BRIEF REPORT

The role of placekeeping ability in fluid intelligence

David Z. Hambrick · Erik M. Altmann

Abstract The question of what underlies individual differences in general intelligence has never been satisfactorily answered. The purpose of this research was to investigate the role of an executive function that we term *placekeeping ability*—the ability to perform the steps of a complex task in a prescribed order without skipping or repeating steps. Participants completed a newly developed test of placekeeping ability, called the UNRAVEL task. The measure of placekeeping ability from this task (error rate) predicted a measure of fluid intelligence (Raven's Advanced Progressive Matrices score), above and beyond measures of working memory capacity, task switching, and multitasking. An existing model of Raven's performance suggests that placekeeping ability supports the systematic exploration of hypotheses under problem-solving conditions.

Keywords Cognitive and attentional control · Executive control · Individual differences · Memory capacity

Research has established that scores on diverse tests of cognitive ability correlate positively with each other (see Jensen, 1999), implying the existence of a general factor of intelligence, which has been termed "psychometric *g*." Nevertheless, the question of what *g* is at the level of the cognitive system remains open. One approach to answering this question—the *cognitive-correlates* approach—involves developing measures to capture specific cognitive processes, and then testing for correlations of these measures with scores on highly *g*-loaded tests. If some cognitive process is important for general

intelligence, then it should correlate positively with individual differences in *g*.

There has been interest in understanding *g* in terms of its relations with *executive functioning* (e.g., Hambrick et al., 2011), the suite of mental operations that coordinate and supervise other mental operations in the service of performing some task (Banich, 2004). One component of executive functioning is the ability to remember the goal of a task (Duncan, Emslie, Williams, Johnson, & Freer, 1996). The goal of making an omelet, for example, organizes the actions of cracking and beating the eggs, heating the skillet and melting butter, pouring the eggs into the skillet, flipping the eggs once they are cooked on one side, and so on.

A related component of general intelligence seems to be the ability to proceed through a sequence of mental or physical actions systematically. Duncan et al. (1996; see also Duncan, 2010) made this connection, and Carpenter, Just, and Shell (1990) argued that performance on Raven's Advanced Progressive Matrices—generally considered the gold-standard measure of psychometric *g*—depends in particular on the ability to serialize and systematically explore a series of subgoals.

In the present study, we explored the relationship between *fluid intelligence* and what we refer to as *placekeeping ability*. Fluid intelligence, or Gf, is the ability to solve novel problems and adapt to new situations. Placekeeping is the ability to accurately maintain one's place in a task sequence, which seems to play a role in a wide variety of everyday tasks. Although placekeeping failures are inconsequential in some situations (e.g., heating the skillet before beating the eggs), in other situations they can be catastrophic—as, for example, when a soldier neglects to make sure a gun is unloaded before cleaning it, or a nurse forgets having administered an insulin injection to a patient.

Existing tests of placekeeping ability are limited in their usefulness for investigating individual differences, because

D. Z. Hambrick (✉) · E. M. Altmann (✉)
Department of Psychology, Michigan State University, East Lansing, MI 48824, USA
e-mail: hambric3@gmail.com
e-mail: ema@msu.edu

errors are sparse and administration times are long. For example, in routine tasks such as making coffee, error rates in healthy populations are only about 2 % (Botvinick & Bylsma, 2005; Cooper & Shallice, 2000). Thus, to study placekeeping errors in coffee making, Botvinick and Bylsma found that they had to have each participant make 50 cups of actual coffee in order to produce enough data to analyze. In event counting (e.g., Carlson & Cassenti, 2004), errors are again infrequent enough that the data are generally total counts reported after a series of events, which are compared to actual counts in order to compute a deviation. However, total counts mask errors that offset one another, as when someone fails to count one event but double-counts another.

To generate richer data on placekeeping, we have recently developed a task that produces a relatively large number of placekeeping errors in a relatively short amount of testing time (Altmann, Trafton, & Hambrick, 2014). The task is defined by the acronym UNRAVEL. Each letter in the acronym identifies a step in a cyclical ("looping") procedure, and the letter sequence stipulates the order in which the steps are to be performed. That is, the U step is performed first, the N step second, the R step third, and so forth, and the participant returns to the U step following the L step. The participant is interrupted at random points and must perform a transcription task, before resuming UNRAVEL with the next step.

The task is illustrated in Fig. 1 in terms of two sample stimuli. Each sample stimulus includes two characters; of these characters, one is a letter and one a digit, one is presented either underlined or italicized, one is colored red or yellow, and one is located outside of a gray box. Each step requires a two-alternative forced choice related to one feature, and the letter of the step mnemonically relates to the choice rule: The U step involves deciding whether the formatted character is underlined or italicized, the N step whether the letter is near to or far from the start of the alphabet, the R step whether the colored character is red or yellow, the A step whether the character outside the box is above or below, the V step whether the letter is a vowel or a consonant, the E step whether the digit is even or odd, and the L step whether the digit is less or greater than 5.

