	53	54	55	56	57	58	59	60
61	62	63	64	65	66	67	68	69	70
71	72	73	74	75	76	77	78	79	80
81	82	83	84	85	86	87	88	89	90
91	92	93	94	95	96	97	98	99	100

Figure A.20: The ARITHMETICS test.

Figure A.21: The MISPLACED OBJECT SEARCH test (in this instance, a basketball is placed above a lamp).

to the next trial. For example, a basketball is not contextually appropriate in an office scene (see Figure A.21).

A.13 The Misplaced Object Recall Test

MISPLACED OBJECT RECALL is the recall counterpart of the MISPLACED OBJECT SEARCH test (Section A.12). Each trial is carried out in 2 stages. First, an object prompt is given at the bottom of the screen and the user must select the scene it was seen in by clicking one of the scene thumbnails (as in Figure A.22). Second, the chosen scene is enlarged and the user must click on the location within the scene in which the object was seen (displayed as blue boxes, as in Figure A.23).

A.14 The Sentence Production Test

In the SENTENCE PRODUCTION test, the user is given an image of a scene containing actors and/or events taking place. They are instructed to build a

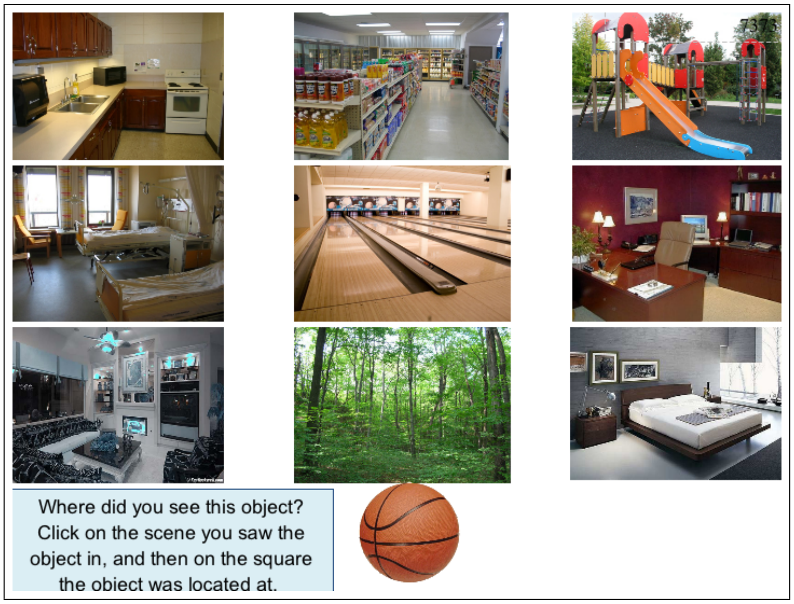


Figure A.22: The MISPLACED OBJECT RECALL test (step 1: selecting the scene from the MISPLACED OBJECT SEARCH test that contained the target misplaced object).

sentence to describe the scene using the list of word icons below the image. Word icons must be dragged into a sentence drop zone; while users are instructed to form detailed, grammatically-correct sentences, not all words in the list need to be used (as in Figure A.24).

A.15 The Square Puzzles Test

The SQUARE PUZZLES test is analogous to a game of moving matchsticks to form squares. A pattern of lines are displayed with the instruction to move a certain number of lines in order to create a number of complete squares (as in Figure A.25). The option to 'Start Over', or reset the pattern, was not completed for C-TOC.v1.

Figure A.23: The MISPLACED OBJECT RECALL test (step 2: selecting a location within the scene in which the misplaced object was seen by clicking on a blue box).

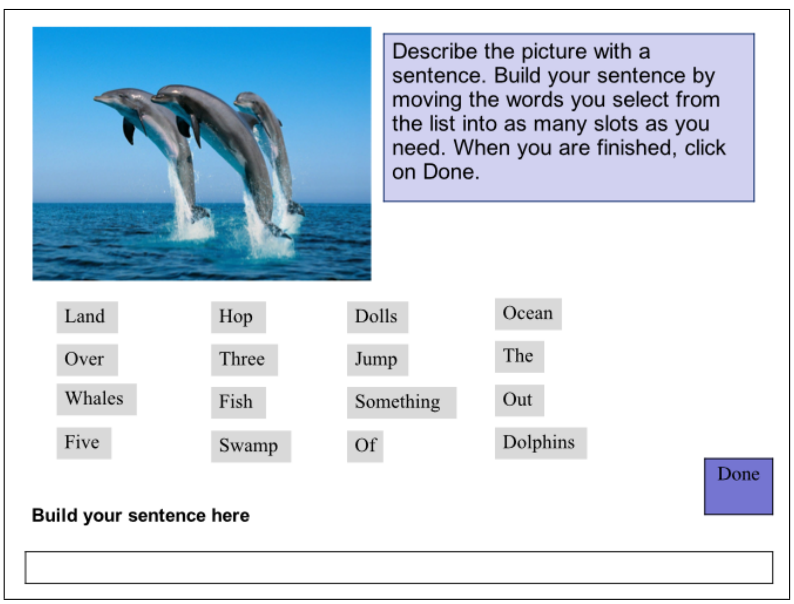


Figure A.24: The SENTENCE PRODUCTION test.

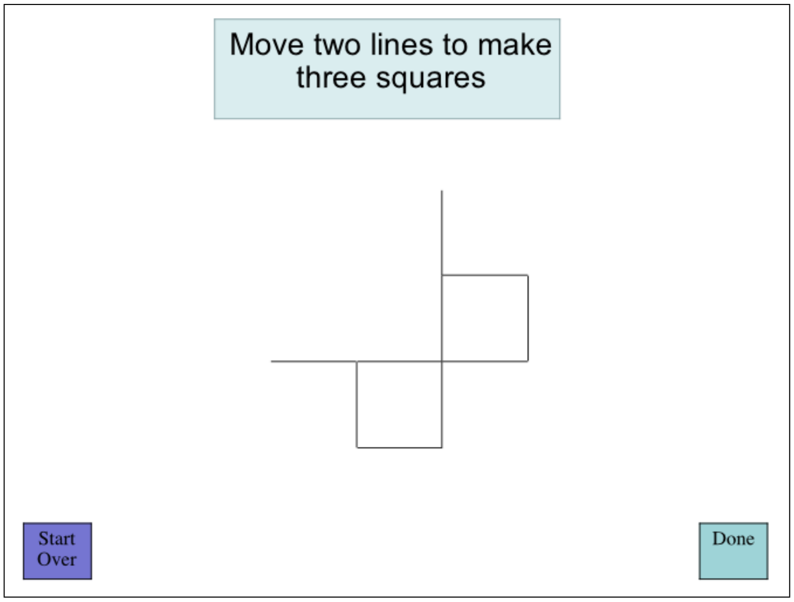


Figure A.25: The SQUARE PUZZLES test.

A.16 The Go-Stop Test

GO-STOP is a Go/No-Go inhibition test: an instruction is initially given to await for words to appear on screen (in between short randomly timed intervals), and to either click or not click depending on certain criteria, (as in Figure A.26). For example, the user may be instructed to click on all colour words except for 'Red'.

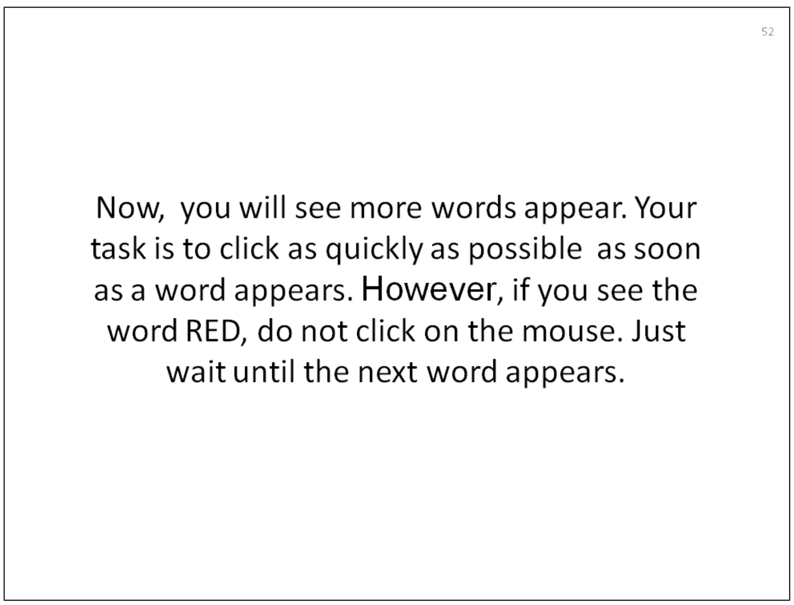


Figure A.26: The GO-STOP test.

Appendix B

Study 1 Resources

This appendix contains resources used in *Study 1* discussed in Chapter 3.

B.1 Usability Interview Script

The following interview script was used in 7 usability interviews.

B.1.1 Prior to Interaction with C-TOC

The following list of questions were adapted and expanded from previous HCI research with older adults by Goodman et al. [22].

Note: For the purposes of this interview, the term computer refers to any of the following: desktop, laptop/notebook, tablet, or handheld (such as a personal digital assistant like a Palm Pilot).

- When did you first use a computer, and for what purpose?
- How did you learn to use a computer? (e.g., computer classes/course, relatives/friends, work, self-taught, other)
- Do you have a computer in your household now? If yes, describe your home computer. How did you come to own your computer? (e.g., bought new, second-hand from friend/relative, selected by friend/relative, other)

- If not, where have they used / where do they use a computer? (e.g., library, at relatives/friends, community centre, other)
- How would you characterise yourself in terms of your knowledge of computers?
- What kinds of computers have you used? (e.g., Windows, Linux, Mac/Apple, Unix, Laptop/Notebook, Tablet, Handheld (PDA/Palm Pilot), Not sure, Other?)
- Do/did you use a computer for work? (either at home or work) If yes, on an average day (or week), approximately how many hours do/did you spend using a computer for work? What work-related tasks did you perform on a computer?
- Do you use a computer for leisure or personal tasks? If yes, on an average day (or week), approximately how many hours do you spend using a computer for this purpose? (e.g., web browsing (includes looking up health information), communicating with friends/relatives (email, chat, social networks), shopping)
- How familiar are you with the following types of computer programs?
  - Word processor (e.g., Microsoft Word)
  - Web Browser (e.g., Mozilla Firefox, Microsoft Internet Explorer)
  - Email (e.g., Microsoft Outlook, Lotus Notes, Thunderbird, gmail)
  - Spreadsheets (e.g., Microsoft Excel, Lotus 1-2-3)
  - Databases (e.g., MySql, Oracle)
  - Games (e.g., Solitaire, Hearts, online games, other)
  - Music/Video/Photos (e.g., iTunes, iPhoto, Quicktime, Windows Media Player)
  - Graphics Software (e.g., Adobe Photoshop / Illustrator)
  - Presentation software (e.g., Microsoft PowerPoint)

- Have you ever read help documentation to learn about using a computer or a particular computer program? If yes, how helpful was this process? Did you experience any common frustrations when reading documentation?
- How would you describe your overall enjoyment using computers?
- Are you comfortable using a mouse? Keyboard? Trackball / trackpad?
- If you have stopped using a computer, can you tell us why? Was it because you forgot how to use the technology or because you no longer needed it?

