This is a preprint of an article submitted for consideration in the European Journal of Cognitive Psychology. European Journal of Cognitive Psychology is available online at http://journalsonline.tandf.co.uk/

Random number generation and working memory.

Towse, John N (Lancaster University, UK)

Cheshire, Andrea (Lancaster University, UK)

Address for correspondence

Dr John N Towse
Department of Psychology
Fylde College
Lancaster University
Lancaster LA1 4YF
United Kingdom
Email: j.towse@lancaster.ac.uk
Tel:+44-1524-593705, Fax:+44-1524-593744

Authors' notes. This paper is dedicated to the memory of Derek Neil (1968-2001), who contributed software to the analysis of random sequences. More information about automated analysis of sequences can be found at http://www.lancs.ac.uk/staff/towse/rgcpage.html. We are grateful for comments on this paper from André Vandierendonck, Iring Koch, and an anonymous reviewer.

Abstract

We demonstrate the close relationship that exists between random sequence generation and working memory functioning. We clarify the nature of this link by examining the impact of concurrent requirements for random sequence response quality. Experiments 1A and 1B show that marking specific response choices for differential treatment, either by requiring an ancillary behaviour or by suppressing these choices from output, impairs overall sequence quality. Contrasting with previous findings, these distinct concurrent tasks have comparable effects. We show that disruption is found only when concurrent demand is high. Experiment 2 demonstrates that increasing the dynamic working memory load by requiring the ancillary response to change during the task leads to additional disruption of randomisation. The results extend and refine our understanding of the contribution of active maintenance of representations in random generation.

Introduction

From time to time, cognitive psychologists ask their experimental participants to produce sequences of random numbers under substantial time pressure. Invariably, the participants struggle with the task, even in relatively benign study configurations. While randomisers understand their task in general, its execution is problematic. That is, although participants may labour under some misunderstandings about randomness (e.g. Bar-Hillel & Wagenaar, 1991), they often recognise the inadequacy of their own response choices, implicating production constraints.

Yet, we should probably not be overly surprised at this task difficulty. It is worth bearing in mind how difficult it is for *researchers* to capture randomness in a single, comprehensive measure, though one can certainly assess many facets of performance (e.g., Ginsburg & Karpiuk, 1991; Towse & Neil, 1998). So pity the poor *participants* who must submit their response sequence to such varied scrutiny! Also, mental processes are more commonly deployed in the service of pattern detection and environmental prediction (e.g. Kareev, 1995). Therefore, producing behaviour that lacks the qualities that otherwise are almost ubiquitous, goes against the grain.

Experimental research into random generation has suggested that response production in adults (and children) relies on “Executive Functions” (EFs) (Baddeley, 1986; Baddeley, Emslie, Kolodny & Duncan, 1998; Miyake et al., 2000; Towse & Mclachlan, 1999; Vandierendonck, 2000). Whilst EFs represent a loosely defined cognitive construct, random generation illustrates the engagement of different higher-order mental operations that modulate other processes. More specifically, in generating random numbers there are natural, over-learned sequences or response chains (e.g. “1, 2, 3”) and rate-limited, inhibitory processes (possibly involving dorsolateral prefrontal cortex; Knoch, Brugger & Regard, 2005) act to minimise their occurrence. Randomisers also need to consider all response alternatives and performance declines as the response set size increases (Towse, 1998). This is consistent with the notion of an additional limited capacity activation function that tries to maintain representations of responses so that they are available for selection.

Despite a generic consensus that random generation involve EFs (see Brugger, 1997), research has been “sporadic” (Heuer, Kohlisch, & Klein, 2005). Thus, it is important to focus on the nature of the executive requirements in randomisation and we concentrate here on the particular role of working memory – i.e. the maintenance and transformation of transient representations (Baddeley. 1986; 2000).

To investigate working memory constraints on randomisation, Towse and Valentine (1997, Expt. 2) described a paradigm that imposed additional requirements. Participants either suppressed particular response values (e.g., randomising numbers between 1 and 12 but avoiding 4 and 8 as responses) or performed an ancillary memory task (e.g., randomising numbers between 1 and 10 but tapping the desk on production of 4 and 8 to demonstrate memory for these ‘marked’ values). Insofar as one must remember an item’s special status in order to withhold it from output, the remember condition formed a useful control for measuring active inhibition itself.

Towse and Valentine (1997) found that sequence randomness declined with both concurrent requirements, measured in terms of digram combinations and the occurrence of rising or falling ordinal sequences. They also reported more adjacent responses were produced in the suppression condition than the remember condition and there were more counting sequences (e.g., “3, 4” or “9, 8”). One interpretation was that the executive requirement to suppress response values - over and above the need to remember them - degraded the control over response production. Consequently, output included more prepotent responses.

The findings raise several additional issues. For example, they do not address whether any concurrent task requirement (suppression or memory) produces the observed pattern, or alternatively, whether the size of the concurrent load is critical. Towse and Valentine (1997) marked either two out of ten responses for differential treatment, or none. Consequently, the observed effect could arise from the additional task requirement itself, or it could be load-dependent. This question relates to the theories of random generation and cognitive control. For example, the need to represent and maintain additional task goals (Duncan, Emslie, Williams, Johnson, & Freer, 1996) might constrain the ability to implement the primary task objective of randomisation. Alternatively, both primary and concurrent requirements might rely on working memory operations, yet impairment may occur only when capacity is sufficiently challenged. This latter type of approach suggests that performance is dependent on not just the presence but also the extent of the concurrent task requirements. In general, this view would be compatible with ‘load theory’ (Lavie, Hirst, De Fockert & Viding, 2004) whereby large cognitive control requirements enhance the effect of distractors. Of course, these two accounts need not be exclusive, and they are examined below.

Experiment 1A

This study explores and extends previous findings of a disruptive effect on random generation from remember and suppress requirements, as well as the overproduction of adjacent values with suppression. We varied the concurrent task load to specify its source; marking a single item and two items for differential treatment, as well as using a control condition. The sample size matched Towse and Valentine (1997) and participants produced 81 responses from among 9 alternatives. This response sequence density, relative to response choice, is comparable with other test protocols.

Participants

Twenty-four naïve adult participants registered through a research notice-board. They were paid a £2 honorarium for a 15-minute experimental session. Twelve participants were assigned at random to the remember and twelve to the suppression conditions.

