BMC Psychology

BioMed Central
The Open Access Publisher

Are women better than men at multi-tasking?

BMC Psychology 2013, **1**:18 doi:10.1186/2050-7283-1-18

Gijsbert Stoet (gijsbert.stoet@glasgow.ac.uk)
Daryl B O'Connor (d.b.o'connor@leeds.ac.uk)
Mark Conner (m.t.conner@leeds.ac.uk)
Keith R Laws (k.laws@herts.ac.uk)

ISSN 2050-7283

Article type Research article

Submission date 3 January 2013

Acceptance date 28 August 2013

Publication date 24 October 2013

Article URL http://www.biomedcentral.com/2050-7283/1/18

Are women better than men at multi-tasking?

Gijsbert Stoet[1*]
[*]Corresponding author
Email: gijsbert.stoet@glasgow.ac.uk

Daryl B O'Connor[2]
Email: d.b.o'connor@leeds.ac.uk

Mark Conner[2]
Email: m.t.conner@leeds.ac.uk

Keith R Laws[3]
Email: k.laws@herts.ac.uk

[1]School of Education, University of Glasgow, Glasgow, Scotland, UK

[2]Institute of Psychological Sciences, University of Leeds, Leeds, West Yorkshire, UK

[3]School of Life and Medical Sciences, University of Hertfordshire, Hatfield, Hertfordshire, UK

Abstract

Background

Their seems to be a common belief that women are better in multi-tasking than men, but there is practically no scientific research on this topic. Here, we tested whether women have better multi-tasking skills than men.

Methods

In Experiment 1, we compared performance of 120 women and 120 men in a computer-based task-switching paradigm. In Experiment 2, we compared a different group of 47 women and 47 men on "paper-and-pencil" multi-tasking tests.

Results

In Experiment 1, both men and women performed more slowly when two tasks were rapidly interleaved than when the two tasks were performed separately. Importantly, this slow down was significantly larger in the male participants (Cohen's $d = 0.27$). In an everyday multi-tasking scenario (Experiment 2), men and women did not differ significantly at solving simple arithmetic problems, searching for restaurants on a map, or answering general knowledge questions on the phone, but women were significantly better at devising strategies for locating a lost key (Cohen's $d = 0.49$).

Conclusions

Women outperform men in these multi-tasking paradigms, but the near lack of empirical studies on gender differences in multitasking should caution against making strong generalisations. Instead, we hope that other researchers will aim to replicate and elaborate on our findings.

Background

In the current study, we address the question whether women are better multi-taskers than men. The idea that women are better multi-taskers than men is commonly held by lay people (for a review see Mäntylä 2013). While the empirical evidence for women outperforming men in multi-tasking has been sparse, researchers have shown that women are involved *more* in multi-tasking than men, for example in house-hold tasks (Offer and Schneider 2011; Sayer 2007). In this paper we address the question if it is true that women actually *outperform* men when multi-tasking.

Multi-tasking is a relatively broad concept in psychology, developed over several decades of research (for a review see Salvucci and Taatgen 2010); this research has enormous relevance for understanding the risk of multi-tasking in real-life situations, such as driving while using a mobile phone (Watson and Strayer 2010).

There are at least two distinct types of multi-tasking abilities. The first type is the skill of being able to deal with multiple task demands *without* the need to carry out the involved tasks simultaneously. A good example of this type of multi-tasking is carried out by administrative assistants, who answer phone calls, fill in paperwork, sort incoming faxes and mail, and typically do not carry out any of these tasks simultaneously.

A second type of multi-tasking ability is required when two types of information must be processed or carried out *simultaneously*. An example of the latter category is drawing a circle with one hand while drawing a straight line with the other hand. While humans have no difficulty carrying out each of these tasks individually, drawing a circle with one hand and drawing a *straight* line with the other simultaneously is nearly impossible (the circle becomes more of an ellipse and the line more of a circle, Franz et al. 1991). Another example is the requirement to process different types of sensory information at the same time (Pashler 1984), such as different auditory streams on different ears (Broadbent 1952). While humans frequently are asked to do such tasks in the psychological laboratory, humans seem to try to avoid these situations in real life, unless they are highly trained (e.g., playing piano, with the left and right hands playing different notes, or having a conversation while driving a car). Arguably, we are not good at doing multiple tasks simultaneously (except when well trained), and that probably explains why this type of multi-tasking is less common than the type in which we serially alternate between two tasks (Burgess 2000). It is because of this that we focus on the first type of multi-tasking in this study. Also, it is important to note that the two types of multi-tasking described above are two extreme examples on a continuum of multi-tasking scenario.

Cognitive scientists and psychiatrists have postulated a special set of cognitive functions that help with the coordination of multiple thought processes, which include the skills necessary for multi-tasking, namely "executive functions" (Royall et al. 2002): task planning, postponing tasks depending on urgency and needs (i.e., scheduling), and ignoring task-irrelevant information (also known as "inhibition"). Healthy adults can reasonably well interleave two novel tasks rapidly (Vandierendonck et al. 2010). The involved (human) brain areas necessary for multi-tasking have been investigated and we can at the very least make a reasonable estimate of which are involved (Burgess et al. 2000). Among primates, humans seem to have a unique way of dealing with task switching (Stoet and Snyder 2003), which we hypothesize reflects an evolutionary unique solution for dealing with the advantages and disadvantages of multi-tasking (Stoet and Snyder 2012). The specific contributions of individual brain areas to executive control skills in humans have been linked to a number of mental disorders, in particular schizophrenia (Evans et al. 1997; Kravariti et al. 2005; Royall et al. 2002; Semkovska et al. 2004; Dibben et al. 2008; Hill et al. 2004; Laws 1999).

Currently, there are few studies on gender and multi-tasking, despite a seemingly confident public opinion that women are better in multi-tasking than men (Ren et al. 2009). Ren and colleagues (2009)

extrapolated the hunter-gatherer hypothesis (Silverman and Eals 1992) to make predictions about male and female multi-tasking skills. The hunter-gatherer hypothesis proposes that men and women have cognitively adapted to a division of labor between the sexes (i.e., men are optimized for hunting, and women are optimized for gathering). Ren and colleagues speculated that women's gathering needed to be combined with looking after children, which possibly requires more multi- tasking than doing a task without having to look after your offspring. In their experiment, men and women performed an Eriksen flanker task (Eriksen and Eriksen 1974) either on its own (i.e., single task condition) or preceded by an unrelated other cognitive decision making task (i.e., multi-tasking condition). They found that in the multi-tasking condition, women were less affected by the task-irrelevant flankers than men. Thus, the latter study supports the hypothesis that women are better multi-taskers.

We tested whether women outperform men in the first type of multi-tasking. In Experiment 1, we tested whether women perform better than men in a computer-based task-switching paradigm. In Experiment 2[a], we tested whether women outperform men in a task designed to test "planning" in a "real-life" context that included standardized tests of executive control functions. Our prediction was that women would outperform men.