B.1.2 During Interaction with C-TOC

During/After Instruction Slides Only:

- Do you have any difficulty reading the text on the instruction screens? (Would it help if the text was larger?)

During/After Opening Slides:

- How easy is it to move to the next screen?
- How easy is it to use the mouse to respond during the practice items? Would you need more practice?
- How good is the help menu? Can you think of anything else that you might want in this menu?
- Is the amount of information on the screen reasonable? (Is it perhaps too cluttered, or too simple?)
- Is the information positioned in an understandable way? If not, how would you arrange the information?
- Can you show me how to access the help menu, go to the previous screen, etc.?

During/After Every Task:

- Did you have any difficulties using the mouse to respond to the test items? What difficulties?
- (If applicable) Would you prefer 'click and snap' or 'drag and drop'?
- Was anything on the screen distracting you from the task?
- Did you at any time have a feeling of being stuck or lost? Not knowing how to move to the next screen? What would make it easier for you?
- Imagine being at home while doing this task (on your own computer). What concerns would you have in terms of completing the it?
- Do you think you would perform better or worse than in the Clinic office?
- Do you feel like you might forget information on how to advance through the test? You might need reminders? Which do you think would work?

B.1.3 After Interaction with C-TOC

- With regards to using the computer, what were the major difficulties in self-administering the C-TOC battery?
- What additional functions, if any, should the C-TOC help menu have for test-takers like you?
- Are their any computer-related aspects of the test battery, ones that we haven't already discussed, that you would want to see changed or improved? What are they?
- How would you feel about taking C-TOC at home on the web using your own computer? What concerns would you have?
- Would you be able to get set up for the C-TOC and do everything by yourself? What would you need help with?

B.2 Cross-Cultural Advisory Panel Questionnaire

The following questionnaire was given to cross-cultural advisory panel members to complete as they worked through the C-TOC prototype independently.

Development of Cognitive Testing on Computer (C-TOC)
Cultural Advisory Meeting Cycle 1
May 27, 2010

C-TOC QUESTIONNAIRE

I. PERSONAL INFORMATION

1. Gender:

❑ F ❑ M

2. Education:

❑ High School
❑ Some College/University
❑ College/University undergraduate degree
❑ Post-graduate degree

3. Occupation: ___________________________________

4. Country of Birth: ______________________________

5. Years lived in Canada: ___________________________

6. Which ethnocultural community/ies do you consider yourself part of:

____

7. Do you have prior experience representing your community/ies in an advisory role?

❑ Yes ❑ No

8. How would you rate your level of computer expertise?

❑ None
❑ Low
❑ Moderate
❑ Extensive

9. How did you learn to use a computer? (check all that apply)

❑ Self-taught
❑ Friends / relatives

❑ Computer classes / school
❑ Work
❑ (Not applicable)
❑ Other:_________________

10. Which computer operating systems do you use? (check all that apply)

	N/A	Seldom	Sometimes	Often
Microsoft Windows XP	❑	❑	❑	❑
Microsoft Windows Vista / Windows 7	❑	❑	❑	❑
Apple Mac OS X	❑	❑	❑	❑
Linux/UNIX	❑	❑	❑	❑
Other *(specify)*:__________________	❑	❑	❑	❑

11. Which software applications do you use? (check all that apply)

	N/A	Seldom	Sometimes	Often
Word processor (*e.g. Microsoft Word*)	❑	❑	❑	❑
Email (*e.g. Microsoft Outlook, Lotus Notes, gMail*)	❑	❑	❑	❑
Web Browser (*e.g. Firefox, Internet Explorer*)	❑	❑	❑	❑
Spreadsheets (*e.g. Microsoft Excel, Lotus 1-2-3*)	❑	❑	❑	❑
Databases (*e.g. mySQL, Oracle*)	❑	❑	❑	❑
Games (*e.g. Solitaire, Hearts, online games, other*)	❑	❑	❑	❑
Music/Video/Photos (*e.g. iTunes, Quicktime, Windows Media Player, iPhoto*)	❑	❑	❑	❑
Graphics Software (*e.g. Adobe Photoshop/Illustrator*)	❑	❑	❑	❑
Presentation Software (*e.g. Microsoft PowerPoint*)	❑	❑	❑	❑

12. What, if any, are typical sources of confusion for you when using computer hardware and/or software?

____

____

II. C-TOC EVALUATION:

Please answer all questions from the perspective of people from your community

1. How familiar are people from your community with computer technology?

- ❑ Very familiar
- ❑ Somewhat familiar
- ❑ Neutral
- ❑ Somewhat unfamiliar
- ❑ Very unfamiliar

2. How clear is the Introduction to the C-TOC?

- ❑ Very clear
- ❑ Somewhat clear
- ❑ Neutral
- ❑ Somewhat unclear
- ❑ Very unclear

Please make suggestions on how to improve the Introduction

3. How adequate is the list of C-TOC Help Menu topics?

- ❑ Very adequate
- ❑ Somewhat adequate
- ❑ Neutral
- ❑ Somewhat inadequate
- ❑ Very inadequate

Please make suggestions on how to improve the Help Menu

4. How clear are the Information sections including Personal, Health, Computer Knowledge, and Family History?

❑ Very clear
❑ Somewhat clear
❑ Neutral
❑ Somewhat unclear
❑ Very unclear

Please make suggestions on how to improve the Information screens

5. How adequate are the mouse practice trials?

❑ Very adequate
❑ Somewhat adequate
❑ Neutral
❑ Somewhat inadequate
❑ Very inadequate

II.i – [SUBTEST NAME]

6. What is your impression of the **[SUBTEST NAME]** subtest?

Given the following responses:

1 = Very unlikely
2 = Somewhat unlikely
3 = Neutral
4 = Somewhat likely
5 = Very likely

How likely are people to understand: *(Circle the best answer)*

The task	1	2	3	4	5
The language	1	2	3	4	5
The contents	1	2	3	4	5
The required mouse action	1	2	3	4	5
The graphical interface *(i.e. on-screen buttons, navigation, layout of items on screen)*	1	2	3	4	5

Please provide additional comments on the **[SUBTEST NAME]** subtest:

III. MAKING C-TOC ACCEPTABLE

1. Think generally about C-TOC: would it better if there were few subtests with many items rather than many tasks with few items?

❑ Strongly agree
❑ Somewhat agree
❑ Neutral
❑ Somewhat disagree
❑ Strongly disagree

2. If there were room in C-TOC only for one **memory** subtest, which would you choose?

❑ Picture/Word Pairs & Word Recognition
❑ Pattern Construction & Recall
❑ Misplaced Objects Search & Recall

3. If there were room in C-TOC only for one **attention** subtest, which would you choose?

❑ Symbol Digit Matching
❑ Trails
❑ Go-Stop

4. If there were room in C-TOC only for one **language** subtest, which would you choose?

❑ Sentence Comprehension
❑ Sentence Production
❑ Similarities

5. What do you think is the ideal length of C-TOC as an initial screen prior to a visit with the specialist?

❑ < 10 minutes
❑ Between 10-20 minutes
❑ Between 20-30 minutes

6. Do people from your community consult a doctor if they have complaints about their memory or thinking?

❑ Very likely
❑ Somewhat likely
❑ Neutral
❑ Somewhat unlikely
❑ Very unlikely

7. Would people from your community be willing to take a computerized test to provide advance information to the doctor?

At the doctor's office:

- ❑ Very likely
- ❑ Somewhat likely
- ❑ Neutral
- ❑ Somewhat unlikely
- ❑ Very unlikely

From their own home:

- ❑ Very likely
- ❑ Somewhat likely
- ❑ Neutral
- ❑ Somewhat unlikely
- ❑ Very unlikely

Please provide additional comments on the acceptability of C-TOC for people from your community. How can we make it more acceptable?

B.3 Cross-Cultural Advisory Panel Focus Group Questions

The following questions were asked of cross-cultural advisory panel members after completing an independent interactive session with the C-TOC prototype.

1. What is your general sense of the purpose and applicability of the tool to your community?
2. How likely are people from your community to
   (a) Consult a doctor for evaluation of cognitive impairments?
   (b) Accept C-TOC as an initial screen for their cognitive complaints?
3. Would they be willing to complete C-TOC? Would they be willing to self-administer C-TOC? Would they be willing to take C-TOC at home?
4. What barriers would you particularly identify to people using C-TOC? (i.e., literacy, computer illiteracy, familiarity with computer technology, concerns about privacy, test motivation and/or anxiety, language).

Appendix C

Study 2 Resources

This appendix contains resources used in *Study 2*, discussed in Chapter 4.

C.1 Recruitment Poster

The following study recruitment poster was posted throughout the community. Locations included the UBC campus, Vancouver Public Library branches, Vancouver community and seniors' centres, and seniors' housing complexes.

The body text of the recruitment poster also appeared in the UBC Professors Emeritus Association newsletter, the UBC HCI-experiments mailing list, and in online classified advertisements (Craigslist, Kijiji).

The UNIVERSITY OF BRITISH COLUMBIA
Department of Computer Science / Medicine
University of British Columbia
Vancouver, BC, V6T 1Z4

Adults Age 70+ Needed for UBC Research Study

Usability Evaluation of an Online Cognitive Health Assessment Tool

Study Recruitment

Principal Investigator: Claudia Jacova, PhD (Medicine)

Co-Investigators: Matthew Brehmer, M.Sc Student (Computer Science),
Joanna McGrenere, PhD (Computer Science)
Charlotte Tang, PhD (Computer Science)

Purpose: This study is designed to investigate how people interact with an online cognitive health assessment tool which involves recall from memory and other cognitive processes. The purpose of this study is to evaluate the usability of the tool's components in order to improve its design.

Participants: We are looking for adults aged 70 and older, who:

- Are healthy, and have normal or corrected-to-normal eyesight,
- Are free of diagnosed cognitive impairments or motor impairments to their hands

Procedure: You will be asked to perform a number of tasks while we record aspects of your performance, including task completion time and response accuracy. You will also be asked interview questions about your experience in performing the tasks, e.g. difficulties encountered. Photographs/Videos may be taken with your permission.

Objective: The research objective is to inform and refine the design of an online tool that is intended for cognitive health care purposes. To achieve this, we need to identify any usability issues associated with the tasks to be performed during use of the tool. With this greater understanding we can continue to design effective and usable health care technologies.