Procedure

Participants initially read a set of standard instructions (see Appendix 1). These requested a random number sequence at a regular speed (1 item every 1.5 seconds) cued by an auditory signal, a tone, from a cassette tape. The experimenter identified the set of 9 response alternatives. Participants in either condition provided 3 random number sequences; with no concurrent task (control), with one number and with two numbers marked for differential treatment. In the remember condition, participants tapped the desk in front of them when they produced a marked value. They were asked that the ancillary task not influence their random responses (they should not produce the number(s) more or less often than they otherwise would). In the suppression condition, participants attempted to avoid saying marked response value(s) (e.g., use the numbers 1 to 10 but avoid 4, or 1 to 11 but avoid 3 and 7). Marked number values varied across participants and sequences, they were never the largest or smallest response choice, and in the case of the two-item condition, the number values were different sides of the median. Each production sequence comprised 81 responses. The order of tasks, either control to one-item to two-item conditions or vice versa, was counterbalanced.

Results

Response compliance

Before examining the quality of random sequences, we report on two forms of response failures. The first was a failure to maintain the response pace (see Table 1). Analysis of variance showed a trend for an increased number of missed responses with a concurrent task, $\underline{F}(1.3,28.6) = 3.51$, $\underline{p}<.07$, $\eta_p^2=.138$[1]. Suppression led to more missed responses than the memory requirements, but not significantly so, $\underline{F}(1,22) = 2.64$, $\eta_p^2=.107$. However, four participants missed approximately 20% of response cues in at least one condition. Their response rate was considerably slower than that specified in the experimental design.

---

[1] In cases where ANOVA sphericity assumptions were violated, Greenhouse-Geisser corrections were applied to the degrees of freedom, and these values are reported in the text.

A second error type concerns response violations; the production of a value outside the permitted range (e.g., the number 10 when randomising between 1 and 9), a prohibited response in the suppression condition, or a marked value response without an accompanying memory signal in the remember condition. Analysis of variance confirmed that more violations occurred when there were concurrent requirements, $\underline{F}(2,34.0) = 10.01$, $\underline{p} < .01$, $\eta_p^2 = .313$, more violations under suppression instructions, $\underline{F}(1,22) = 5.28$, $\underline{p} < .05$, $\eta_p^2 = .194$, and an interaction between these factors, $\underline{F}(2,34.0) = 4.19$, $\underline{p} < .05$, $\eta_p^2 = .160$. The increase in response violations across load was more marked in the suppression condition than the remember condition.

Table 1 about here.

Response quality

Randomness is unusual because it concerns the absence of any structural feature. Randomness indices generally assess the extent to which responses are non-random. They actually measure the orderliness or regularity (see also Reed & Johnson, 1994 for a perspective from the sequence learning literature on response structure). However, regularity may occur in myriad different forms, and lack of response structure in one respect may be independent of another. From analysis of different indices, Towse & Neil (1998) offered evidence that there were potentially at least four psychological factors in randomisation. We therefore sample from (and only from) measures representative of these factors, explained in more detail below and in Appendix 2; We employed three measures of stereotyped sequences (Turning Point Index (TPI), Random Number Generation (RNG), & Adjacency (A)); one measure of the tendency to draw upon all responses evenly (Redundancy (R)); and short-term and long-term repetition avoidance (Phi2 & Phi7 respectively). Of principal interest, given previous results, are measures of sequence stereotypy; others are reported for completeness.

TPI scores measure stereotyped behaviour by assessing changes between ascending and descending runs of numbers. Analysis of variance showed a significant effect of task load, $F(2,44) = 6.41$, $p<.01$, $\eta_p^2=.226$, mainly attributable to greater non-random bias on the two-item condition. This differed from the control and one-item condition, ($t(23)=3.11$, $p<.01$, $d=.634$, & $t(23)=2.87$, $p<.01$, $d=.585$ respectively) but the one-item condition did not differ from control, $t(23)=.032$, $d=.006$. There was no significant difference between suppression and remember conditions, $F<1$, $\eta_p^2=.005$, and no interaction, $F<1$, $\eta_p^2=.019$. As explained in Appendix 2, RNG scores assess the reliance on repeating response pairs. Analysis suggested no significant effect of task load, $F(2,44) = 2.24$, $\eta_p^2=.092$, or difference between concurrent conditions, $F<1$, $\eta_p^2=.001$, or an interaction, $F(2,44) = 2.18$, $\eta_p^2=.090$.

R scores assess the degree to which response alternatives are chosen equally. R varied marginally with task load, $F(1.4, 30.4) = 3.08$, $p<.08$, $\eta_p^2=.123$ (compared with control performance, R scores increased with a one-item load but decreased with a two-item load). There was no significant difference between remember and suppression conditions, $F<1$, $\eta_p^2=.037$, nor an interaction, $F(1.4,30.4)=1.15$, $\eta_p^2=.049$. Phi2 scores essentially measure response repetition. These are often infrequent in human random sequences, as can be seen with first-order differences in Figure 1, where ‘0’ indicates immediate repetitions (Brugger, Monsch, Salmon, & Butters, 1996). Phi2 scores did not vary significantly with task load, $F<1$, $\eta_p^2=.032$, or task type, F(1,22) = 2.11, η_p^2=.088, and these factors did not interact, F<1, η_p^2=.031. Phi-7 scores reflect repetitions across (five) intervening items. These values were also invariant across experimental condition, with no significant effect of task load, F<1, η_p^2=.036, task type, F<1, η_p^2=.01, or interaction, F<1, η_p^2=.011.

Figure 1 about here.

Towse and Valentine (1997) reported that two-item suppression, relative to remember demands, specifically increased A scores (a response value adjacent to its predecessor, i.e., '1' and '-1' in Figure 1). Analysis of A scores indicated a significant difference between task load conditions, F(1.4,30.6) = 7.68, p<.01, η_p^2=.259, but no difference between remember and suppress requirements, F<1, η_p^2=.029 and no interaction, F<1, η_p^2=.006. The effect of task load is attributable to the two-item condition: this differed from the one-item and control conditions (t(23)=3.19, p<.01, d=.651, and t(23)=2.77, p<.05, d=.565 respectively), while the one-item did not differ from control condition, t(23)=.81, d=.165.