Experiment 1

In this experiment, we used a task-switching paradigm to measure task-switching abilities. Task-switching paradigms are designed to measure the difficulty of rapidly switching attention between two (or more) tasks. Typically, in these types of studies, performing a task consists of a simple response (e.g., button press with left or right hand) to a simple stimulus (e.g., a digit) according to simple rules (e.g., odd digits require left hand response, even digits a right hand response).

In task-switching paradigms, there are usualy two different tasks (e.g., in task A deciding whether digits are odd or even, and in task B deciding whether digits are lower or higher than the value 5). An easy way to think of task-switching paradigms is to call one task "A" and another task "B". A block of just ten trials of task A can be written as "AAAAAAAAAA" and a block of just ten trials of task B can be written as "BBBBBBBBBB". Most adults find carrying out sequences of one task type relatively simple. In contrast, interleaving trials like "AABBAABBAABB" is difficult, as demonstrated for the first time in 1927 by Jersild (1927). Today, the slowing down associated with carrying out a block of mixed trials compared to a block of pure trials is known as "mixing cost". Further, within mixed blocks, people slow down particularly on trials that immediately follow a task switch (in AA**B**B**A**A there are two such trials, here indicated in bold font); the latter effect is known as "switch cost".

Researchers have given switch costs more attention than mixing costs, especially since the mid-1990s (Vandierendonck et al. 2010)[b]. In the current experiment, we measured both types of costs.

Methods

Participants

We recruited participants via online advertisements and fliers in West Yorkshire (UK). Our recruitment procedure excluded participants with health problems and disorders that could potentially affect their performance, which included color-vision deficits, as tested with the Ishihara color test (Ishihara 1998) before each experimental session. Altogether, we selected 240 participants stratified by gender and age (Figure 1).

Figure 1 The distribution of participants by gender and age. The average age of women was 27.4 years (*SD* = 6.0); the average age of men was 27.8 years (*SD* = 6.4).

Research ethics

Research was in accordance with the declaration of Helsinki, and approval of ethical standards for Experiment 1 was given by the ethics committee of the Institute of Psychological Sciences, University of Leeds. All participants gave written or verbal consent to participate.

Apparatus and stimuli

The experiment was controlled by a Linux operated PC using PsyToolkit software (Stoet 2010). A 17" color monitor and a Cedrus USB keyboard (model RB-834) were used for stimulus presentation and response registration, respectively. Of the Cedrus keyboard, only two buttons were used. These were the buttons closest to the participant (3.2 $\times$ 2.2 cm each, with 4.3 cm between the two buttons), which we will further refer to as the left and right button, respectively.

A rectangular frame (7 $\times$ 8 cm) with an upper and lower section (Figure 2a) was displayed. The words "shape" and "filling" were presented above and below the frame, respectively. Further four imperative stimuli were used in different trials (Figure 2b). These four were the combination of two shapes (diamond and rectangle) and a filling of two or three circles. The frame and the imperative stimuli were yellow and were presented on a black background. Feedback messages were presented following trials that were not performed correctly ("Time is up" or "That was the wrong key").

Figure 2 Schematic representation of the task-switching paradigm. **A**: Example trial. During a block of trials, a rectangular frame with the labels "shape" and "filling" was visible. On each trial, a different imperative stimulus (i.e., a stimulus that requires an immediate response) was presented in the top or bottom part of this frame. The location (i.e., in top or bottom part of frame) determined whether the participant had to apply the shape or filling task rules to it. **B**: There were four different imperative stimuli, which needed to be responded to as follows. In the shape task, a "diamond" required a left-button response, and a rectangle a right-button response. In the filling task, a filling of two circles required a left-button response, and a filling of three circles a right-button response. Congruent stimuli are those that required the same response in both tasks, whereas incongruent stimuli required opposite responses in the two tasks. Thus, the imperative stimulus in panel A is incongruent: It appears in the top of the frame, thus is should be responded to in accordance to the shape task, and because it is a diamond (the filling of three circles is irrelevant in the shape task) it should be responded to with a left-button response (see Additional file 1 for demonstration).

Procedure

Participants were seated in a quiet and dimly lit room, and received written and verbal instructions from the experimenter. They were instructed to respond to stimuli on the computer screen. There were two different tasks, namely a shape and a filling task. In the shape task, participants had to respond to the shape of imperative stimuli (diamonds and rectangles required a left and right response, respectively). In the filling task, participants had to respond to the number of circles within the shape (two and three circles required a left and right response, respectively). The essential feature of this procedure was that both task dimensions (shape and filling) were always present and that the two dimensions required opposite responses on half the trials (incongruent stimuli). This meant that participants were forced to think of which of the two tasks needed to be carried out and to attend to the relevant stimulus dimension.

Participants were informed which task to carry out based on the imperative stimulus location: If the stimulus appeared in the upper half of the frame, labeled "shape", they had to carry out the shape task, and when it appeared in the bottom half of the frame, labeled "filling", they had to carry out the filling task.

Participants first went through 3 training blocks (40 trials), and then performed 3 further blocks (192 trials total) that were used in the data analysis. The first two blocks were blocks with just one of the two tasks (pure blocks), and in the third block the two tasks were randomly interleaved (mixed block). In the mixed block, task-switch trials were those following a trial of the alternative task, and task-repeat trials were those following the same task. The order of blocks was identical for all participants. The computer used a randomisation function to choose which task would occur on a given trial. Further, it is important to note that participants had training in both tasks before the blocks started that were used for data analysis; this means that even in the first pure block of the analyzed data, participants were aware that incongruent stimuli were associated with opposite responses in the alternative task.

In each trial, the frame and its labels (as displayed in Figure 2a) were visible throughout the blocks. When an imperative stimulus (one of the four shown in Figure 2b) appeared (they were chosen at random by the software), participants had up to 4 seconds to respond. The imperative stimulus disappeared following a response or following the 4 seconds in case no response was given. Incorrect responses (or failures to respond) were followed by a 5 seconds lasting reminder of the stimulus-response mapping, and then followed by a 500 ms pause. The intertrial interval lasted 800 ms. A demonstration of the task is available in the Additional file 1.

When we report response times in task switching trials or in pure blocks, we always report the average of both tasks. For example, when reporting the response times in the pure blocks, we will report the average of the pure block of the shape task and pure block of the filling task.

Results

Response time analyses were based on response times in correct trials following at least one other correct trial. Further, we excluded all participants who performed not significantly different from chance level in all conditions. This exclusion is necessary, given that response time analyses in cognitive psychology are based on the assumption that response times reflect decision time. When participants guess, for example because they find the task difficult, the response times are no longer informative of their decision time.

The procedure for testing if participants performed better than chance was carried out as follows. Given that there were only two equally likely response alternatives on each trial, participants had 50% chance to get a response correct. To determine if a participant performed significantly better than chance level, we applied a binomial test to the error rates in each condition. Based on this analysis, we concluded that nine participants (5 men and 4 women, aged 18-36) did not perform better than chance in at least one experimental condition. We found that each of these nine participants worked at chance level in the incongruent task-switching condition (with error rates ranging from 29% to 60%), and for five of them, this was the only condition they failed in. None of these nine failed in the pure task blocks. We excluded these participants from all reported analyses.