In two previous experiments (Altmann et al., 2014), we found that the UNRAVEL task not only produced a reasonably high error rate, but that the error rates that were produced varied widely across participants, ranging from 0 % to 12 % in the first experiment, and from 0 % to 18 % in the second experiment. Furthermore, internal-consistency reliability was reasonably high (αs = .54 and .72, respectively), suggesting that the measure is suitable for use in individual-differences research.

In the present study, we investigated the relationship between placekeeping in the UNRAVEL task and Gf as measured by Raven's Advanced Progressive Matrices, the gold-standard assessment of Gf (Jensen, 1999). We also measured three other executive functions—working memory capacity,

task switching, and multitasking—and collected college entrance exam (ACT) score as an index of crystallized intelligence (Gc). Each of these other factors has been demonstrated or hypothesized to relate to Gf (see Hambrick et al., 2011). Moreover, UNRAVEL could measure any of these factors to some degree, and this relationship could account for any correlation that we observe between UNRAVEL error rate and Raven's score. To address this possibility, we performed regression analyses to investigate whether UNRAVEL error rate would predict Gf above and beyond measures of these other factors. Finally, to further test the reliability of UNRAVEL performance, we administered two sessions of the task on separate days, allowing us to evaluate test–retest as well as internal-consistency reliability.

Method

Participants

The participants were undergraduate students recruited through the subject pool at Michigan State University. A total of 158 participants contributed data on all of the measures. Of these, we excluded 21 who had below-threshold accuracy on the UNRAVEL or task-switching tasks (see the task descriptions, below), leaving 137. Of these, we excluded an additional five participants who had outlying scores, where an outlying score for a measure was defined as one differing by more than 3.5 standard deviations from the mean for that measure. Data from the remaining 132 participants were submitted to analysis.

The average self-reported ACT score for the sample was 26 (SD =3.3), with a range from 17 to 34. The national average is approximately 21 (SD =4).

Procedure

The study took place in three sessions on separate days, with each session lasting approximately 1 h. We aimed to have all three sessions occur within a 5-day period, but participant scheduling constraints led to some variability (Session 3 – 1 range = 2 to 13 days, M = 4.4, SD = 1.4). In Session 1, participants reported their ACT score, performed four blocks of UNRAVEL, and completed the Raven's test; in Session 2, they performed a task-switching procedure, a test of working memory capacity, and a test of multitasking ability; and in Session 3, they performed four more blocks of UNRAVEL. Participants were tested individually.

Materials

ACT Participants reported their total score on the ACT college admissions exam. The total score is the average of the scores

(a) Sample stimuli for UNRAVEL task:

Stimulus 1   Stimulus 2

X

1

8

A
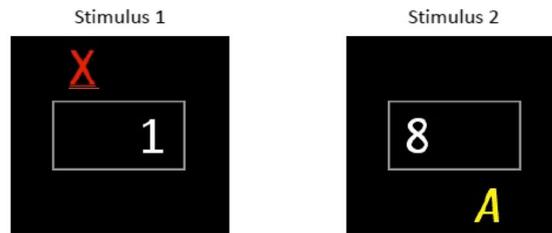
(b) Choice rules and candidate responses for UNRAVEL task, and responses to the stimuli in (a):

Step	Candidate responses		Choice rules	Responses to sample stimuli	
				Stimulus 1	Stimulus 2
U	u	i	character is Underlined or in Italics	u	i
N	n	f	letter is Near to or Far from start of alphabet	f	n
R	r	y	character is Red or Yellow	r	y
A	a	b	character is Above or Below the box	a	b
V	v	c	letter is Vowel or Consonant	c	v
E	e	o	digit is Even or Odd	o	e
L	l	m	digit is Less than or More than 5	l	m

(c) Sample stimulus for interrupting task:

Please type in code #1:   vneombrfyaucli

vne   [Return]

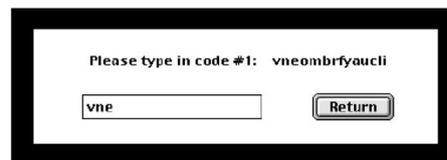**Fig. 1** (**a**) Two sample stimuli for the UNRAVEL task (the X is presented in red, and the A in yellow). (**b**) Response mappings for the UNRAVEL task, along with responses for the two sample stimuli shown in panel a. (**c**) Sample stimulus for the interrupting task, shown when the participant has typed three letters of the first of two 14-letter "codes."

on the four required sections of the test: English, Math, Reading, and Science.