Condition completion order was an incidental variable. Nonetheless we report its effect on randomisation quality. Analysis indicated only that participants produced lower A scores (adjacent responses) on all sequences when the control condition was attempted first. We take this to suggest that the control condition offered a beneficial introduction to the task, which then persisted.

Independence of task components

Although instructions asked participants to ignore concurrent task requirements in terms of random response choices, one can ask nonetheless whether flagging a particular response option as ‘special’ makes its selection more or less alluring. The response frequency of the single remember item (M=12.1, SD=3.78) differed from the overall average (9 items per choice), t(11) = 2.83, p<.05, d=.817. The (average) response frequency of the two remember items (M=9.42, SD=1.48) was not significantly greater than overall mean, t(11) = 1.13, d=.326. In the suppression condition, a somewhat analogous question is whether the prohibition on particular responses leads to the over-selection of neighbouring values, as participants hover around the illicit choices. Although responses to randomly selected neighbouring values were higher than the overall mean (M=9.17, SD=1.35 and M=9.29, SD=1.76, in the single and two-item condition), these effects were not significant (t(11) = .43, d=.124, and t(11) =.57, d=.165, respectively). Finally, the production frequency of those single response values that had been marked as different for the preceding sequence did not differ from global average frequencies for either the remember or suppress conditions. Thus, there was no evidence for ‘proactive interference’ from one condition to another.

Experiment 1B

In Experiment 1A several participants did not adhere to the pacing requirements. Some individuals may have achieved response suppression only by generating no response at all, masking task difficulty. Therefore, in Experiment 1B, extra effort was devoted to explaining the importance of the response pace. Revised instructions emphasised and reminded participants of the timing cues. We used 100 responses per condition from 10 choices to increase sampling density and even more closely replicate Towse and Valentine (1997), once again using the same sample size to facilitate comparison with that study and with Experiment 1A.

Participants and Procedure

Twenty seven adult participants registered through a research notice-board. All were naïve (one completed a separate spatial randomisation task 4 months previously) and paid £2. The procedure mostly followed Experiment 1A, although participants produced 100 responses from the numbers 1-10. In addition to the previous instructions, participants were told “it’s really important to try and respond within the time limit. Make sure you produce one response for each tone that you hear.” Later, they were told “I’m going to write down your responses. If you slow down and don’t give a response in the time available, I have to mark that as a missed response. I’d like you to avoid missed responses as far as possible”. Timekeeping instructions were re-iterated between sequence productions.

Results

Response compliance

Three participants were dropped from analysis because more than 10% of a sequence included illegitimate or missed responses. Table 1 details adherence to instructions, indicating greater compliance with instructions than Experiment 1A, especially with respect to timekeeping. Analysis of variance on response rate failures revealed a significant main effect of load, $\underline{F}(1.4,31.0)=6.25$, $\underline{p}<.05$, $\eta_p^2=.221$, no reliable difference between the tasks, $F(1,22)=1.44$, $\eta_p^2=.061$, and no interaction, $F(1.4,31.0)=2.54$, $\eta_p^2=.103$. Analysis of variance on the number of response violations indicated a significant load effect, $F(1.2,26.4)=9.02$, $p<.01$, $\eta_p^2=.291$, more violations in the suppression condition, $F(1,22)=12.4$, $p<.01$, $\eta_p^2=.360$, and a significant interaction, $F(1.2,26.4)=5.61$, $p<.05$, $\eta_p^2=.203$. The concurrent load effect on response errors was more marked in the suppression than in the memory condition.

Response quality

Analysis of TPI scores (see Appendix 2 for details) showed no significant effect of load, $F(1.6,34.4)=2.14$, $\eta_p^2=.089$, task, $F<1$, $\eta_p^2=.004$, or interaction, $F<1$, $\eta_p^2=.002$. Likewise, RNG scores did not show an effect of load, $F(1.5,31.9)=2.09$, $\eta_p^2=.087$, task, $F(1,22)=3.02$, $p<.10$, $\eta_p^2=.121$, or an interaction, $F<1$, $\eta_p^2=.034$.

Indices that reflected other randomness factors did not show significant experimental effects; for the manipulation of load [R scores, $F(2,44)=1.36$, $\eta_p^2=.058$, RNG scores, $F(1.5,31.9)=2.09$, $\eta_p^2=.087$, Phi2 values, $F(2,44)=1.47$, $\eta_p^2=.063$, or Phi7 values, $F(2,44)=1.51$, $\eta_p^2=.065$], or task [R scores, $F<1$, $\eta_p^2<.001$, RNG scores, $F(1,22)=3.02$, $p<.10$, $\eta_p^2=.121$, Phi2 scores, $F<1$, $\eta_p^2=.023$ or Phi7 scores, $F<1$, $\eta_p^2=.003$]. Interaction terms were all non significant (Fs<2.03, $\eta_p^2<.085$).

However, as in Experiment 1A, the measure of adjacent values (A) revealed significantly greater response regularity with a concurrent load, $F(2,44)=10.13$, $p<.001$, $\eta_p^2=.315$. Again, there was no difference between memory and suppression requirements, $F<1$, $\eta_p^2=.004$, and no interaction, $F(2,44)=1.97$, $\eta_p^2=.082$. Figure 2 illustrates this result. Pairwise comparisons again confirmed that both control and single-item conditions differed from the two-item condition, ts(23)>3.86, $p<.01$, ds>.788, while the control and single item values did not differ, t(23)=1.08, d=.220. Analysis of order effects (whether the control condition was first or last) showed no reliable differences for any of the above randomisation indices.

Figure 2 about here.

We then combined control and two-item A scores from Experiment 1A and 1B to produce a sample size double that used by Towse and Valentine (1997). Analysis confirmed a strong load effect, F(1,44)=18.0, $p<.01$, $\eta_p^2=.290$, but still no difference between concurrent tasks, F<1, $\eta_p^2=.007$, and no interaction, F<1, $\eta_p^2=.010$.