Commitment: Your participation in this study will involve 1 session that will require no more than 2 hours of your time and you will be monetarily compensated for your time.

To Participate:
Please contact Matthew at ###_###_#### for more information.

C.2 Participant Consent Form

The following is a copy of the consent form participants were required to sign in order to participate in the study. Whenever possible, participants were emailed a PDF copy of the form 1-2 days prior to their scheduled session.

The UNIVERSITY OF BRITISH COLUMBIA
Department of Computer Science[1] / Medicine[2]
University of British Columbia
Vancouver, BC, V6T 1Z4

Consent Form

Research Project Title: **Development of a Computer-Based Screening Test to Support Evaluation of Cognitive Impairment and Dementia**
(Part 1C - Usability Evaluation of an Online Cognitive Assessment Tool)

Principal investigator: Claudia Jacova, PhD, ###_###_#### (Medicine)

Co-Investigators: Matthew Brehmer, MSc Student, ###_###_####
Joanna McGrenere, PhD, ###_###_####
Charlotte Tang, PhD, ###_###_####
Ging-Yuek Robin Hsiung, MD, MHSc, FRCPC, ###_###_####
Lynn Beattie, MD, FRCPC, ###_###_####
Philip Lee, MD, FRCPC, ###_###_####
Dean Foti, MD, FRCPC, ###_###_####
Sherri Hayden, PhD, R.Psych, ###_###_####

In this study, we aim to identify usability issues associated with selected task components of a novel computer-based cognitive test battery, called Cognitive Testing on Computer (C-TOC). You are being invited to participate in this study because you are 18 years of age or older without any diagnosed cognitive impairments or motor impairments to your hands. Your participation will help us probe the usability of C-TOC task components.

Your participation in this research study is entirely voluntary. This consent form, a copy of which has been given to you, is only part of the process of informed consent. It should give you the basic idea of what the research is about and what your participation will involve. If you would like more detail about something mentioned here, or information not included here, you should feel free to ask. Please take the time to read this carefully and to understand any accompanying information.

If you wish to participate, you will be invited to sign this form but you should understand that you are free to withdraw your consent at any time and without giving any reasons for your decision.

Purpose: This study is designed to investigate how people interact with an online cognitive health assessment tool which involves recall from memory and other cognitive processes. The purpose of this study is to evaluate the usability of the tool and improve its design.

Procedure: Your participation in this study will involve 1 session that will require no more than 2 hours of your time. During this session, you will be asked to perform a number of tasks on a desktop computer. We will record aspects of your performance, including task completion time and accuracy. This test is not meant to test your skills or experience with computers; it is only being carried out to probe the usability of C-TOC task components. You will also be asked interview questions about your experience in performing the tasks, e.g. difficulties encountered. In all circumstances, you do not need to answer any questions that you do not feel comfortable answering.

Objective: The research objective is to inform and refine the design of an online tool that is intended for cognitive health care purposes. To achieve this, we need to identify all usability issues that may affect people's performance on the tasks that are presented in the online tool. This knowledge will help us design effective and usable health care technologies.

Option for Photographing/Videotaping:
For the purpose of data analysis, we would like to videotape and/or photograph your computer session and your interview. Please note that this is an optional procedure, which you are free to decline, and a refusal to videotape or photograph will in no way affect your eligibility for this study. Only the investigators of this study will have access to the recordings. The recordings will be stored in a secured departmental network of Computer Science for three years after the study, which will then be permanently erased. Participants' identity will be protected by masking in publications and presentations. Please check and initial the ones you agree.

- I agree that the researchers may videotape my computer session. __________
- I agree that the researchers may videotape my interview. __________
- I agree that the researchers may use the photographs taken during the study without modification, except for masking identities, for illustrative purposes in the disseminatiion of the study's results, including but not limited to, presentations and publication of papers and/or theses. __________

What are the Possible Harms and Side Effects of Participating?
You may experience fatigue from performing the computer tasks and answering the questions.

What are the Benefits of Participating in this Research?
There may be no immediate, direct benefit to you as a result of participating in this study. However the findings from this study can help us improve future health care technologies that may benefit you, your family members and the community in the longer term.

What Happens If I Decide to Withdraw My Consent to Participate?
Your participation in this research is entirely voluntary. You may withdraw from this study at any time, and are not required to provide any reason for withdrawing. If you choose to enter the study and then decide to withdraw at a later time, all data collected about you during your enrollment in the study will be retained for analysis. By law, this data cannot be destroyed. If you wish to withdraw your consent, we ask that you notify Dr. Claudia Jacova at ###_###_#### or Matthew Brehmer at ###_###_####.

What Happens If Something Goes Wrong?
Signing this consent form in no way limits your legal rights against the study sponsor, investigators, or anyone else.

Will My Taking Part in this Study be Kept Confidential?
Your confidentiality will be respected. The Investigators in this study will be responsible for maintaining your confidentiality at all times. Study records will be labeled only with an assigned numeric code. They will not include information that identifies you by name, initials, or date of birth. This code number and the connection of the code number to your name and identifying

information will be stored in a private, password-protected computer in the Department of Computer Science at the UBC ICICS/CS Building. Access to personal identifying information will be restricted to the Principal Investigator, Co-Investigators, and their research study staff.

Results from this study may be presented at meetings and may be published, but no information that discloses your identity will be released or published without your specific consent to the disclosure. However, research records and medical records identifying you may be inspected in the presence of the Investigator or his or her designate, and the UBC Research Ethics Board for the purpose of monitoring the research. However, no records which identify you by name or initials will be allowed to leave the Investigators' offices.

Who do I Contact if I have any Questions or Concerns about the Study?
If you have any questions or desire further information with respect to this research, you should contact Dr. Claudia Jacova at ###_###_#### or Matthew Brehmer at ###_###_####. If you have any concerns about your rights as a research subject and/or your experiences while participating in this study, you should contact the Research Subject Information Line at the University of British Columbia's Office of Research Services at ###_###_####.

Subject Consent to Participate:

- I have read and understood the subject information and consent form.
- I have had sufficient time to consider the information provided and to ask for advice if necessary.
- I have had the opportunity to ask questions and have had satisfactory responses to my questions.
- I understand that all of the information collected will be kept confidential and that the results will only be used for scientific objectives such as research and publications.
- I understand that I can refuse to answer any questions that I do not feel comfortable answering from this study.
- I understand that my participation in this study is voluntary and that I am completely free to refuse to participate or to withdraw from this study at any time.
- I understand that I am not waiving any of my legal rights as a result of signing this consent form.
- I understand that there is no guarantee that this study will provide any benefits to me.
- I have read this form and I freely consent to participate in this study.
- I have been told that I will receive a dated and signed copy of this form.

Signatures

Printed Name of Participant Signature and Date

Principal Investigator or designated representative Signature and Date

C.3 Participant Screening Materials

Participants were screened using the MOCA [40] and the NAART [53].

C.3.1 Cognitive Impairment

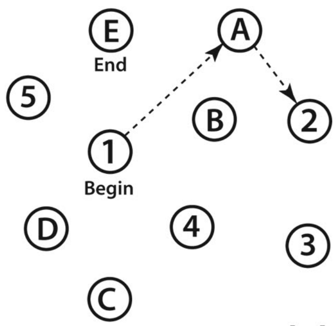
The MOCA [40] was adminsitered to help ensure that participants had no existing cognitive impairment. Participants required a score of 26 or higher (out of 30) on the MOCA, which corresponds to NCI.

Participants who did not meet this criteria were allowed to finish the study, but their data were not included in the analysis. These participants completed a shorter version of the study (see Section C.5.2).

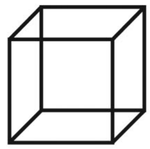
MONTREAL COGNITIVE ASSESSMENT (MOCA)
Version 7.1 Original Version

NAME :
Education : Date of birth :
Sex : DATE :

		POINTS
VISUOSPATIAL / EXECUTIVE		



[ ]



Copy cube

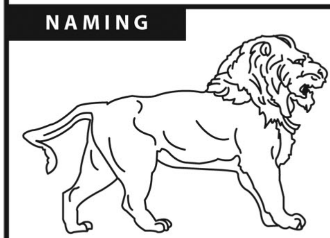
[ ]

Draw CLOCK (Ten past eleven)
(3 points)

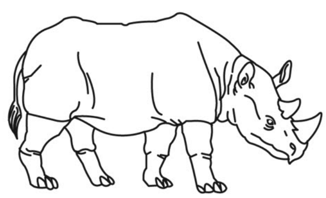
[ ] Contour [ ] Numbers [ ] Hands

___/5

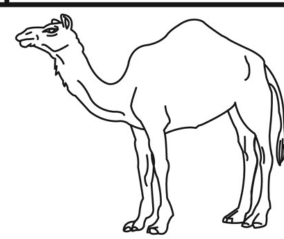
NAMING



[ ]



[ ]



[ ]

___/3

MEMORY Read list of words, subject must repeat them. Do 2 trials, even if 1st trial is successful. Do a recall after 5 minutes.

	FACE	VELVET	CHURCH	DAISY	RED
1st trial					
2nd trial					

No points

ATTENTION Read list of digits (1 digit/ sec.).

Subject has to repeat them in the forward order [ ] 2 1 8 5 4
Subject has to repeat them in the backward order [ ] 7 4 2

___/2

Read list of letters. The subject must tap with his hand at each letter A. No points if ≥ 2 errors
[ ] F B A C M N A A J K L B A F A K D E A A A J A M O F A A B

___/1

Serial 7 subtraction starting at 100 [ ] 93 [ ] 86 [ ] 79 [ ] 72 [ ] 65
4 or 5 correct subtractions: **3 pts**, 2 or 3 correct: **2 pts**, 1 correct: **1 pt**, 0 correct: **0 pt**

___/3

LANGUAGE Repeat : I only know that John is the one to help today. [ ]
The cat always hid under the couch when dogs were in the room. [ ]

___/2

Fluency / Name maximum number of words in one minute that begin with the letter F [ ] _____ (N ≥ 11 words)

___/1

ABSTRACTION Similarity between e.g. banana - orange = fruit [ ] train – bicycle [ ] watch - ruler

___/2

DELAYED RECALL

	FACE	VELVET	CHURCH	DAISY	RED	
Has to recall words **WITH NO CUE**	[ ]	[ ]	[ ]	[ ]	[ ]	Points for UNCUED recall only
Optional Category cue						
Multiple choice cue						

___/5

ORIENTATION [ ] Date [ ] Month [ ] Year [ ] Day [ ] Place [ ] City

___/6

 www.mocatest.org Normal ≥ 26 / 30

TOTAL ___/30
Add 1 point if ≤ 12 yr edu

Administered by: ______________________________

Montreal Cognitive Assessment
(MoCA)

Administration and Scoring Instructions

The Montreal Cognitive Assessment (MoCA) was designed as a rapid screening instrument for mild cognitive dysfunction. It assesses different cognitive domains: attention and concentration, executive functions, memory, language, visuoconstructional skills, conceptual thinking, calculations, and orientation. Time to administer the MoCA is approximately 10 minutes. The total possible score is 30 points; a score of 26 or above is considered normal.