The next set of analyses were carried out to confirm that the used paradigm showed the typical effects of task-switching and task-mixing paradigms as described in the introduction (Figure 3). Throughout, we only report statistically significant effects (α criterion of .05).

We analyzed task-switch and incongruency costs in response times in the mixed blocks. We carried out a mixed-design ANOVA with the within-subject factors "switching" and "congruency" and between-subject factor "gender". We found a significant effect of switching, $F(1,229) = 743.90$, $p < .001$: Par-

Figure 3 The response times and error rates + 1 standard error of the mean in the pure, task-switching and task-mixing conditions. Further, data is split up for congruent and incongruent stimuli, and for men and women.

ticipants responded 247 $\pm$ 9 ms more slowly in the task-switch (1010 $\pm$ 14 ms) than in the task-repeat (763 $\pm$ 10) condition[c]. Further, participants were 35 $\pm$ 5 ms slower in incongruent (904 $\pm$ 11 ms) than in congruent (869 $\pm$ 11 ms) trials, $F(1,229) = 52.48$, $p < .001$.

We repeated the same analysis on the error rates. Again, we found a significant effect of switching, $F(1,229) = 53.20$, $p<.001$, with people making 1.97 $\pm$ 0.27 error percentage points (ppt) more in the task-switch (4.62 $\pm$ 0.27%) than in the task-repeat (2.65 $\pm$ 0.18%) condition. Further, people made 3.77 $\pm$ 0.31 ppt more errors in incongruent (5.52 $\pm$ 0.30%) than in congruent (1.75 $\pm$ 0.18%) trials,$F(1,229) = 143.90$, $p < .001$. Finally, the interaction between switching and congruency was significant, $F(1,229) = 14.65$, $p < .001$.

Next, we analyzed task-mixing costs using a similar approach as above. Now, we contrasted trials in the pure blocks with task-repeat trials in mixed block. We observed a slow down of 319 $\pm$ 8 ms due to mixing, $F(1,229) = 1555.34$, $p<.001$, with an average response time in mixed trials of 763 $\pm$ 10 ms and in pure trials of 444 $\pm$ 5 ms. This effect interacted significantly with the gender of participants. The slow down due to mixing was 336 $\pm$ 11 ms in men and 302 $\pm$ 12 ms in women (the effect size of this gender difference expressed as Cohen's $d = 0.27$). We also found an effect of congruency, $F(1,229) = 24.46$, $p<.001$, with people responding 18 $\pm$ 4 ms slower in incongruent (613 $\pm$ 7 ms) than congruent (594 $\pm$ 7 ms) trials. Finally, there was a significant interaction between mixing and congruency, $F(1,229) = 10.37$, $p = .001$.

We carried out the same analysis using error rate as dependent variable, and we found a significant effect of task-mixing again. People made 0.55 ppt more errors in the task mix condition (2.65 $\pm$ 0.18%) than in the pure condition (2.10 $\pm$ 0.13%), $F(1,229) = 9.17$, $p = .003$. People made 1.77 $\pm$ 0.20 ppt more mistakes in the incongruent (3.26 $\pm$ 0.19%) than in the congruent (1.49 $\pm$ 0.13%) condition, $F(1,229) = 80.86$, $p < .001$. The factors switching and congruency interacted, $F(1,229) = 26.94$, $p<.001$. In the error rates, there were no effects of gender. Even so, it might be of interest to report that women's mixing cost in error rates was 0.50 $\pm$ 0.28 percentage points and that of men 0.60 $\pm$ 0.23 percentage points.

Altogether, the ANOVAs of task-switching, task-mixing, and congruency confirmed the well known picture of task-switching data. The novelty is the gender difference in task-mixing costs. Although men and women did not show an overall speed difference, we wanted to ensure that the gender difference was not simply related to overall speed (e.g., people with larger switch costs might also have had a different baseline speed). To do so, we analyzed relative mixing costs as well. Relative mixing costs is the percentage slowing down in mixed compared to pure task blocks. For example, if a person responds on average in 500 ms in mixing blocks and 400 ms in pure blocks the person gets 25% slower due to mixing tasks.

We found that when analyzing the relative slow down due to mixing in relationship to performance in pure blocks, there was a significant effect of gender. Women's relative slow down (69.1 $\pm$ 2.6%) was, in correspondence to the ANOVA of the absolute response time, less than that of men (77.2 $\pm$ 2.6%), $t(229)=2.18$, $p = .030$; in other words, both the analysis of absolute and relative mixing costs show the same phenomenon.

Experiment 2

In Experiment 1, we found that men's and women's performance differed in a computer-based task measuring the capacity to rapidly switch between different tasks. One of the difficulties with computer-based laboratory tasks is their limited ecological validity. Experiment 2 aimed to create a multi-tasking situation in a "real-life" context that included standardized neurocognitive tests.

The approach of this experiment is based on tasks common in cognitive neuropsychology. From a neuropsychological perspective, Burgess (Burgess et al. 2000) described multi-tasking as the ability to manage different tasks with different (sometimes unpredictable) priorities that are initiated and monitored in parallel. Furthermore, goals, time, and other task constraints are seen as self defined and flexible. Shallice and Burgess (Shallice and Burgess 1991) devised the Six Elements Test to assess precisely these abilities (later modified by others, Wilson et al. 1998). In this task, participants receive instructions to do three tasks (simple picture naming, simple arithmetic and dictation), each of which has two sections, A and B. The subject has 10 minutes to attempt at least part of each of the six sections, with the proviso that they cannot do sections A and B of the same task after each other.

Burgess and colleagues (Burgess 2000; Burgess et al. 2000) have highlighted various features of multi-tasking behaviour, including: (1) several discrete tasks to complete; (2) interleaving required for effective dovetailing of task performance; (3) performing only one task at a particular time; (4) unforeseen interruptions; (5) delayed intentions for the individual to return to a task which is already running; (6) tasks that demand different task characteristics (7) self-determining targets with which the individual decides for him/herself; and (8) no minute-by-minute feedback on how well an individual performs. As Burgess and colleagues note, most laboratory-based tasks do not include all of these features when assessing multi-tasking. If this is indeed the case, there is a real advantage in studying multi-tasking using this approach.

Methods

Participants

We recruited 47 male and 47 female participants, largely undergraduate students of Hertfordshire University. The mean age was 24.2 years ($SD = 8.1$, range 18–60) for men, and 22.6 years ($SD = 5.6$, range 18–49) for women; there was no significant age difference between these two groups, $t(92) = 1.1$, $p = .28$.

Research Ethics

Research was in accordance with the declaration of Helsinki, and approval of ethical standards for Experiment 2 was given by the ethics committee of the School of Life and Medical Sciences, University of Hertfordshire. All participants gave written or verbal consent to participate.