Independence of task components

With both one and (the average of) two marked remember values, production frequency was significantly higher than the overall mean (10), (M=14.8, SD=1.95, t(11)=8.6, $p<.001$, d=2.48 and M=13.0, SD=2.00, t(11)=5.20, $p<.001$, d=1.50, respectively). In the single-item suppression condition, the frequency of a randomly selected neighbour was higher than the average response frequency of 10, (M=11.3, SD=1.87, t(11)=2.46, $p<.05$, d=.710). With two suppression items, mean production of neighbourhood values did not differ from the overall mean (M=10.6, SD=1.88, t(11)=1.15, d=.332). The production frequency of those response choices marked as different on the preceding condition did not differ from the overall mean for either the remember or suppress conditions.

Discussion

The data are valuable in at least three of respects. First, they partially replicate previous research, demonstrating specific disruption to randomisation from concurrent tasks. Second, they extend that work, suggesting the point at which concurrent tasks interfere. Third, results also challenge aspects of the original findings.

Using a similar design and sample size to Towse and Valentine (1997), Experiment 1A and 1B confirm that randomisation performance is affected by instructions both to remember and to avoid certain response values. Such effects are localised largely in the production of adjacent values and the chaining of response sequences (that changes occur only for certain indices simply points to the specific nature of disruption).

Towse and Valentine (1997) compared a control condition with no concurrent requirement to a condition in which two items were identified for special treatment (associated with a remember or suppression action). The present studies extended this design by incorporating a further condition in which just one item was singled out for special treatment. The results show that the one item concurrent load was largely equivalent to the control condition, while both differed from the two-item concurrent load. Although one might treat the lack of difference between the control and single-item conditions cautiously, the two-item condition is statistically distinct from the others. The data imply that a concurrent task *per se* is less critical to performance than the exact load of that task. Such a conclusion resonates with, for example, seminal working memory studies reporting impairment on a reasoning task by an additional 6-item memory load but not a 2-item load (Hitch & Baddeley, 1976).

In contrast to the previous reports, we found an equivalent disruption effect from each concurrent task, for both Experiment 1A and 1B which had the same sample size as previous analysis, as well as when these datasets were combined together. Thus, the impact of the concurrent requirements does not appear to be task specific, as initially argued (Towse & Valentine, 1997) but instead arises from being functionally distracted from randomisation. We conclude that there is less reason to argue that acting upon a representation (by suppressing its production) has noticeable effects beyond the maintenance of representations themselves during randomisation. That is not to say that the concurrent tasks are the same – there are differences in the opportunities for and occurrences of response compliance failures for example – but we do not find consistent differences here in randomisation production.

Inspection of First Order Differences (FOD) gives a possible clue as to the difference between the present findings and previous reports. In both current datasets, two-item suppression produces an increase in “-1” and “1” FOD values (albeit in Experiment 1B more symmetrical). This matches Towse and Valentine (1997). However, the remember instructions also lead to “-1” and “1” increases here, but had very little impact in Towse and Valentine (1997). That is, the suppression effect is actually highly consistent across all three analyses, but the effect of remember requirements was smaller in the original study than has been obtained here. Previous results appear to have underestimated the impact of concurrent memory operations. This is investigated further in the next experiment.

Experiment 2

Experiment 1 establishes that randomisation is affected by a substantive concurrent task contingent upon primary task response choices. The effect is robust, although largely localised to measures of sequence stereotypy. In other words, when participants are required to accompany some of their random responses with another behaviour (i.e., tapping a desk when certain responses are produced), the quality of random generation declines. However, in the data reported so far, the memory items are constant through the task - it is always the same one or two items that require an identical motoric response. Experiment 2 investigated whether a dynamic concurrent requirement – one in which the response changed through the task and thereby imposed greater working memory requirements, would produce more disruption.

To produce a dynamic concurrent task, participants engaged in a sequence of response movements around the perimeter of a laminated A4 board. There were 8 locations in a rectangular outline. Participants made a series of hand taps, with the target location moving in a clockwise direction following each marked response. For example, if the numbers “2” and “7” were marked for tapping, when the participant chose “2” as a random response, they tapped the bottom left location. On the next production of “2” they tapped the middle left location. If they then selected “7” for the first time, they tapped the bottom left location, and so on. Participants therefore needed to maintain and update a memory representation of the tap location for the two marked numbers.

On the assumption that working memory representations are utilised in random generation performance, we predicted that the dynamic location tracking requirements would impair randomisation to a greater extent than the constant remember requirements, even though the latter induces a noticeable decline in performance relative to control conditions. Since we focus on the relative impact of the dynamic vs. static memory requirement, a control condition was not employed here.

Participants

Seventeen naïve adult participants registered through a research notice board. They were paid a £2 honorarium for a 15-minute experimental session.

Procedure

The random generation task was explained as before, emphasising timekeeping (see Experiment 1B). Participants produced two sequences – with remember or track instructions – in counterbalanced order. They were asked to make randomisation choices without reference to the concurrent task (i.e. not to use marked responses differently from others) and produced a sequence of 100 single digit numbers.

Remember instructions. The experimenter identified two response choices (values varied across individuals). Participants tapped the desk in front of them whenever they chose either number as part of their random sequence.

Track instructions. The experimenter identified two response choices (these also varied across individuals and were always different from those used in the remember condition). Participants tapped the appropriate location (marked on a laminated sheet in front of them) whenever they chose either number as part of their random sequence. They started at the bottom left position, advancing locations in a clockwise direction, tapping boxes in a separate sequence for each of the two numbers.

Results

One participant was excluded from analysis after failing to follow task instructions with respect to timekeeping and concurrent tasks. We then calculated the average number of response rate failures and rule violations (a response choice outside the permitted range or a failure to tap for a marked response). We also noted separately the number of tracking errors for both marked values: occasions when a participant failed to select the correct spatial location. As Table 2 shows, compliance with randomisation requests was good, and the trend for fewer response rate and response violation errors in the tracking condition was not significant, $\underline{t}(15)=1.96$, $\underline{p}<.10$, $\underline{d}=.49$ and $\underline{t}(15)=1.35$, $\underline{d}=.338$. Table 2 shows that participants clearly had problems tracking the appropriate position to tap when they produced one of the marked response choices, with participants both repeating and skipping locations. This is suggestive of a possible, albeit weak trade-off, with randomisation compliance in the tracking condition occurring at the expense of tracking accuracy.

Table 2 and Figure 3 about here.