1. Alternating Trail Making:

Administration: The examiner instructs the subject: *"Please draw a line, going from a number to a letter in ascending order. Begin here* [point to (1)] *and draw a line from 1 then to A then to 2 and so on. End here* [point to (E)]."

Scoring: Allocate one point if the subject successfully draws the following pattern:
1 −A- 2- B- 3- C- 4- D- 5- E, without drawing any lines that cross. Any error that is not immediately self-corrected earns a score of 0.

2. Visuoconstructional Skills (Cube):

Administration: The examiner gives the following instructions, pointing to the **cube**: "*Copy this drawing as accurately as you can, in the space below*".

Scoring: One point is allocated for a correctly executed drawing.
• Drawing must be three-dimensional
• All lines are drawn
• No line is added
• Lines are relatively parallel and their length is similar (rectangular prisms are accepted)

A point is not assigned if any of the above-criteria are not met.

3. Visuoconstructional Skills (Clock):

Administration: Indicate the right third of the space and give the following instructions: *"Draw a **clock**. Put in all the numbers and set the time to 10 past 11".*

Scoring: One point is allocated for each of the following three criteria:
• Contour (1 pt.): the clock face must be a circle with only minor distortion acceptable (e.g., slight imperfection on closing the circle);
• Numbers (1 pt.): all clock numbers must be present with no additional numbers; numbers must be in the correct order and placed in the approximate quadrants on the clock face; Roman numerals are acceptable; numbers can be placed outside the circle contour;
• Hands (1 pt.): there must be two hands jointly indicating the correct time; the hour hand must be clearly shorter than the minute hand; hands must be centred within the clock face with their junction close to the clock centre.

A point is not assigned for a given element if any of the above-criteria are not met.

4. Naming:

Administration: Beginning on the left, point to each figure and say: *"Tell me the name of this animal".*

Scoring: One point each is given for the following responses: (1) lion (2) rhinoceros or rhino (3) camel or dromedary.

5. Memory:

Administration: The examiner reads a list of 5 words at a rate of one per second, giving the following instructions*: "This is a memory test. I am going to read a list of words that you will have to remember now and later on. Listen carefully. When I am through, tell me as many words as you can remember. It doesn't matter in what order you say them".* Mark a check in the allocated space for each word the subject produces on this first trial. When the subject indicates that (s)he has finished (has recalled all words), or can recall no more words, read the list a second time with the following instructions: *"I am going to read the same list for a second time. Try to remember and tell me as many words as you can, including words you said the first time."* Put a check in the allocated space for each word the subject recalls after the second trial.
At the end of the second trial, inform the subject that (s)he will be asked to recall these words again by saying, *"I will ask you to recall those words again at the end of the test."*

Scoring: No points are given for Trials One and Two.

6. Attention:

Forward Digit Span: Administration: Give the following instruction: "*I am going to say some numbers and when I am through, repeat them to me exactly as I said them*". Read the five number sequence at a rate of one digit per second.

Backward Digit Span: Administration: Give the following instruction: "*Now I am going to say some more numbers, but when I am through you must repeat them to me in the backwards order.*" Read the three number sequence at a rate of one digit per second.

Scoring: Allocate one point for each sequence correctly repeated, (*N.B.*: the correct response for the backwards trial is 2-4-7).

Vigilance: Administration: The examiner reads the list of letters at a rate of one per second, after giving the following instruction: "*I am going to read a sequence of letters. Every time I say the letter A, tap your hand once. If I say a different letter, do not tap your hand*".

Scoring: Give one point if there is zero to one errors (an error is a tap on a wrong letter or a failure to tap on letter A).

Serial 7s: Administration: The examiner gives the following instruction: *"Now, I will ask you to count by subtracting seven from 100, and then, keep subtracting seven from your answer until I tell you to stop."* Give this instruction twice if necessary.

Scoring: This item is scored out of 3 points. Give no (0) points for no correct subtractions, 1 point for one correction subtraction, 2 points for two-to-three correct subtractions, and 3 points if the participant successfully makes four or five correct subtractions. Count each correct subtraction of 7 beginning at 100. Each subtraction is evaluated independently; that is, if the participant responds with an incorrect number but continues to correctly subtract 7 from it, give a point for each correct subtraction. For example, a participant may respond "92 – 85 – 78 – 71 – 64" where the "92" is incorrect, but all subsequent numbers are subtracted correctly. This is one error and the item would be given a score of 3.

7. Sentence repetition:

Administration: The examiner gives the following instructions: *"I am going to read you a sentence. Repeat it after me, exactly as I say it* [pause]: ***I only know that John is the one to help today."*** Following the response, say: *"Now I am going to read you another sentence. Repeat it after me, exactly as I say it* [pause]: ***The cat always hid under the couch when dogs were in the room."***

Scoring: Allocate 1 point for each sentence correctly repeated. Repetition must be exact. Be alert for errors that are omissions (e.g., omitting "only", "always") and substitutions/additions (e.g., "John is the one who helped today;" substituting "hides" for "hid", altering plurals, etc.).

8. Verbal fluency:

Administration: The examiner gives the following instruction: *"Tell me as many words as you can think of that begin with a certain letter of the alphabet that I will tell you in a moment. You can say any kind of word you want, except for proper nouns (like Bob or Boston), numbers, or words that begin with the same sound but have a different suffix, for example, love, lover, loving. I will tell you to stop after one minute. Are you ready?* [Pause] *Now, tell me as many words as you can think of that begin with the letter F.* [time for 60 sec]. *Stop."*

Scoring: Allocate one point if the subject generates 11 words or more in 60 sec. Record the subject's response in the bottom or side margins.

9. Abstraction:

Administration: The examiner asks the subject to explain what each pair of words has in common, starting with the example: *"Tell me how an orange and a banana are alike"*. If the subject answers in a concrete manner, then say only one additional time: *"Tell me another way in which those items are alike"*. If the subject does not give the appropriate response *(fruit)*, say, *"Yes, and they are also both fruit."* Do not give any additional instructions or clarification. After the practice trial, say: *"Now, tell me how a train and a bicycle are alike"*. Following the response, administer the second trial, saying: *"Now tell me how a ruler and a watch are alike"*. Do not give any additional instructions or prompts.

Scoring: Only the last two item pairs are scored. Give 1 point to each item pair correctly answered. The following responses are acceptable:
Train-bicycle = means of transportation, means of travelling, you take trips in both;
Ruler-watch = measuring instruments, used to measure.
The following responses are **not** acceptable: Train-bicycle = they have wheels; Ruler-watch = they have numbers.

10. Delayed recall:

Administration: The examiner gives the following instruction: *"I read some words to you earlier, which I asked you to remember. Tell me as many of those words as you can remember."* Make a check mark (√) for each of the words correctly recalled spontaneously without any cues, in the allocated space.

Scoring: **Allocate 1 point for each word recalled freely without any cues.**

Optional:
Following the delayed free recall trial, prompt the subject with the semantic category cue provided below for any word not recalled. Make a check mark (√) in the allocated space if the subject remembered the word with the help of a category or multiple-choice cue. Prompt all non-recalled words in this manner. If the subject does not recall the word after the category cue, give him/her a multiple choice trial, using the following example instruction, *"Which of the following words do you think it was, NOSE, FACE, or HAND?"*
Use the following category and/or multiple-choice cues for each word, when appropriate:

FACE:	category cue: part of the body	multiple choice: nose, face, hand
VELVET:	category cue: type of fabric	multiple choice: denim, cotton, velvet
CHURCH:	category cue: type of building	multiple choice: church, school, hospital
DAISY:	category cue: type of flower	multiple choice: rose, daisy, tulip
RED:	category cue: a colour	multiple choice: red, blue, green

Scoring: **No points are allocated for words recalled with a cue.** A cue is used for clinical information purposes only and can give the test interpreter additional information about the type of memory disorder. For memory deficits due to retrieval failures, performance can be improved with a cue. For memory deficits due to encoding failures, performance does not improve with a cue.

11. Orientation:

Administration: The examiner gives the following instructions: "Tell me the date today". If the subject does not give a complete answer, then prompt accordingly by saying: *"Tell me the [year, month, exact date, and day of the week]."* Then say: *"Now, tell me the name of this place, and which city it is in."*

Scoring: Give one point for each item correctly answered. The subject must tell the exact date and the exact place (name of hospital, clinic, office). No points are allocated if subject makes an error of one day for the day and date.

TOTAL SCORE: Sum all subscores listed on the right-hand side. Add one point for an individual who has 12 years or fewer of formal education, for a possible maximum of 30 points. A final total score of 26 and above is considered normal.

C.3.2 English Language Fluency

The NAART [53] was administered to help ensure participants had sufficient English fluency to follow our instructions. The NAART is a quick to administer test measuring verbal intelligence, which requires participants to read a list of 30 words increasing in difficulty. Participants were required to read at least 50% of words used in the NAART correctly.

Participants who did not meet this criteria were allowed to finish the study, but their data were not included in the analysis. These participants completed a shorter version of the study (see Section C.5.2).

The NAART word list used in the experiment follows.

debt
debris
aisle
reign
depot
simile
lingerie
recipe
gouge
heir
subtle
catacomb
bouquet
gauge
colonel
subpoena
placebo
procreate
psalm
banal
rarefy
gist
corps
hors d’oeuvre
sieve
hiatus
gauche
zealot
paradigm
façade

C.4 Experimental Software

The experimental software was written as an AIR application using the Adobe Flex 4.0 SDK. As of August 22, 2011, it is available for download at http://cs.ubc.ca/~brehmer/research/expt_software.zip. The package contains installer files for PC and Mac, as well as a certificate needed for installation. This will install Adobe AIR, which allows web applications to run as standalone desktop applications.

The application records all timing and accuracy data in 4 log files, saved in the My Documents folder (PC) or the Documents folder (Mac). Each time the application is run, the log files are generated with filenames that include group and subject identifiers, the current date, and the current time.

Three of the log files are in tab-delimited format, corresponding to results from the VERBAL (SC) and SPATIAL (SQP) primary tasks, as well as the ACTIVE interruption task (nBack). These files can be easily imported into SPSS with import script files, also provided in the above package.

The fourth text file contains a time-stamped log of all participant activity (i.e., clicking and dragging objects, etc.) while the application is open.

C.4.1 Screen Capture Software

CamStudio, a screen-capture application, was used to record all experimental sessions, which was useful for scoring primary task accuracy.

C.5 Experimenter Script

For consistency, experimenters followed a script in each study session.

C.5.1 Participant Instructions & Interview

Participants who met the cutoff criteria on the MOCA and NAART tests completed the full version of the study. The following script is what was read to them.