Material

We used three different tasks. The "Key Search task" was taken from the Behavioral Assessment for Dysexecutive Syndrome (BADS, Wilson et al. 1998). This is a specific test of planning and strategy, in which participants are required to sketch out how they might route an attempt to search a "field" for a missing set of keys. This task is normally used as a measure of problems in executive function, and low scores are indicative of frontal lobe impairment. In the healthy population, this task reveals no evidence of a gender difference according to test norms and personal communication with Jon Evans (one of the

test designers). The test designers reported a high (r = .99) correlation between raters (Wilson et al. 1998).

The Map search task was taken from the "Tests of Everyday Attention" (Robertson et al. 1994). The task requires individuals to find restaurant symbols on an unfamiliar color map of Philadelphia (USA) and its surrounding areas. Again, this task reveals no evidence of a gender difference according to the test norms and personal communication with test designer Ian Robertson.

The third task was custom designed and involved solving simple arithmetical questions presented on paper as shown in Figure 4. We did pilot these mathematics questions (unlike the first two tests, this test is not standardised, and after piloting we moderated these questions to make sure they could be largely successfully attempted while doing the other tasks).

Figure 4 Example of the arithmetic task.

Although there are reports that men outperform women on more complex mathematics problems, this is typically not the case for simple calculations like this (Halpern et al. 2007).

A scoring system established within the BADS marks these plans according to set rules such as parallel patterns and corner entry. A panel of 3 scorers agreed on the scores for each test to ensure reliable scoring. Examples of key search strategies are shown in Figure 5.

Figure 5 Examples of the key search task. The example on the left is from a male participant, the example on the right from a female participant.

Procedure

Each participant was given 8 minutes to attempt the three tasks described above (Arithmetic, Map, Key Search). The layout of the position of the map task, maths task and key search was counterbalanced to avoid any bias affecting which tasks participants chose to do. They were instructed that each task held equal marks; it was left to participants to decide how they would organize their time between each task. The participants were also informed that they would receive a phone call at some unknown time point (always after 4 minutes) asking them 8 simple general-knowledge questions (e.g., "What is the capital of France"), it was again left to participants to decide whether or not they answered the phone call. Without or with answering the phone call, they were multi-tasking; answering the call just added to that multi-tasking 'burden' as such. If they attempted to multi-task while answering the phone call, this was recorded. We recorded time spent on each task as well as performance.

Results

We compared test scores (Table 1) and response times (Table 2) of men and women using t tests. We found that women (10.26 $\pm$ 0.58) scored significantly higher than men (8.13 $\pm$ 0.68) on the key search task. Importantly, this finding cannot simply be explained as a preference difference for the speed with which the task was carried out, as no response time differences were found (Table 2).

No differences emerged in the numbers of men and women who answered the phone (79% of men and 81% of women, $\chi^2(1) = 0.06$, p = .80). Those who answered the phone heard 8 simple general knowledge questions and the correct answers did not differ between men (3.35 $\pm$ 0.35) and women (3.84 $\pm$ 0.34), $t(73) = 1.0$, p = .32; nor did time spent on the phone differ between men (97.68 $\pm$ 3.13 seconds) and women (106.87 $\pm$ 3.65 seconds), $t(73) = 1.91$, p = .06. Of those that did answer the phone,

Table 1 Scores of men and women in Experiment 2

Task	Men	Women	*t* test	*p* value	Cohen's *d*
Arithmetic correct	19.68 (1.07)	17.29 (1.08)	1.57	.12	0.33
Map task (% correct)	75 (3.82)	72.00 (3.72)	0.52	.60	0.11
Key search score	8.13 (0.68)	10.26 (0.58)	5.6	.02	0.49

Standard errors in parentheses

Table 2 Response times (RT, seconds) of men and women in Experiment 2

Task	Men	Women	*t* test	*p* value	Cohen's *d*
Arithmetic	312 (13)	341 (17)	1.33	.19	0.28
Map task	160 (16)	180 (14)	0.91	.37	0.19
Key search	36 (4)	36 (5)	0.03	.98	0.01

Standard errors in parentheses. The sum of the three individual tasks exceeds the 480 allocated seconds, because sometimes the participants carried these tasks out concurrently and so were double scored on time.

we also measured whether they actively multi-tasked while on the phone or concentrated purely on this phone - and there was no significant difference 73% of men and 84% of women multi-tasked, $\chi^2(1) = 1.41$, $p = .24$.

Discussion

Using two very different experimental paradigms, we found that women have an advantage over men in specific aspects of multi-tasking situations. In Experiment 1, we measured response speed of men and women carrying out two different tasks. We found that even though men and women performed the individual tasks with the same speed and accuracy, mixing the two tasks made men slow down more so than women. From this, we conclude that women have an advantage over men in multi-tasking (of about one third of a standard deviation). In Experiment 2, we measured men and women's multi-tasking performance in a more ecologically valid setting. We found that women performed considerably better in one of the tasks measuring high level cognitive control, in particular planning, monitoring, and inhibition. In both experiments, the findings cannot be explained as a gender difference in a speed-accuracy trade off. Altogether, we conclude that, under certain conditions, women have an advantage over men in multi-tasking.

Relation to other work

As noted in the introduction, there is almost no empirical work addressing gender differences in multi-tasking performance. For example, even though there are numerous task-switching papers, none has focused on gender differences[d]. In fact, most task-switching studies do not explore individual differences, and accordingly are carried out with small samples.

Because they are typically carried out in psychology undergraduate programmes (with less than 20% male students), there are few male participants. The novelty of our study is not only the relatively large number of participants, but also the good gender balance. Despite the few studies about gender differences in multi-tasking, there has been an interesting discussion very recently about a study by Mäntylä (2013) which received much attention. Probably the main reason for the attention in the media for this study was the conclusion that men performed *better* than women in a multi-tasking paradigm. The finding of that study thus not only contrasts with the widely held belief that women are better at task switching, it also contrasts with our current data and the experiment by Ren and colleagues (2009).

In the study by Mäntylä (2013), men and women's accuracy in a visual detection task was measured. Participants had to detect specific numerical patterns in three different counters presented on a computer screen. Simultaneously, participants had to carry out an N-back task (stimuli appeared above the aforementioned counters). Men had a higher accuracy score of detecting the correct numerical patterns than women. The latter study is of great interest, because it addresses gender differences in multi-tasking of the second type, namely when tasks need to be carried out simultanously. Of interest is that for this specific type of multi-tasking, men had an advantage over women, and the degree of the advantage was directly related to men's advantage in spatial skills. But as argued in the introduction, this type of multi-tasking is potentially of less relevance to daily life contexts in which people often carry out tasks sequentially. In a comment on the study by Mäntylä (2013), Strayer and colleagues (2013) argue that gender is a poor predictor of multi-tasking. They present data to back this up from their own work on multi-tasking when driving. Arguably, studies showing no gender differences might simply have received less attention due to a publication bias for positive effects. We think that their Strayer et al.'s comments are valuable to the discussion, although their findings seem to primarily apply to the concurrent multi-tasking situations. That said, we found only one study that reported no gender differences in a task-switching paradigm in which people switched between two tasks. Buser and Peter (Buser and Peter 2012) had three groups of participants solving two different types of puzzles (sudoku and word-search). The group that did the two puzzles without switching between them solved the puzzles best, while switching between the puzzles while solving them impaired performance. The degree of impairment was similar for men and women, irrespective of whether the switching was voluntary or imposed. This situation is somewhat similar to Experiment 2, and thus, especially gender differences in this type of task-switching need further study to draw strong conclusions.