Analysis of randomisation revealed performance was significantly poorer in the tracking compared with the remember condition as measured by the RNG index, $\underline{t}(15)=3.02$, $\underline{p}<.01$, $\underline{d}=.755$, and by the A index, $\underline{t}(15)=4.72$, $\underline{p}<.001$, $\underline{d}=1.18$, see Figure

3. TPI scores were marginally less random in the track compared with the remember condition, t(15)=1.97, p<.10, d=.493. Other dimensions of randomisation quality were not affected; task differences were not significant as measured by R scores, t(15)=1.64, d=.41, Phi2 scores, t(15)=.91, d=.228, or Phi7 scores, t(23)=1.66, d=.415. In summary, the requirement to track locations for two numbers, in comparison to making a hand tap, led participants to produce more adjacent response choices and digram pairs (as well as a trend for longer runs of successively increasing or decreasing sequences).

Comparison of randomisation scores according to task order (which was counterbalanced) produced two significant effects. When participants completed the static remember condition first, their Phi7 scores were less biased and their TPI scores were more biased specifically in the remember condition.

We next examined whether the marked values were differentially selected. With two marked remember values, participants preferentially chose those items in randomisation (M=11.7, SD=1.77), t(15)=3.75, p<.01, d=.938. This preference was also evident in the track condition (M=12.6, SD=1.42), t(15)=7.20, p<.001, d=1.8.

Discussion

Experiments 1A and 1B established that there is a disruptive effect on random number generation from remembering that two of the numbers are different, insofar as they elicit a hand tap when they are chosen. We assume that this concurrent task involves working memory, because participants must keep the two marked numbers in a raised state of activation. Indeed, there is evidence to support this insofar as the marked numbers are preferentially chosen as random responses.

The current experiment shows that increasing the working memory demands of the concurrent task, not by changing the marked values but instead the consequent actions, has a further disruptive effect on random generation. When participants need to represent and update target positions for their hand tap, random generation suffers. They produced more stereotyped response sequences, in terms of adjacent numbers and other paired combinations. The concurrent task itself was hard, and participants' made mistakes in selecting a location (returning to a previous position or omitting a position from the sequence). Thus, relative to the remember condition, both primary and concurrent task performance declined. The data support the contention that in selecting random numbers, participants draw upon working memory to represent and update task-relevant information, such that access to these representations can be usurped by the memory demands of the concurrent task. For example, participants might normally utilise information about selection history (i.e., which numbers and sequences have already been produced) to modulate current choices (see also Tune, 1964), but the tracking task degrades this process.

General Discussion

Across several datasets, random generation has been shown to be a highly demanding cognitive task that is very difficult to perform well. Interference occurs with additional requirements, provided these are sufficiently taxing. That is, distraction itself is not necessarily disruptive, but it can be if demands are sufficient. This finding has potential implications for the interpretation of random generation as an interference task. Performance does not change linearly with load; a simple task may be qualitatively different from a more complex one. It is also concluded that random generation draws on working memory. We have demonstrated that concurrent memory requirements, especially when these change, hamper sequence selection.

Results confirm one specific conclusion from Towse and Valentine (1997) that asking participants to suppress response values impairs the quality of random sequences. Yet they also indicate, contrary to previous conclusions, that marking response choices out for special treatment can produce an equivalent impairment. The dynamic memory requirements of random generation are underlined by the further effect obtained when participants need to remember and update spatial position information during the task. The results suggest that the impact of suppression requirements lie substantially with the need to keep in mind the particular inhibitory requirement (see Wegner, 1994), rather than the act of suppression *per se*.

The evidence from Experiment 2 supports findings that other random sequencing tasks, such as random interval generation impair both verbal short-term memory (e.g., Vandierendonck, De Vooght & Van der Goten, 1998) and spatial memory (e.g., Vandierendonck, Kemps, Fastame, & Szmalec, 2004). The present data extends these analyses by showing the converse result –a memory requirement (in this case with a spatial component) that affects randomisation. Thus, they all point to a functional overlap between memory and randomisation processes (see also Baddeley et al., 1998).

Such findings advance the case for regarding working memory as broadly relevant to random generation, but also demonstrate more specifically that memory processes are integral to executive task performance. This is especially true for memory representations that change, and thus require continuous updating and the inhibition of no-longer-relevant representations (e.g., Hasher, Lustig & Zacks, in press; Palladino, Cornoldi, De Beni & Pazzaglia, 2001), even though Experiment 1 shows that changing memory representations in the concurrent task are not necessary for randomisation to be disrupted. We suggest that response selection in randomisation incorporates both immediate and more distant past choices, and consideration of the preceding sequence can be regarded as a working memory function (see also Kareev, 1995). More generally, the findings are broadly compatible with (and link to) load theory, which argues that increases in working memory demand produce greater distraction effects (Lavie et al., 2004). In the context of the present results, it is the extent of the load and not its precise form (i.e., the requirement to remember or suppress) that is most relevant, and which produces non-linear effects.

While the experiments have demonstrated the particular impact of concurrent tasks, we do not claim to have employed, in the strict sense, a dual task approach. Although the remember instruction plainly produces a separate response, it is linked because it is contingent on particular randomisation behaviour. The tasks are therefore loosely related, and the additional activity completed intermittently. In a prototypical dual task situation, two separate tasks are formally independent, even though they may produce cross-talk (e.g., Baddeley et al., 1998; Vandierendonck et al., 2004). The broad compatibility in the results, despite differences in the exact nature of the concurrent task, reinforce the generality of the present conclusions.

The case for linking random generation paradigms to the working memory framework, in particular the concept of a central executive, was persuasively made by Baddeley (1986), by reviewing the then-extant data (e.g., Baddeley, 1966). While legitimate concerns can be expressed about exactly how to characterise this relationship, because for example EFs are heterogeneous (Miyake et al., 2000) and random generation itself is multiply determined (Towse, 1998; Towse & Houston-Price, 2001), the present data demonstrate how each concept has something to offer the other. The concept of working memory, especially the emphasis on active maintenance of representations, clearly helps to provide a general explanatory cloak by which to understand randomisation. At the same time, the need in random generation to maintain representations of legitimate responses as well as integrate previous with current choices, exemplifies the subtle functions of working memory.