Interruption Lab Study

Experiment Script

Introduction

Hello, my name is _____. Thank you for participating in our study. The first thing I'd like you to do is read over this consent form and sign it.

- *Give participant a consent form; ask if they have any questions or require clarification.*

Thank you. I would like you to complete a short assessment test before we begin.

- *Administer the MoCA (following the MoCA administration and scoring instructions); Should the participant score less than 26, proceed to alternate script.*

Now I'd like you to read the words on this list. This is just a test of your knowledge of words. I'd like you to read them out loud one at a time going down the list by column. They get harder and harder as you go along. If you are not sure how to pronounce a word, just give it a try. Don't worry; most people don't know a lot of these words.

- *Administer the NAART; participant reads the 2-page list of irregularly-pronounced words, increasing in difficulty. Participant begins on page 1, column 1; followed by column 2, (flips over page) page 2: column 1, column 2; should the participant get less than 50% of the words correct on page 1, proceed to alternate script.*

Now I'll give you an overview of today's study. Don't worry if this is too much information to take in all at once.

Today, you are going to be working with two types of puzzle tasks. Both tasks involve moving shapes around the screen with the mouse. You will only need to use the left-mouse button. No keyboard use will be necessary. You will perform a number of instances of both of these puzzle tasks, occurring one after the other, which we'll call trials.

During some trials, you will be interrupted. These interruptions will last for approximately 20 seconds, delaying your execution of the puzzle task trial. During these interruptions, you will be required to complete one of two interrupting tasks. These interrupting tasks are purposefully disruptive and are meant to distract you and keep you from thinking about the puzzle task. We are interested in your performance both on the puzzle tasks and on the interrupting tasks.

I would like you to imagine a real-world example, in which your main task is one of web browsing or word processing and an interrupting task is an urgent phone call or email that you absolutely must respond to as quickly as possible, postponing your completion of your main task.

There are several phases to this experiment. First, you will experience the two interrupting tasks by themselves until you are familiar with them. Following this, you will practice completing several trials of the first of the two puzzle tasks. Then, you will then perform 3 sets of trials of this task. During these trials you must remain ready to respond to the interrupting tasks when they occur.

Upon completion of these trials, you will repeat this process for the second of the two puzzle tasks.

Again, don't worry if this overview was a lot of information. You will get to practice on example tasks at each step of the study.

- *Ensure settings are correct and start recording (fullscreen) in CamStudio. Open the BrainFreeze application. Referring to the experiment lookup table, enter the group ID, subject ID, interruption condition order, and main-task order on the splash page's form. Click the "Full-Screen" button, then the "Next" button to enter the main experiment portal screen.*

Interruption Tasks Examples

You will now perform the two interrupting tasks by themselves. At the top left of the screen, you can see two buttons: *Interruption 1 Example* and *Interruption 2 Example*. Please wait for me to explain before clicking.

In *Interruption 1*, you will be shown an automated sequence of a dozen cartoon images. Your task is to watch these images passively. Be patient, and you do not need to click on the images.

- *Demonstrate printed example of Interruption 1*

  Images will appear at a constant rate in this box in the center of the screen. Afterwards, you will be asked to click in order to dismiss the interruption.

When ready, click on *Interruption 1 Example.*

- *Participant performs Interruption 1 Example.*

Good. In Interruption 2, you will once again be shown an automated sequence of a dozen cartoon images; however you will not remain passive. Instead, your task is to pay close attention to this sequence, and to keep track of what you saw two images prior to the current image. Allow me to explain.

- *Demonstrate 3 printout examples of Interruption 2*

  When the current image repeats what you saw two images ago, click inside the box where the image is shown. You must be quick to click before the sequence advances. Otherwise, do not click. Here is an example sequence of images and the responses you would receive by clicking on them. You will be shown a green check if your click is the correct response. You will be shown a red 'X' if you make an incorrect response: if the image you click on was not the repeat what was displayed 2 images ago OR if you don't click when you should have clicked: in other words if the image displayed repeats what was displayed 2 images ago and you don't click.

After 12 images have been displayed, you will be prompted to dismiss the interruption. When ready, click on *Interruption 2 Example*.

- *Participant performs Interruption 2 Example.*

I encourage you now to perform this example again to gain additional practice.

- *Ask if participant understands the difference between the two interrupting tasks. Use visual aids for elaboration if necessary.*

- *Participant performs Interruption 2 Example again. Repeat a 3rd time if requested (they must have performed very poorly during the first 2 attempts); if clicking randomly after 3rd attempt, proceed to alternate script.*

We are now ready to start the first of the two puzzle tasks.

Task A (Sentence Comprehension)

In Task A, you will read instructions to arrange shapes on the screen, and subsequently carry out these instructions to the best of your recollection and understanding.

- *Demonstrate printed example of Task A, pointing out every relevant UI element (continue button, instruction text, etc.)*

  The first part of each puzzle trial involves an instruction, like the one shown here. When you're ready, clicking "Continue" will advance you to the second part of the trial, in which you carry out the instruction. You cannot go back to read the instruction. You will move shapes like these around the screen by clicking on them and dragging them with the mouse. When finished, click "Next" to advance to the next trial.

It is best if you experience an example of this, so when ready, click the *Task A Example* button.

- *Participant completes Task A Example*

Good. Now you will perform a sequence of 3 practice trials of this task. During this sequence you will also need to respond to both types of interrupting tasks. An interruption may occur at any time during a trial. Upon completion of an interrupting task, you will click to dismiss the interruption and return to the puzzle task trial which was interrupted, continuing exactly where you left off. You are encouraged to complete each trial both as quickly and as accurately as possible. When ready, click on *Task A Practice*.

- *Participant completes Task A Practice*

You will now complete the experimental trials, again please complete each trial as quickly and as accurately as possible. You will perform 3 sets of 10 trials for a total of 30 trials. After each set, you will be given an opportunity to rest and I will ask you to complete a short survey. Once again, I will remind you that an interruption may occur at any time during a trial. Many trials will be uninterrupted. We are interested in your performance both on the puzzle tasks and on the interrupting tasks. In addition, recall that the interrupting tasks have been intentionally designed to distract you and keep you from thinking about the puzzle task, so don't allow them to frustrate you too much, and just do the best that you can.

- *Ensure that your marking sheet is ready and labeled with group and subject ID (i.e. g2s14). For each task A.{N}, where N = {1,2,3}:*

When ready, click on *Task A.{N}.*

- *Participant completes trial bank A.{N}.*

Please fill out this section of the questionnaire.

- *Label the header of the questionnaire accordingly. Participant fills out section of questionnaire relevant to A.{N}*

This concludes Task A. [*If Task A occurs first*: You have completed the first of two main tasks. We will now complete the second task.] [*Otherwise, stop recording in CamStudio and save the video file.]*

Task B (Square Puzzles)

In Task B, you will move lines to create complete squares in a specified number of moves. A complete solution will contain no incomplete squares and will require no fewer and no more moves than the number specified.

- *Demonstrate printed example of Task B, pointing out every relevant UI element (continue button, instruction text, etc.)*

You will see an instruction at the bottom of the screen. Move the lines around the white area by clicking on them and dragging them with the mouse. Lines cannot be rotated. Remember that squares have equal height and width (one vertical line by one horizontal line). Every move counts, so try not to rely on trial and error. You cannot "undo" moves, so plan carefully. For example, if I were to move a line at the beginning of a trial, but later decide to move the line back to its original position, this would be counted as 2 moves. When finished, click "Next" to advance to the next trial.

Can you show me how you would solve this puzzle?

- *Correct the participant if he/she is incorrect. Clarify their mistakes. Ask if they understand.*

It is best if you experience an example of this, so when ready, click the *Task B Example* button.

- *Participant completes Task B Example*

Good. Now you will perform a sequence of 3 practice trials of this task. During this sequence you will also need to respond to both types of interrupting tasks. An interruption may occur at any time during the trial. Upon completion of an interrupting task, you click to dismiss the interruption and return to the puzzle task trial which was interrupted, continuing exactly where you left off. You are encouraged to complete each trial both as quickly and as accurately as possible, without performing more than the specified number of moves. Recall that a complete solution contains no incomplete squares. When ready, click on *Task B Practice*.

- *Participant completes Task B Practice*

You will now complete the experimental trials, again please complete each trial as quickly and as accurately as possible. Remember that you cannot "undo" moves, so don't rely on trial and error; plan carefully. You will perform 3 sets of 8 trials, for a total of 24 trials. After each set, you will be given an opportunity to rest and you will be asked to complete a short survey. Once again, I will remind you that an interruption may occur at any time during the trial. Many trials will be uninterrupted. We are interested in your performance both on the puzzle tasks and on the interrupting tasks. In addition, recall that the interrupting tasks have been intentionally designed to distract you and keep you from thinking about the puzzle task, so don't allow them to frustrate you too much, and just do the best that you can.

- *Ensure that your marking sheet is ready and labeled with group and subject ID (i.e. g2s14). For each task B.{N}, where N = {1,2,3}:*

When ready, click on *Task B.{N}.*

- *Participant completes trial bank B.{N}.*

Please fill out this section of the questionnaire.

- *Label the header of the questionnaire accordingly. Participant fills out section of questionnaire relevant to B.{N}*

This concludes Task B. [*If Task B occurs first*: You have completed the first of two main tasks. We will now complete the second task. *Otherwise, stop recording in CamStudio and save the video file to My Documents/My Videos.*]

Conclusion

We're almost done. Just a couple of things left.

I'd like to ask you several questions about your experiences today.

- *Cue tape recorder if consent is given. Otherwise take notes. Use the printouts of the main tasks to serve as illustration.*

Q1: Overall, which of the two puzzle tasks did you find more challenging, regardless of whether or not they were interrupted? Why?

(If it is unclear as to whether the participant clicked randomly on n-Back):
Q2: Did you devote attention to the demanding interruption, as I asked? Or did you click randomly in an effort to maintain your place in the main puzzle task?

Q2: In terms of the interruptions, in which of the two puzzle tasks was your performance impacted more negatively by interruptions? Why?

Q3: When you were informed of an imminent interruption, there was a brief delay between the onset of the red flashing interruption notification and the beginning of the interruption task. Did you find the delay helpful? How So? (*If response is vague*): In other words, did you adopt any strategies for remembering your current state in the puzzle task during this delay?

Q3a: (*elaboration*). Did these strategies differ between the two puzzle tasks?

Q4: After completing an interruption and returning to the puzzle task, what strategy or strategies did you use to resume the puzzle task? For example, did you do anything in particular to remember where you had left off?

Q4a: (*elaboration*). Did these strategies differ between the two puzzle tasks?

Q4b: (*elaboration*). Was your strategy different dependent on whether the interruption was of the passive type versus the interruption was one in which you were required to monitor and click on the sequence of images?