Finally, our finding that men and women did not differ in the effect of phone calls might be linked to a study by Law and colleagues (2004). They stated that the effects of interruptions are "quite subtle" and that more research on their effect on multi-tasking is necessary.

Limitations

We would like to consider a number of limitations of our current study that have implications for the interpretation of our results. First, as already mentioned above, there are many different ways to test multi-tasking performance. Because this is an emerging field with a small extant knowledge base we cannot exclude the possibility that our findings only hold true for the two specific paradigms we employed. Given the aforementioned work by Mäntylä (2013) and others that did not find the effect, and the general sparsity of the reports on the effect, this is a possibility that must be seriously considered.

A second limitation is that we did not formally record levels of education or control for general cognitive ability. Although we think it is not very likely, we appreciate the comment of one of the reviewers that if their were different levels of education this could potentially affect cognitive performance. The only way to exclude this possibility is to formally record the highest level of education of all participants.

A third limitation is that the power of the Experiment 2 may be low. Again, it is difficult to say although evidently powerful enough to detect moderate differences on the key search task - so it may be a task-related issue and further work needs to investigate task-based constraints in multi-tasking. For example, we did not conclude that there was a gender difference in arithmetic performance or time spent on the phone, but this could potentially be due to a lack of statistical power. In the case of the arithmetic task, there are good reasons not to expect a gender difference on simple arithmetic problems, even though we acknowledge the complexity of the study of gender differences in mathematical ability (c.f.,Halpern et al. 2007).

A final limitation is that although we checked that no gender differences emerged on the Key Search with both the test authors and with the published norms, we cannot eliminate the possibility that a difference may have emerged tested alone. We could have retested the individual tasks with another sample of participants. Also, we could have run a repeated measures design (same participants on the individual tasks), although this would defeat the novelty aspect of the task. The best way to address this issue is for another research group to replicate the finding.

Conclusions

Our findings support the notion that woman are better than men in some types of multi-tasking (namely when the tasks involved do not need to be carried out simultaneously). More research on this question is urgently needed, before we can draw stronger conclusions and before we can differentiate between different explanations.

Endnotes

[a]The two experiments were carried out by independent groups of researchers. We only realised the similarity between the two experiments and their findings afterwards. We believe that the two experiments complement each other: While Experiment 1 uses a laboratory based reaction time experiment, Experiment 2 uses a much more ecologically valid approach.
[b]This is likely because of the availability of computers to measure response times. In the 1920s, it would have been hard, if not impossible, to accurately measure task-switching costs, while measuring mixing costs could be done with the paper-and-pensil tests used by Jersild (1927).
[c]Throughout the results section, we report means ± 1 standard error of the mean.
[d]To the best of our knowledge.

Competing interests

The authors declare that they had no competing interests.

Authors' contributions

GS,DO, and MC carried out Experiment 1. KL carried out Experiment 2. The four authors wrote the article together. All authors read and approved the final manuscript.

Acknowledgements

Experiment 1 was made possible with a grant from the British Academy to Stoet, O'Connor, and Conner and with the assistance of Weili Dai, Caroline Allen, and Tansi Warrilow.

References

Broadbent, DE (1952). Failures of attention in selective listening. *Journal of Experimental Psychology, 44*(6), 428–433.

Burgess, PW (2000). Real-world multitasking from a cognitive neuroscience perspective. In Monsell S, Driver J (Eds.), *Attention and Performance, Volume XVIII*. (pp.465–472). Cambridge, MA: MIT press.

Burgess, PW, Veitch, E, Costello, AD, Shallice, T (2000). The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia, 38*(6), 848–863.

Buser, T, & Peter, N (2012). Multitasking. *Experimental Economics, 15*, 641–655.

Dibben, CRM, Rice, C, Laws, KR, McKenna, PJ (2009). Is executive impairment associated with schizophrenic syndromes? A meta-analysis. *Psychological Medicine, 39*, 381–392.

Eriksen, BA, Eriksen, CW (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics, 16*, 143–149.

Evans, JJ, Chua, SE, McKenna, PJ, Wilson, BA (1997). Assessment of the dysexecutive syndrome in schizophrenia. *Psychological Medicine, 27*(3), 635–646.

Franz, EA, Zelaznik, HN, McCabem, G (1991). Spatial topological constraints in a bimanual task. *Acta Psychologica, 77*(2), 137–151.

Halpern, DF, Benbow, CP, Geary, DC, Gur, RC, Hyde, JS, Gernsbacher, MA (2007). The science of sex differences in science and mathematics. *Psychological Science, 8*, 1–51.

Hill, K, Mann, L, Laws, KR, Stephenson, CME, Nimmo-Smith, I, McKenna, PJ (2004). Hyperfrontality in Schizophrenia. A meta-analysis of functional imaging studies. *Acta Psychiatrica Scandinavica, 110*, 243–256.

Ishihara, S (1998). *Ishihara's Tests for Colour Blindness.* London: Hodder Arnold.

Jersild, AT (1927). Mental set and shift. *Archives of Psychology, 89*, 1–81.

Kravariti, E, Dixon, T, Frith, C, Murray, R, McGuire, P (2005). Association of symptoms and executive function in schizophrenia and bipolar disorder. *Schizophrenia Research, 74*(2–3), 221–231.

Law, AS, Logie, RH, Pearson, DG, Cantagallo, A, Moretti, E, Dimarco, F (2004). Resistance to the impact of interruptions during multitasking by healthy adults and dysexecutive patients. *Acta Psychologica, 116*(3), 285–307.

Laws, KR (1999). A meta-analytic review of Wisconsin Card sorting studies in Schizophrenia: a general intellectual deficit in disguise? *Cognitive Neuropsychiatry, 4*, 1–35.

Mäntylä, T (2013). Gender differences in multitasking reflect spatial ability. *Psychological Science, 24*, 514–520.

Offer S, & Schneider, B (2011). Revisiting the gender gap in time-use patterns: multitasking and well-being among mothers and fathers in dual-earner families. *American Sociological Review, 76*(6), 809–833.

Pashler, H (1984). Processing stages in overlapping tasks: Evidence for a central bottleneck. *Journal of Experimental Psychology-human Perception And Performance, 10*(3), 358–377.