Inhibition is very important in randomisation, yet requires modulation via working memory of active representations that extend in time. Inhibition, as part of selection from among competing response candidates, involves the assertion of control over representations in a way that may correspond to dynamic memory tracking or updating. Thus, the current data support the more generic contention that the active maintenance and transformation of information is closely tied into executive functioning. So while the random generation task has been characterised as a "blunt instrument whose detailed theoretical interpretation is at best equivocal" (Baddeley et al., 1998, p.849), such a cautionary approach can be complemented by the present positive view of what may be achieved, with the current paradigms offering an insight into the interface between memory and cognition.

References

Baddeley, A. (2000). The episodic buffer: a new component of working memory. *Trends in Cognitive Sciences, 4*(11), 417-423.

Baddeley, A. D. (1966). The capacity for generating information by randomization. *Quarterly Journal of Experimental Psychology, 18*, 119-129.

Baddeley, A. D. (1986). *Working Memory*. Oxford: Clarendon Press.

Baddeley, A. D., Emslie, H., Kolodny, J., & Duncan, J. (1998). Random generation and the executive control of working memory. *Quarterly Journal of Experimental Psychology, 51A*(4), 819-852.

Bar-Hillel, M., & Wagenaar, W. A. (1991). The perception of randomness. *Advances in Applied Mathematics, 12*, 428-454.

Brugger, P. (1997). Variables that influence the generation of random sequences: An update. *Perceptual and Motor Skills, 84*, 627-661.

Brugger, P., Monsch, A. U., Salmon, D. P., & Butters, N. (1996). Random number generation in dementia of the Alzheimer type: A test of frontal executive functions. *Neuropsychologia, 34*(2), 97-103.

Duncan, J., Emslie, H., Williams, P., Johnson, R., & Freer, C. (1996). Intelligence and the frontal lobe: The organisation of goal directed behavior. *Cognitive Psychology, 30*, 257-303.

Ginsburg, N., & Karpiuk, P. (1994). Random generation: analysis of the responses. *Perceptual and Motor Skills, 79*, 1059-1067.

Hasher, L., Lustig, C., & Zacks, R. (in press). Inhibitory Mechanisms and the Control of Attention. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake & J. N. Towse (Eds.), Variation in Working Memory. New York: Oxford University Press.

Heuer, H., Kohlisch, O., & Klein, W. (2005). The effects of total sleep deprivation on the generation of random sequences of key-presses, numbers and nouns. *Quarterly Journal of Experimental Psychology, 58A*(2), 275-307.

Hitch, G. J., & Baddeley, A. D. (1976). Verbal reasoning and working memory. *Quarterly Journal of Experimental Psychology, 28*, 603-621.

Kareev, Y. (1995). Through a narrow window: Working memory capacity and the detection of covariation. *Cognition, 53*, 263-269.

Knoch, D., Brugger, P., & Regard, M. (2005). Suppressing versus Releasing a Habit: Frequency-dependent Effects of Prefrontal Transcranial Magnetic Stimulation. Cerebral Cortex, 15, 885-887.

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, 133(3), 339-354.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology, 41*, 49-

100.

Palladino, P., Cornoldi, C., De Beni, R., & Pazzaglia, F. (2001). Working memory and updating processes in reading comprehension. Memory & Cognition, 29, 344-354.

Reed, J., & Johnson, P. (1994). Assessing implicit learning with indirect tests: Determining what is learned about sequence structure. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 20(3), 585-594.

Towse, J. N. (1998). On random generation and the central executive of working memory. *British Journal of Psychology, 89*(1), 77-101.

Towse, J. N., & Houston-Price, C. M. T. (2001). Reflections on the concept of the central executive. In J. Andrade (Ed.), *Working memory in perspective* (pp. 240-260). Hove, England: Psychology Press.

Towse, J. N., & Mclachlan, A. (1999). An exploration of random generation among children. *British Journal of Developmental Psychology, 17*(3), 363-380.

Towse, J. N., & Neil, D. (1998). Analyzing human random generation behavior: A review of methods used and a computer program for describing performance. *Behavior Research Methods, Instruments & Computers, 30*(4), 583-591.

Towse, J. N., & Valentine, J. D. (1997). Random generation of numbers: A search for underlying processes. *European Journal of Cognitive Psychology, 9*(4), 381-400.

Tune, G. S. (1964). A brief survey of variables that influence random generation. *Perceptual and Motor Skills, 18*, 705-710.

Vandierendonck, A. (2000). Is judgment of random time intervals biased and capacity-limited? *Psychologial Research, 63*(2), 199-209.

Vandierendonck, A., De Vooght, G., & Van der Goten, K. (1998). Does random time interval generation interfere with working memory executive functions? *European Journal of Cognitive Psychology*, 10(4), 413-442.

Vandierendonck, A., Kemps, E., Fastame, M. C., & Szmalec, A. (2004). Working memory components of the Corsi blocks task. *British Journal of Psychology, 95*(1), 57-79.

Wegner, D. M. (1994). Ironic processes of mental control. *Psychological Review, 101*, 34-52.

Appendix 1

Your task is to produce numbers in a random order. I shall tell you shortly which numbers to use. To give you an idea of what the task requires, imagine you roll a fair die. Each side of the die is equally likely to be selected with every roll, and each roll is independent of the preceding ones. I would like you to attempt to produce a set of numbers as if you were simulating a fair die.

Your sequences will be recorded and analysed to measure how close you were to simulating a random sequence of numbers. For example, if you produce more adjacent number responses (for example ‘2-3’, ‘7-6) than would occur from a random source these patterns will be noted as being non-random. If you choose particular numbers too often or not often enough, this will also be detected. Often sequences are non random in that choices are do not repeat often enough (i.e. people don’t say the same number twice in succession often enough, or the repeat the number with just a single intervening item, etc.). Thus, the point is to try and make the sequence of numbers as unpredictable or as jumbled up as possible.

You will hear a series of tones at the rate of 1 per 1.5 seconds. Please produce a number each time you hear the computer give a signal, and continue until told to stop (this will be after 81 responses).

The most important part of the task is to keep pace with the tones, to give a number at the right time. Remember, there is no ‘right’ or ‘wrong’ answer to give, so there is no need to be anxious. Just try to produce a random sequence of numbers as best you can.