- *Turn off tape recorder.*

Thank you. Now if you will sign this receipt, I can give you your compensation for participating.

- *Participant signs receipt and receives compensation.*

Thank you very much for your time.

- *Participant leaves*
- *Create a subfolder for your participant in the expt_logs folder in DropBox, name it according to group and subject ID (i.e. g2s14). Move the experiment logs files from My Documents to the new subfolder.*
- *Transfer audio file to My Documents / My Music; rename as group and subject ID (i.e. g2s14)*
- *Collect and organize scoring sheets, NASA-TLX survey sheets (make sure they are labeled for each condition), NAART scoring sheet, MoCA, and consent form; store in a folder labeled by group and subject ID (i.e. g2s14);*

C.5.2 Alternate Participant Instructions & Interview

Participants who did not meet the cutoff criteria on the MOCA and NAART tests completed a shorter version of the study. They were not informed that they were completing the shorter version. The following script is what was read to them.

Alternate Experiment Script

Alternate Introduction

Now I'll tell you about the study. Today, we are going to be working on puzzle tasks. These tasks involve moving shapes around the screen with the mouse. This experiment involves only the use of the left-mouse button. No keyboard use will be necessary. I will tell you more about this task shortly. You will perform a number of instances of this task, occurring one after the other, which we'll call trials.

During some trials, you will be interrupted. These interruptions will last for approximately 20 seconds, delaying your execution of the puzzle. During this time, you will be required to complete an interrupting task. We are interested in your performance both on the puzzle task and on the interrupting task.

There are several phases to this experiment. First, you will experience the interrupting tasks by themselves until you are familiar with them. Following this, you will practice completing several trials of the puzzle task. Then, you will then perform 3 sets of trials of the puzzle task. During these trials you must remain ready to respond to the interrupting tasks when they occur.

- *Open the BrainFreeze application. Group ID= X, subject id = X, interruption condition ordering = N-L-H, and main-task order = B-A on the splash page's form (only task B will be used). Click the "Full-Screen" button, then the "Next" button to enter the main experiment portal screen.*
  Alternate Interruption Tasks Examples

You will now perform the interrupting tasks by themselves. At the top left of the screen, you can see two buttons: *Interruption 1 Example* and *Interruption 2 Example*.

In *Interruption 1*, you will be shown an automated sequence of a dozen cartoon images. Your task is to watch these images passively. You do not need to click on the images.

- *Demonstrate printed example of Interruption 1*

  Images will appear in this box in the center of the screen. Afterwards, you will be asked to click in order to dismiss the interruption.

When ready, click on *Interruption 1 Example.*

- *Participant performs Interruption 1 Example.*

Good. Interruption 2 looks similar to Interruption 1, but your task is to pay close attention to the images, and to click on an image if you saw the same image two images back in the sequence. Allow me to explain.

- *Demonstrate 3 printout examples of Interruption 2*

  When the current image repeats what you saw two images ago, click inside the box where the image is shown. Otherwise, do not click. Here is an example sequence of images and the responses you would receive by clicking on them. You will be shown a green check if your click is the correct response. You will be shown a red 'X' if you make an incorrect response.

After 12 images have been displayed, you will be prompted to dismiss the interruption. When ready, click on *Interruption 2 Example*.

- *Participant performs Interruption 2 Example.*

I encourage you now to perform this example again to gain additional practice.

We are now ready to start the puzzle task.

Alternate Task B (Square Puzzles)

You have been selected to perform Task B (Other participants will perform Task A). In this puzzle task, you will move lines to create complete squares.

- *Demonstrate printed example of Task B, pointing out every relevant UI element (continue button, instruction text, etc.)*

You will see an instruction at the bottom of the screen. Move the lines around the white area by clicking on them and dragging them with the mouse. Lines cannot be rotated. When finished, click "Next" to advance to the next trial.

It is best if you experience an example of this, so when ready, click the *Task B Example* button.

- *Participant completes Task B Example*

Good. Now you will perform a sequence of 3 practice trials of this task. During this sequence you will also need to respond to both types of interrupting tasks. An interruption may occur at any time during the trial. Upon completion of an interrupting task, you will dismiss the interruption by clicking "Continue" and return to the primary task trial which was interrupted, continuing exactly where you left off. You are encouraged to complete each trial both as quickly and as accurately as possible. When ready, click on *Task B Practice*.

- *Participant completes Task B Practice*

You will now complete the experimental trials, again please complete each trial as quickly and as accurately as possible. You will perform 3 sets of 8 trials, for a total of 24 trials. After each set of trials, you will be given an opportunity to rest. Once again, I will remind you that an interruption may occur at any time during the trial. Many trials will be uninterrupted. In addition, recall that the interrupting tasks have been intentionally designed to distract you and keep you from thinking about the puzzle task, so don't allow them to frustrate you too much, and just do the best that you can.

- *Ensure that your marking sheet is ready. For each task B.{N}, where N = {1,2,3}:*

When ready, click on *Task B.{N}.*

- *Participant completes trial bank B.{N}.*

This concludes the experimental tasks.

Alternate Conclusion

We're almost done. Just a couple of things left.

I'd like to ask you a couple questions about your experiences today.

Q1: After completing an interruption and returning to the primary task, what strategy or strategies did you use to resume the puzzle task? For example, did you do anything in particular to remember where you had left off?

Q1b: (*elaboration*). Was your strategy different dependent on whether the interruption was of the passive type versus the interruption was one in which you were required to monitor and click on the sequence of images?

Thank you. Now if you will sign this receipt, I can give you your compensation for participating.

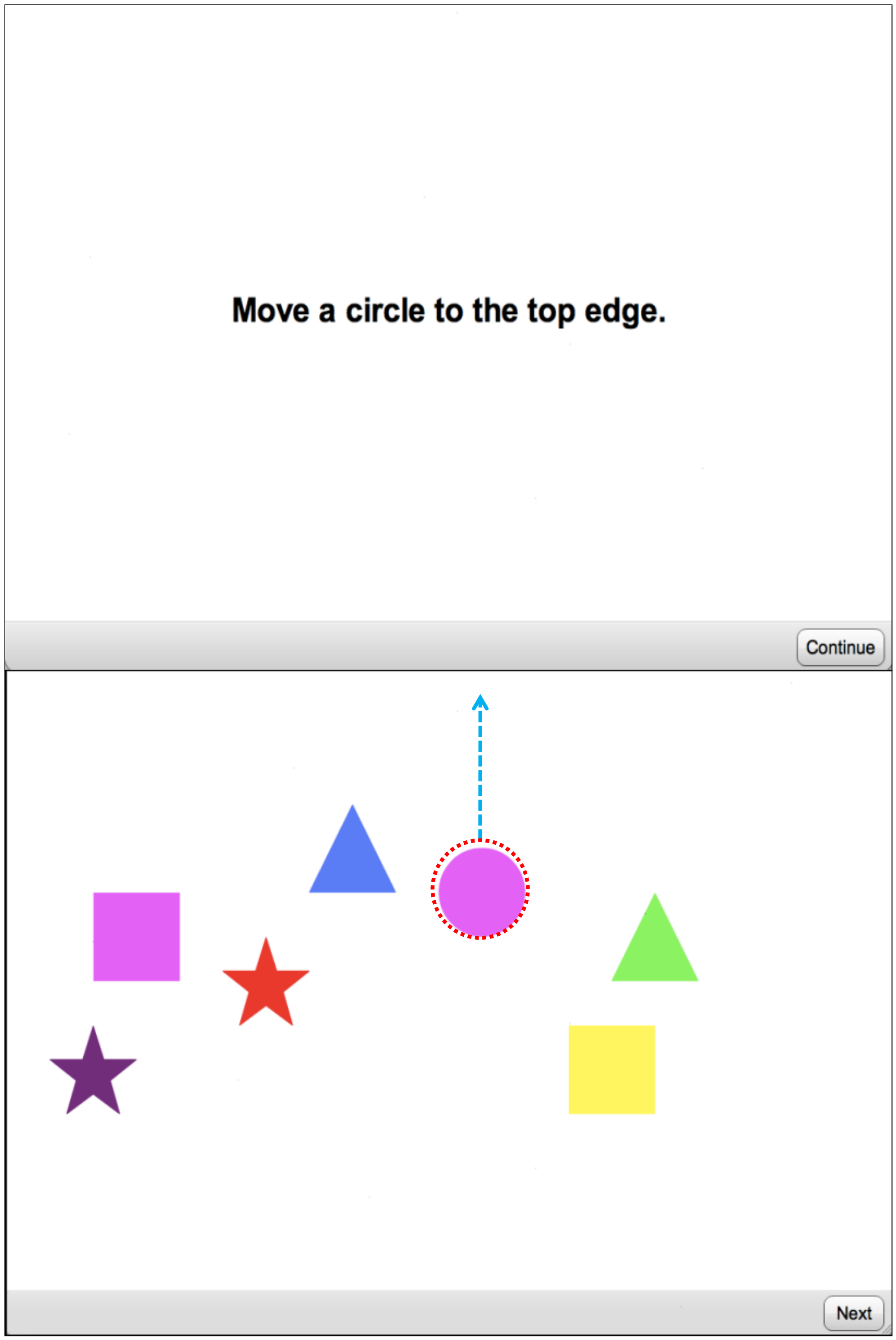
- *Participant signs receipt and receives compensation.*

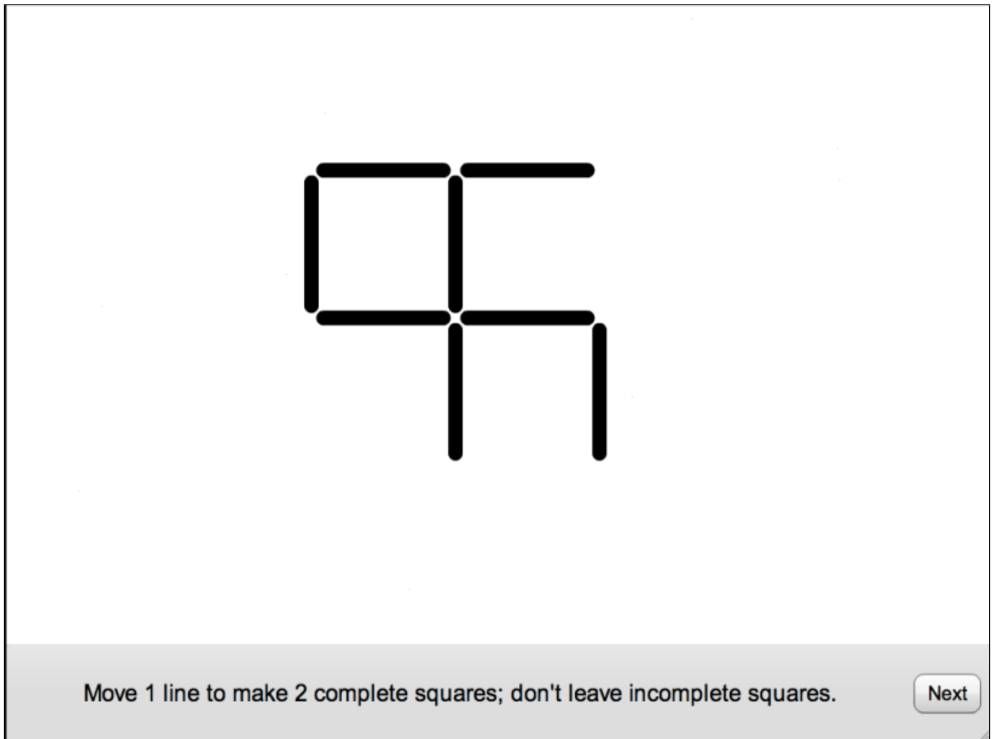
Thank you very much for your time.