Ren, D, Zhou, H, Fu, X (2009). A deeper look at gender difference in multitasking: gender-specific mechanism of cognitive control. In *Fifth international conference on natural computation* (pp. 13–17). Washington: IEEE Computer Society.

Robertson, IH, Ward, A, Ridgeway, V, Nimmo-Smith, I (1994). *Tests of everyday attention.* Bury St Edmunds, U.K.: Thames Valley Test Company.

Royall, DR, Lauterbach, EC, Cummings, JL, Reeve, A, Rummans, TA, Kaufer, DI, LaFrance, WC, Coffey, CE (2002). Executive control function: a review of its promise and challenges for clinical research. a report from the Committee on Research of the American Neuropsychiatric Association. *Journal of Neuropsychiatry and Clinical Neurosciences, 14*, 377–405.

Salvucci, DD, & Taatgen, NA (2010). *The multitasking mind.* Oxford, UK: Oxford University Press.

Sayer, LC (2007). Gender differences in the relationship between long employment hours and multitasking. In BA Rubin (Ed.), *Workplace Temporalities (Research in the Sociology of Work)* (pp. 403–435). Amsterdam: Elsevier.

Semkovska, M, Bedard, MA, Godbout, L, Limoge, F, Stip, E (2004). Assessment of executive dysfunction during activities of daily living in schizophrenia. *Schizophrenia Research, 69*(2-3), 289–300.

Shallice, T, & Burgess, PW (1991). Deficits in strategy application following frontal lobe damage in man. *Brain, 114*, 727–741.

Silverman, I, & Eals, M (1992). Sex differences in spatial abilities: evolutionary theory and data. In Barkow J, Cosmides L, Tooby J (Eds.), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture* (pp.487–503). New York: Oxford University Press.

Stoet, G (2010). PsyToolkit – A software package for programming psychological experiments using Linux. *Behavior Research Methods, 42*(4), 1096–1104.

Stoet, G, & Snyder, LH (2003). Task preparation in macaque monkeys (*Macaca mulatta*). *Animal Cognition, 6*, 121–130.

Stoet, G, & Snyder, LH (2012). The role of executive control in tool use. *Behavioral and Brain sciences, 35*, 240–241.

Strayer, DL, Medeiros-Ward, N, Watson, JM (2013). Gender invariance in multitasking: a comment on Mäntylä (2013). *Psychological Science, 24*, 809–810.

Vandierendonck, A, Liefooghe, B, Verbruggen F (2010). Task switching: interplay of reconfiguration and interference control. *Psychological Bulletin, 136*(4), 601–626.

Watson, JM, & Strayer, DL(2010). Supertaskers: profiles in extraordinary multitasking ability. *Psychonomic Bulletin & Review, 17*(4), 479–485.

Wilson, BA, Evans, JJ, Emslie, H, Alderman, N, Burgess, P (1998). The development of an ecologically valid test for assessing patients with a dysexecutive syndrome. *Neuropsychological Rehabilitation, 8*(3), 213–228.

Additional file

Additional_file_1 as JAR
Additional file 1: Demonstration of task-switching paradigm (Java application which runs on all desktop computers with Java installed).

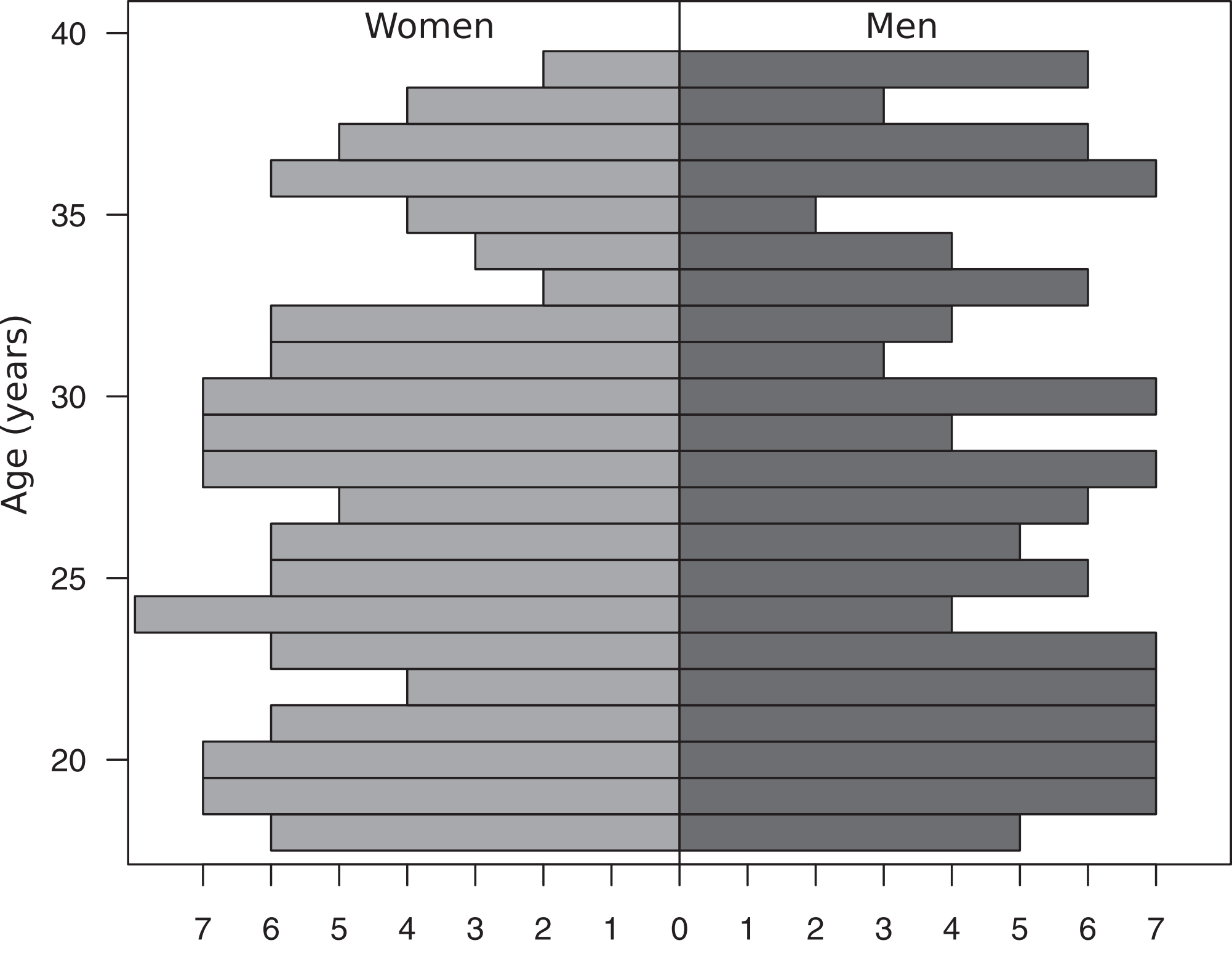
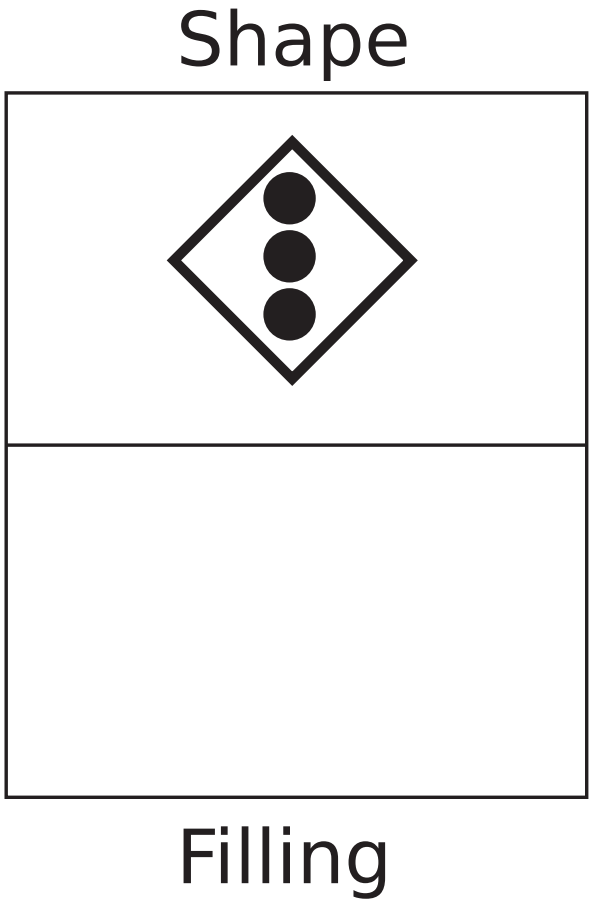