Appendix 2

Explanations of randomisation indices (for a more extensive treatment, see Ginsburg & Karpiuk, 1994; Towse & Neil, 1998).

RNG – An assessment of the distribution of all response pairs in the sequence. Values lie between 0 and 1, and the RNG score rises as particular pair combinations are repeated. For example, if certain stereotyped sequences are repeatedly used such as adjacent values “3, 4” or even numbers “4, 6”, this will be reflected in higher RNG scores. RNG scores necessarily vary with the number of response permutations and the sequence length. Formally, RNG scores are:

$$\text{RNG} = \frac{\sum n_{ij} \log n_{ij}}{\sum n_i \log n_i}$$

where n_{ij} is the frequency count from each cell in the matrix of possible combinations and n_i represents the frequency of occurrence of response *i*.

A – this Adjacency measure involves calculation of specific paired combination values. Whereas RNG scores measure all possible response pairs, the A score reflects the percentage of adjacent response values (e.g., “1, 2” or “4, 3) in the sequence, and is formally:

$$\frac{\text{number of adjacent pairs}}{\text{number of response pairs}} \times 100$$

TPI – The Turning Point Index measures the number of times responses involve a change between ascending (e.g. “1, 4, 9”) and descending (e.g. “7, 5, 2”) sequences (the count of local peaks and troughs in a time-series plot), and then compares this with an expected value;

$$TPI = \frac{\text{number of observed turning points}}{\frac{2}{3}(N-2)} \times 100$$

where N is the sequence length. A TPI value less than 100 indicates fewer changes in the ordinal progression of responses than would be expected of a random sequence, and a value greater than 100 indicates more changes than would be expected. Thus, a low TPI value would suggest a response strategy involving runs of ordinal numbers.

R – This is a measure of the distribution of response frequencies. That is, it indicates whether some responses are produced more often than others. A (minimum) value of 0 indicates responses are used equally, whereas a (maximum) value of 100 indicates a single value is used for all responses. Thus larger numbers indicate more bias in the response set. R is calculated as

$$\mathrm{R} = \left(1 - \frac{\log_2 N - \frac{1}{N}\left(\sum n_i \log_2 n_i\right)}{\log_2 a}\right) \times 100$$

where a is the number of response alternatives (other symbols as used above).

Phi2 and Phi7 – The Phi index measures response repetitions over different sequence lags, with a range of scores between -100 (repetitions occur less frequently than would be expected from random sequences) and 100 (repetitions occur more frequently than would be expected from random sequences). The Phi2 measure focuses on immediate repetitions of a response, the Phi7 measure assesses repetitions that occur after five intervening responses have been produced. The Phi Index is;

$$\phi = \sqrt{\frac{\chi^2}{T}} \times 100$$

where a chi-squared statistic is determined by comparing observed repetitions against expected values (given known lower-order frequency counts), and T is an artificial sequence length value obtained after transforming the original sequence into binary strings.

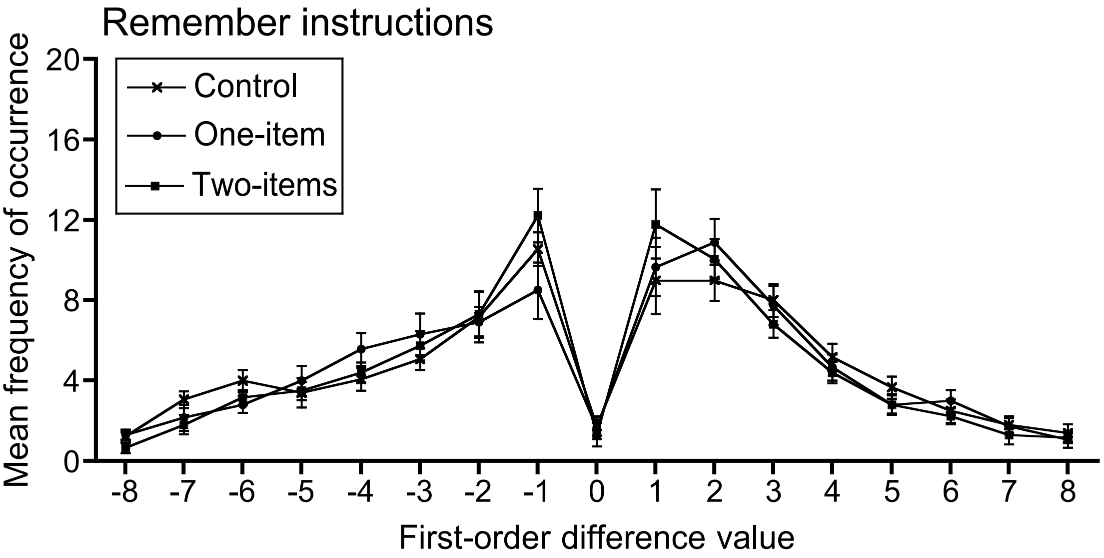
Table 1. Adherence to response requirements in Experiment 1

	Control	One-item	Two-items
Experiment 1A			
Response rate failures (memory)	1.42 (2.94)	2.17 (5.08)	2.17 (3.35)
Response rate failures (suppression)	3.33 (5.76)	4.67 (6.79)	7.25 (7.02)
Response violations (memory)	0.08 (0.29)	0.50 (0.67)	0.75 (0.87)
Response violations (suppression)	0.25 (0.62)	1.25 (1.91)	3.17 (3.30)
Experiment 1B			
Response rate failures (memory)	0.25 (0.45)	0.17 (0.58)	0.50 (1.17)
Response rate failures (suppression)	0.00 (0.00)	0.50 (0.67)	1.17 (1.11)
Response violations (memory)	0.08 (0.29)	0.17 (0.38)	0.42 (0.51)
Response violations (suppression)	0.25 (0.61)	1.00 (0.93)	3.08 (3.12)

Table 2. Adherence to response requirements in Experiment 2

	Remember	Tracking
Response rate failures	1.56 (2.19)	0.81 (2.04)
Response violations	1.38 (2.87)	0.38 (0.62)
Concurrent tracking errors		8.00 (3.71)

Figure 1. Mean first-order difference scores (and standard errors) in Experiment 1A. Plot points depict performance under control conditions, and where a concurrent task relates to one (single) or two (double) items. Upper panel shows data from the remember condition, lower panel shows data from the suppression condition.