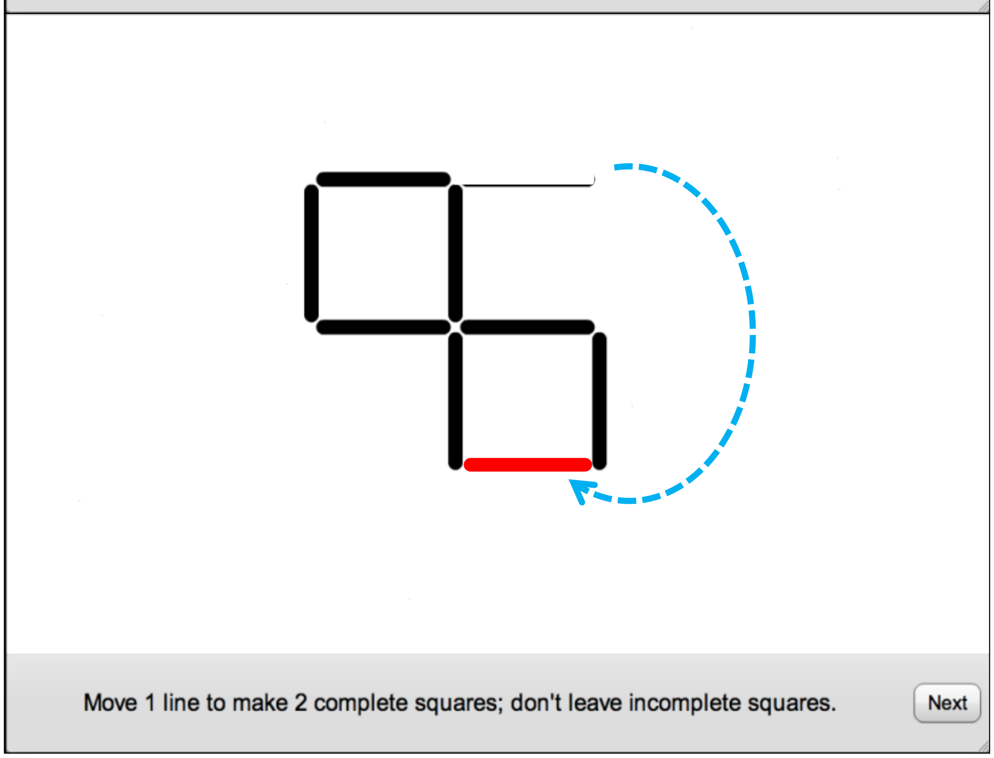
- *Participant leaves*
- *Create a subfolder for your participant in the expt_logs/excluded folder in DropBox, name it 'excluded' and include group and subject ID (i.e. excluded_g2s14). Move the experiment logs files from My Documents to the new subfolder.*

C.5.3 Visual Examples of Main & Interrupting Tasks

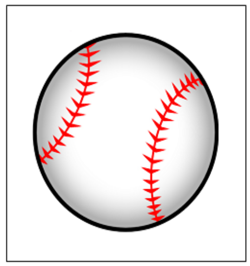
Where specified in the experiment script (Section C.5.1), participants were shown the following printed examples of the main (VERBAL, SPATIAL) and interrupting (PASSIVE, ACTIVE) tasks.


Move a circle to the top edge.
Continue
Next

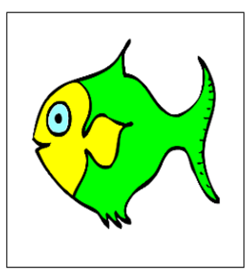

Move 1 line to make 2 complete squares; don't leave incomplete squares.
Next

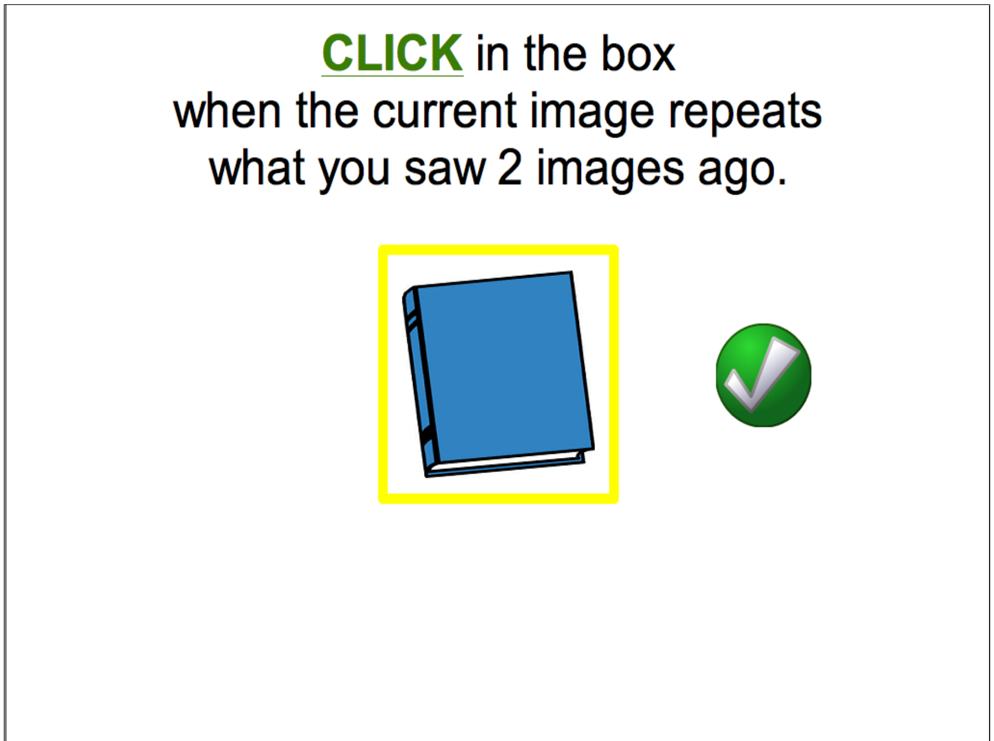

Move 1 line to make 2 complete squares; don't leave incomplete squares.
Next

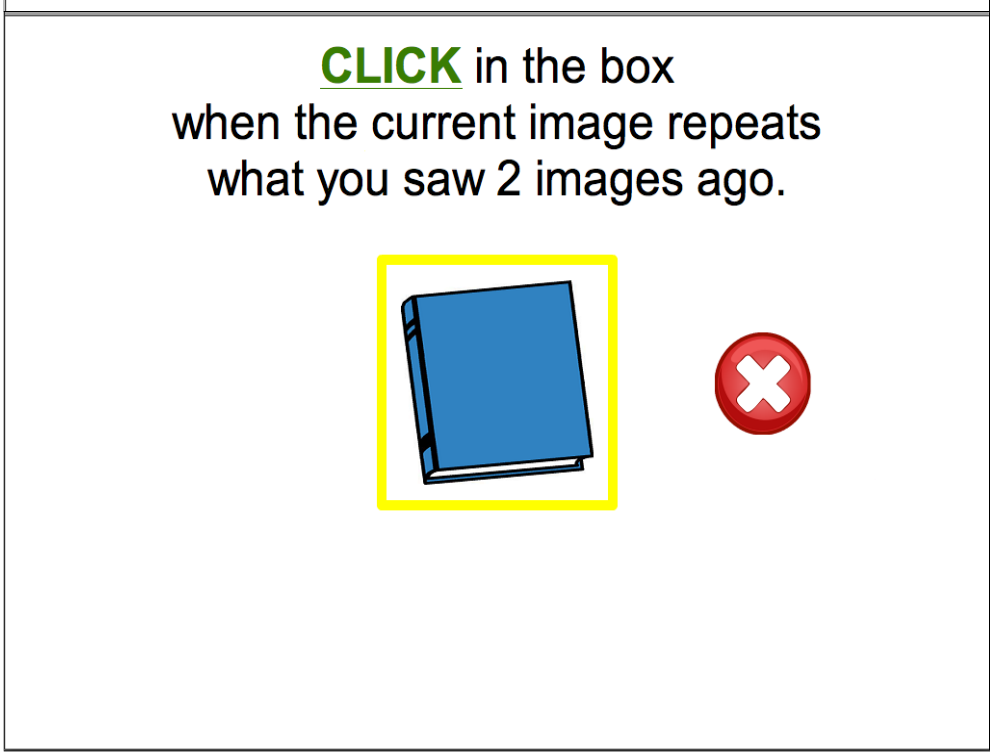
WATCH these images.
DO NOT CLICK on them.

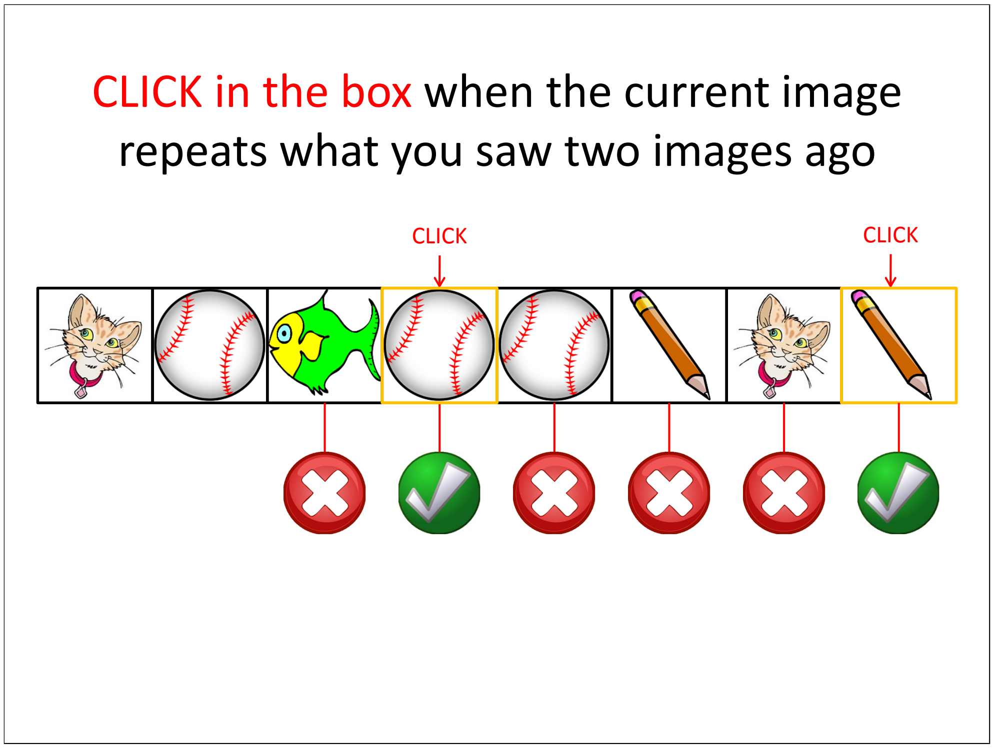


CLICK in the box
when the current image repeats
what you saw 2 images ago.