Women
Men
40
35
30
25
20
Age (years)
7
6
5
4
3
2
1
0
1
2
3
4
5
6
7
participant count

Figure 1

A



B

Shape task	Left	Left	Right	Right
Filling task	Left	Right	Left	Right

Figure 2

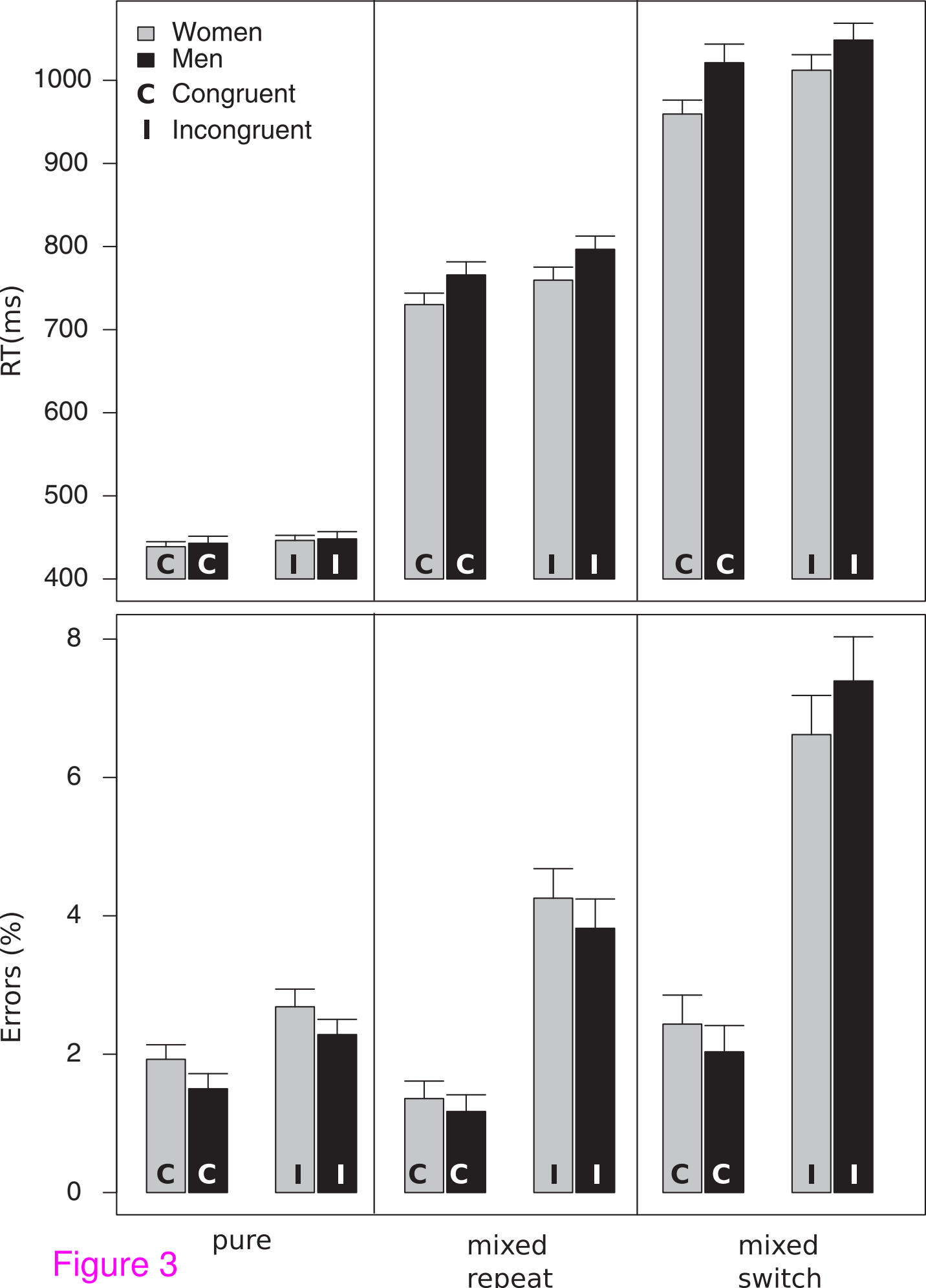
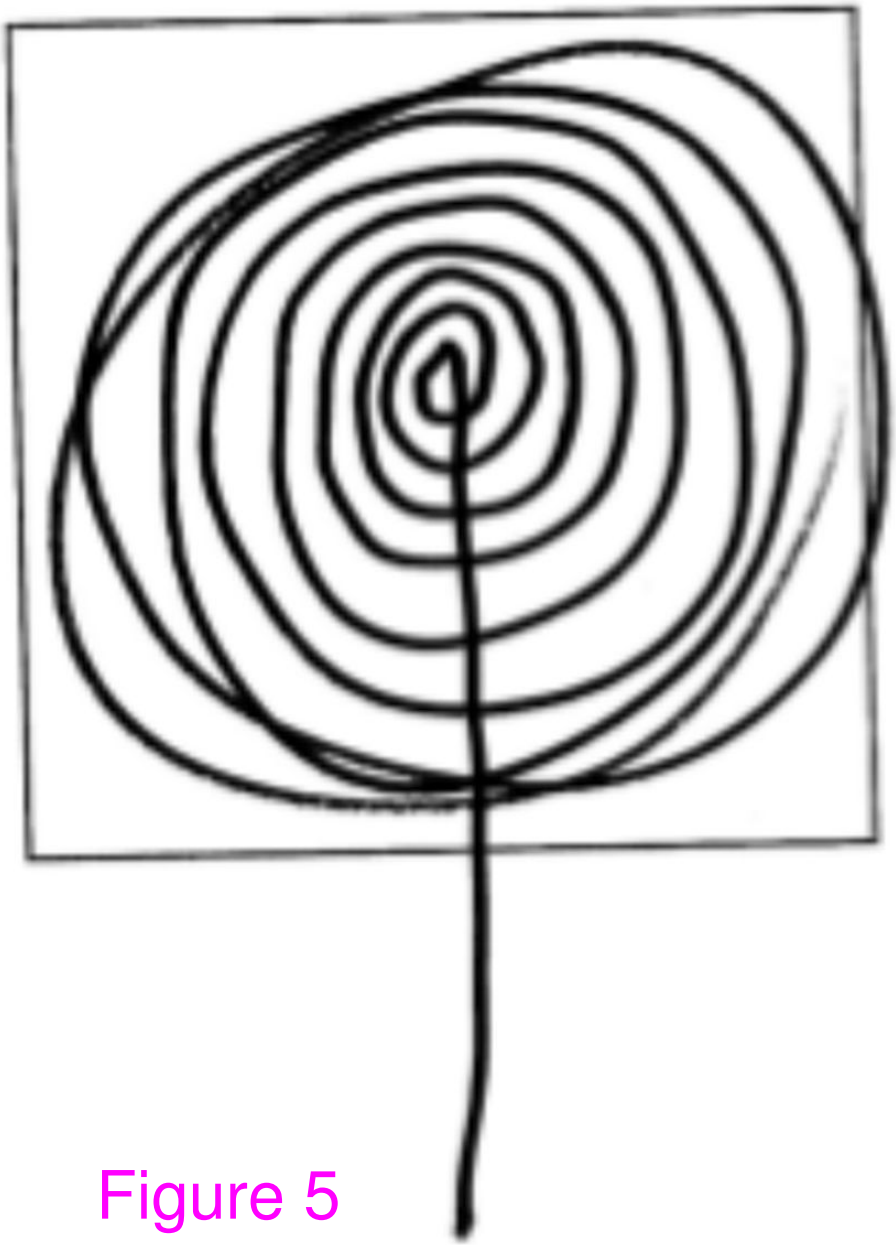