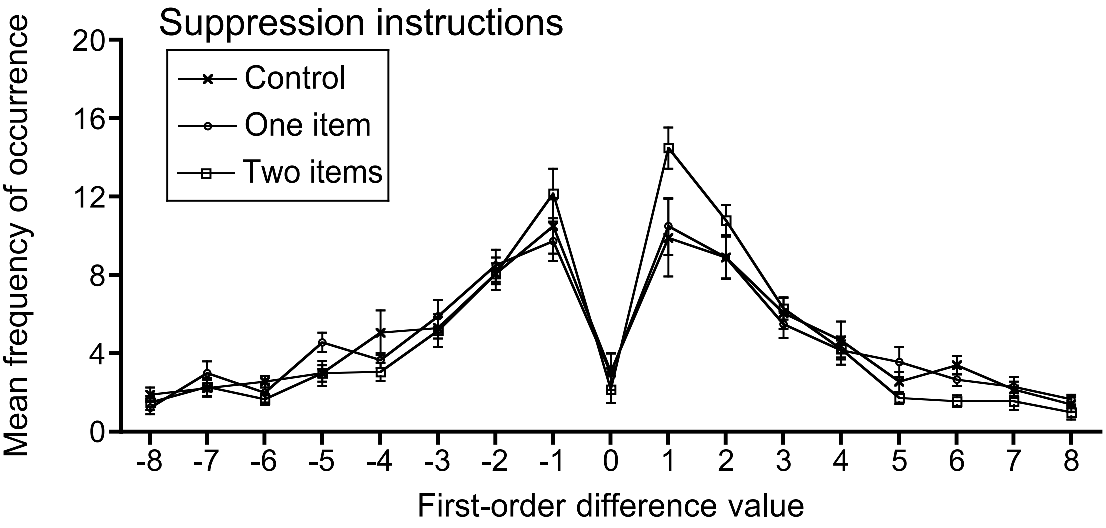


Figure 2. Mean first-order difference scores (and standard errors) in Experiment 1B. Plot points depict performance under control conditions, and where a concurrent task relates to one (single) or two (double) items. Upper panel describes performance in the remember condition, lower panel in the suppression condition.

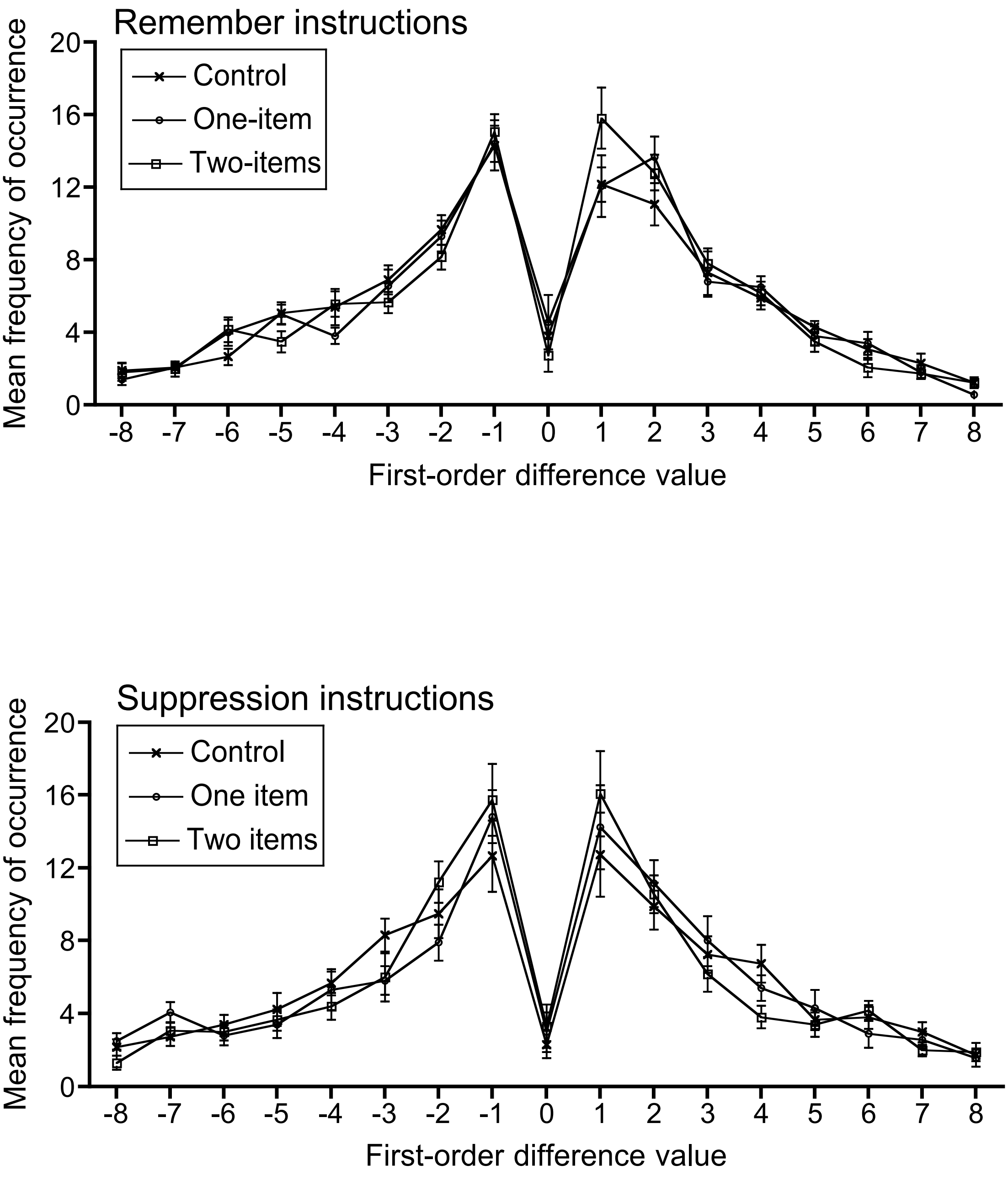


Figure 3. Mean first order difference scores (and standard errors) in Experiment 2.