CLICK in the box
when the current image repeats
what you saw 2 images ago.


CLICK in the box
when the current image repeats
what you saw 2 images ago.


CLICK in the box when the current image repeats what you saw two images ago
CLICK
CLICK

C.6 Trial Instructions & Accuracy Scoring

C.6.1 Verbal Task

Trial Instructions

The 3 isomorphic trial blocks were presented in a random order to each participant. Corresponding trials in the each bank contained the same number of figures, however in unique initial configurations, and an isomorphic instruction, as denoted by the options presented in brackets in the following list:

1. "Move a [square / circle / triangle] to the [left / right / bottom] edge" – *(1 figure, 3 points)*

2. "Place the [blue / yellow] [triangle / square / circle] to the left of the [yellow / blue] [triangle / square / circle]" – *(2 figures, 6 points)*

3. "Cover the [yellow / blue / red] [triangle / circle] with the [red / blue] [circle / triangle]" – *(2 figures, 6 points)*

4. "Place the [orange / pink / red] [circle / square / star] below the [red / yellow] [circle / square / star]" – *(2 figures, 6 points)*

5. "Arrange figures of the same [shape / colour] in [vertical columns / horizontal rows]" – *(9 figures, 27 points)*

6. "Place the [blue / yellow / orange] [triangle / circle] in between the [pink / red / blue] [square / circle] and the [purple / orange / red] [star / square / triangle]" – *(3 figures, 9 points)*

7. "If there is a black [square / star], move all the figures to the [bottom / right / left] edge . Otherwise, move all the figures to the [top / left / right] edge" – *(7 figures, 21 points)*

8. "Line up all figures in order of size from smallest to largest, in a [horizontal row / vertical column] from [left to right / right to left / top to bottom]" – *(6 figures, 18 points)*

9. "Arrange all figures of the same [colour / shape] in [vertical columns / horizontal rows] . Then place the pink [square / rectangle] above the [column / row] with the [most / least] number of figures" – *(16 figures, 48 points)*

10. "Move the red figure to the [top / bottom] left , the blue figure to the top [right / left] , and the [green / yellow] figure to the [bottom / top] right" – *(3 figures, 9 points)*

Scoring Scheme

The VERBAL task had pre-defined scoring criteria, based on clinical scoring schemes and developed in consultation with Dr. Claudia Jacova. The scoring scheme was as follows:

1. 1 point awarded for moving the correct figure (when specified, 0.5 points for colour, 0.5 points for shape)

2. 1 point awarded for the correct action (i.e., cover, align horizontal, align vertical, move to edge / corner, place beside, between

3. 1 point awarded for correct relative position (i.e., on top, below, right, left, aligned by shape / colour, direction of alignment) or for satisfying a conditional instruction (i.e., if x, then y)

4. the total possible trial block score is 153 points

Scoring Sheets

The following 3 VERBAL task score sheets are unique, corresponding to the 3 isomorphic trial blocks.

p_id	SC		tokens	SC_A	instruction	comments on scoring
ord						
cond						
1	3		1		*Move a* ***square*** *to the* ***left*** *edge*	
2	6		2		*Place the* ***blue triangle*** *to the* ***left*** *of the* ***yellow triangle***	
3	6		2		***Cover*** *the* ***yellow triangle*** *with the* ***red circle***	
4	6		2		*Place the* ***orange circle below*** *the* ***red circle***	
5	27		9		*Arrange figures of the* ***same shape*** *in* ***vertical columns***	
6	9		3		*Place the* ***blue triangle*** *in* ***between*** *the* ***pink square*** *and the* ***purple star***	
7	21		7		***If*** *there is a* ***black square*** *,* ***move all*** *the figures to the* ***bottom edge*** *. Otherwise,* ***move all*** *the figures to the* ***top edge*** *.*	
8	18		6		***Line up all figures*** *in order of size from* ***smallest*** *to* ***largest,*** *in a* ***horizontal row*** *from* ***left*** *to* ***right***	
9	48		16		*Arrange* ***all figures*** *of the same* ***colour*** *in* ***vertical columns*** *. Then place the* ***pink square above*** *the column with the* ***most number*** *of figures*	
10	9		3		*Move the* ***red figure*** *to the* ***top left*** *, the* ***blue figure*** *to the* ***top right*** *, and the* ***green figure*** *to the* ***bottom right***	
Total	153					

p_id	SC		tokens	SC_B	instruction	comments on scoring
ord						
cond						
1	3		1		*Move a* ***circle*** *to the* ***right edge***	
2	6		2		*Place the* ***yellow square*** *to the* ***right*** *of the* ***blue square***	
3	6		2		*Cover the* ***blue circle*** *with the* ***red triangle***	
4	6		2		*Place the* ***pink square above*** *the* ***yellow square***	
5	27		9		*Arrange figures of the* ***same colour*** *in* ***vertical columns***	
6	9		3		*Place the* ***yellow circle*** *in* ***between*** *the* ***red circle*** *and the* ***orange square***	
7	21		7		***If*** *there is a* ***black square*** *, move* ***all the figures*** *to the* ***right*** *edge. Otherwise, move a* ***ll the figures*** *to the* ***left*** *edge*	
8	18		6		***Line up all figures*** *in order of size from* ***smallest*** *to* ***largest,*** *in a* ***horizontal row*** *from* ***right*** *to* ***left***	
9	48		16		*Arrange figures of the* ***same shape*** *in v* ***ertical column*** *s. Then place the* ***pink rectangle above*** *the column with the* ***most number*** *of figures*	
10	9		3		*Move the* ***red figure*** *to the* ***top left*** *, the* ***blue figure*** *to the* ***top right*** *, and the* ***green figure*** *to the* ***bottom right***	
Total	153					

p_id	SC		tokens	SC_C	instruction	comments on scoring
ord						
cond						
1	3		1		Move a **triangle** to the **bottom** edge	
2	6		2		Place the **yellow circle** to the **right** of the **blue circle**	
3	6		2		**Cover** the **red circle** with the **blue triangle**	
4	6		2		Place the **red star below** the **yellow star**	
5	27		9		Arrange figures of the **same shape** in **horizontal rows**	
6	9		3		Place the **orange circle** in **between** the **blue circle** and the **red triangle**	
7	21		7		**If** there is a **black star,** move **all** the figures to the **left edge.** Otherwise, move **all** the figures to the **right edge**	
8	18		6		**Line up all figures** in order of size from **smallest** to **largest,** in a **verical column** from **top** to **bottom**	
9	48		16		Arrange figures of the **same shape** in **horizontal rows**. Then place the **pink rectangle beside** the **smallest row** of figures	
10	9		3		Move the **red figure** to the **bottom left,** the **blue figure** to the **top left,** and the **yellow figure** to the **top right**	
Total	153					

C.6.2 Spatial Task

Trial Instructions

The trial blocks were presented in a random order to each participant. Corresponding trials in the each bank contained the same number of lines, however appearing in rotated initial configurations, and an identical instruction, as follows:

Trial 1 – "Move 1 line to make 2 complete squares; don't leave incomplete squares" – *(2 squares, 4 points)*

Trial 2-5 – "Move 2 lines to make 3 complete squares; don't leave incomplete squares" – *(3 squares, 6 points)*

Trial 6 – "Move 3 line to make a number of complete squares; don't leave incomplete squares" – *(4 squares, 8 points)*

Trial 7-8 – "Move 3 line to make a number of complete squares; don't leave incomplete squares" – *(5 squares, 10 points)*

Scoring Scheme

The SPATIAL task had pre-defined scoring criteria, based on clinical scoring schemes and developed in consultation with Dr. Claudia Jacova. The scoring scheme was as follows:

1. 2 points awarded for every complete square
2. 1 point deducted for every incomplete square
3. 1 point deducted for every additional move (after the specified number of moves are completed)
4. the minimum score is 0 points
5. the total possible trial block score is 56 points

Scoring Sheets

The following SPATIAL task score sheet was used for the 3 isomorphic trial blocks, as the instructions were identical for each block.

p_id	SQP points	mvs	SQP1	comments on scoring
bank				
cond				
1	4	1		
2	6	2		
3	6	2		
4	6	2		
5	6	2		
6	8	3		
7	10	3		
8	10	3		
Total	56			

C.7 Study Questionnaire

The questionnaire was adapted from the NASA-TLX, a standardised instrument for assessing various dimensions of workload [23]. Six questions were posed regarding mental and physical demand, annoyance, perceived performance, and fatigue; responses were along a 10-point scale.

Participants completed this questionnaire after each condition, for both primary tasks (a total of 6 times).

DO NOT WRITE IN THIS SECTION (Experimenter Use Only)
Group ID: ____ Subject ID: ____ Task: ____ Bank: ____ Condition: ____ Order: ____

Survey

With respect to your **<u>how you feel right now</u>**, please answer the following question by marking an 'X' along the scale beside the corresponding question.

How fatigued are you feeling at this time?

FATIGUE

Low High

With respect to **<u>the last set of trials</u>**, including the primary puzzle task and any interrupting secondary tasks that may have occurred, please answer the following questions by marking an 'X' along the scale beside the corresponding question.

How much mental activity was required to perform the **primary puzzle task during this set of trials** (e.g. thinking, remembering, looking, searching, deciding, etc.)?

MENTAL DEMAND

Low High

How **annoyed** (i.e. pestered, harassed, disturbed, or irritated) were you **during this set of trials in general**?

GENERAL ANNOYANCE

Low High

IF INTERRUPTIONS OCCURRED during this set of trials, how annoyed (i.e. pestered, harassed, disturbed, or irritated) were you **by the interruptions**? Otherwise, do not answer.

INTERRUPTION ANNOYANCE

Low High

How **successful** do you think you were in accomplishing the goals of this set of trials?

PERFORMANCE

Low High

How much **physical activity** was required to perform **this set of trials**? (e.g. moving the mouse, clicking the mouse button, etc.)

PHYSICAL DEMAND

Low High