Women
Men
C Congruent
I Incongruent
RT(ms)
1000
900
800
700
600
500
400
Errors (%)
8
6
4
2
0
C C I I C C I I C C I I
pure
mixed repeat
mixed switch

Figure 3

	x	7	=	35
42	÷		=	7
	-	17	=	26
19	+		=	67

Figure 4



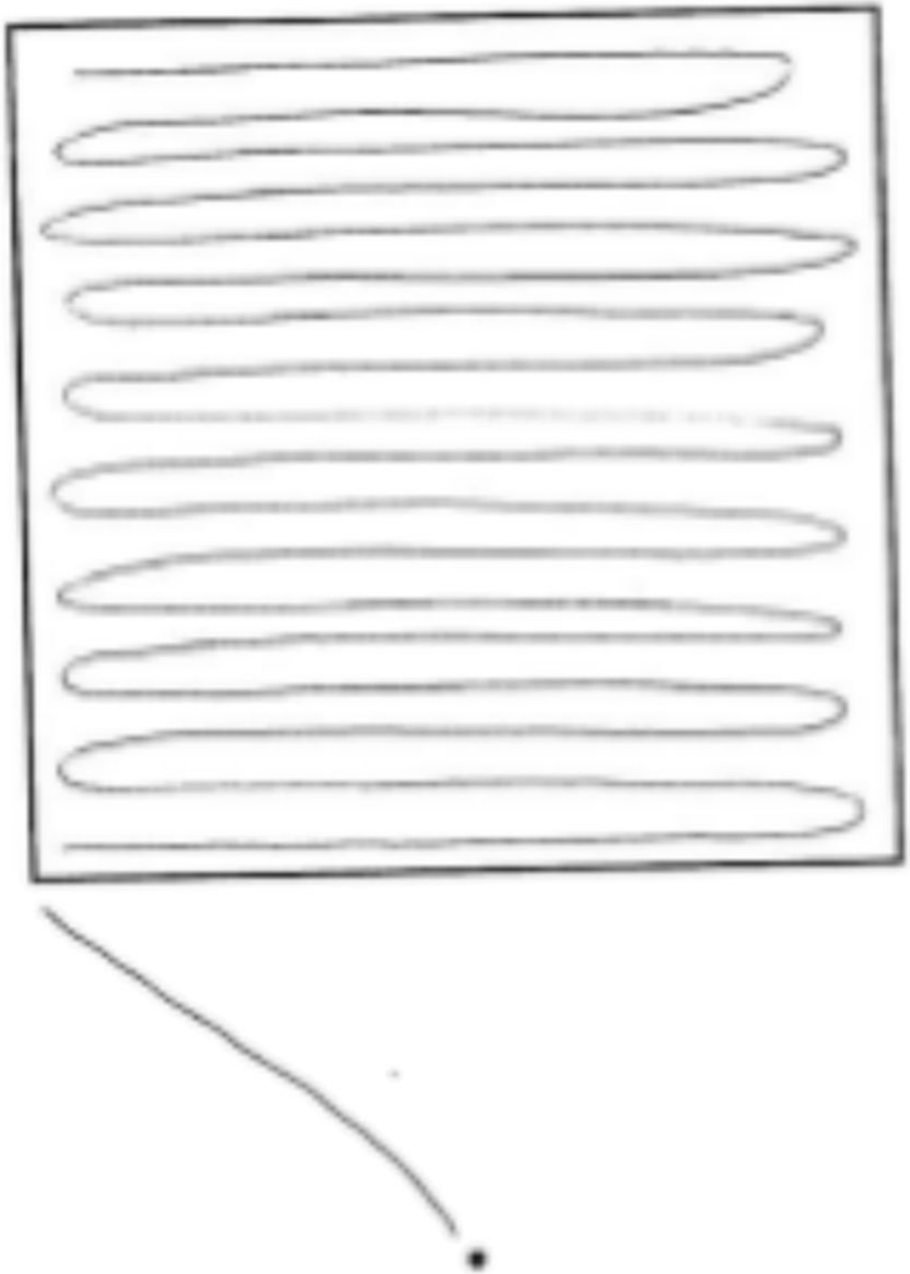


Figure 5

Additional files provided with this submission:

Additional file 1: 1841347955883140_add1.jar, 999K
http://www.biomedcentral.com/imedia/5080532111079484/supp1.jar