UNRAVEL The procedure was as follows (adapted from Altmann et al., 2014): The Session 1 administration began with a step-by-step introduction to the UNRAVEL step sequence. The introduction emphasized the acronym, showing how each step in turn corresponded to a constituent letter, and then presented a summary screen showing the choice rules for each step and the letters spelling out the word. A different pair of response keys was used for each choice rule (see Fig. 1, panel B), so that we could infer from each response the step that the participant thought was correct. After the introduction, to ensure that the participants understood the task, they performed 16 trials during which the computer required participants to make the correct response on each trial before allowing him or her to move on. This 16-trial sequence was interrupted twice, to illustrate for participants that they were supposed to pick up after an interruption where they had left off. The experimenter remained present during this period, to help if necessary. A sheet of paper with the choice rules for the UNRAVEL sequence remained visible to the side of the computer throughout the entire session.

In preparation for the experimental phase of the session, participants were reminded to "please try to keep your place in the UNRAVEL sequence" and to "please try to pick up in the sequence where you left off" after an interruption. The experimental phase consisted of four blocks, each with ten interruptions. The number of trials between interruptions was randomized, with a mean of six trials between interruptions, and there was one run of trials before the first interruption, so that each block contained about 66 trials. A session took about 30 min to complete.

Sequence errors were coded with respect to the previous step. For example, if steps U, R, and A were performed in succession, R would be a sequence error, because N was skipped, but A would be correct, because A follows R in the UNRAVEL sequence. After each block, the participant was given his or her score, computed as the percentage of trials during that block for which the step and response were both correct. If the score was above 90 %, the participant was asked to go faster. If the score was below 70 %, the participant was asked to be more accurate, and that block was excluded from the analysis. The threshold accuracy for a participant was defined as 70 % or higher on at least three of the four blocks in a session. Participants who fell below this threshold in either session were judged not to be following the accuracy instruction and were excluded from the analysis.

For both Session 1 and Session 3, the skewness and/or kurtosis statistics (Ms = 1.18 and 1.58, respectively) were somewhat high for UNRAVEL error rate. However, the correlations of the untransformed and arcsine-root-transformed variables with Raven's were almost identical (mean difference in $rs < .02$). Therefore, we used the untransformed variables in all of the analyses reported next, so that the results can be interpreted in the original units.

Raven's Advanced Progressive Matrices We used the 18 odd-numbered items of Raven's Advanced Progressive Matrices to measure Gf. Each item was displayed on a computer screen and consisted of a series of patterns, arranged in three rows and three columns. The pattern in the lower right was always missing, and the participant's task was to choose the alternative that logically completed the series. The score was the number correct. Cronbach's alpha for Raven's score was .65.

Task switching Participants alternated between two simple two-alternative forced choice tasks. The tasks, task cues, stimuli, and response mappings are described in Altmann (2004).

The participant performed the same task on a run of consecutive trials, with the task potentially switching between runs. A task cue was presented at the start of the run to indicate the task for that run. The cue appeared for 100 ms and was followed by a blank screen for 100 ms. After the blank screen, the run of trials began. Each run consisted of an average of five trials.

After the last trial of a run, the screen went blank for a short interval, after which the task cue for the next run appeared. The duration of this blank interval, which we refer to as the *response–cue interval*, or RCI, was either 100 or 800 ms, randomly selected.

Participants performed ten test blocks of 30 runs each. At the end of each block, the participant had a chance to rest and was told the score for that block. If the score for that block was 100 %, participants were asked to see if they could go faster; if it was below 90 %, they were asked to be more accurate. The threshold accuracy for a participant was defined as 90 % or higher, averaged over the test blocks. Participants who fell below this threshold were judged not to be following the accuracy instruction and were excluded from the analysis.

The measure of interest was switch cost, computed as the response latency on the first trial of a switch run minus the response latency on the first trial of a repeat run (after Altmann, 2004). On a switch run, the task cue differed from that for the previous run, whereas on a repeat run, the task cue was the same as for the previous run. A datum was the median response latency in a given cell of the design for a given participant. We used the correlation of switch cost across the two levels of RCI as the estimate of internal-consistency

reliability. The correlation was .45, indicating reasonably high reliability for a derived score (see, e.g., Ettenhofer, Hambrick, & Abeles, 2005).

Working memory capacity (WMC) We used the automated version of the operation span task (Unsworth, Heitz, Schrock, & Engle, 2005) to measure working memory capacity. The participant's task was to solve a series of math equations, and after solving each equation, to remember a letter. After from three to seven equation–letter trials, participants were prompted to recall the letters in the order in which they had been presented. The score was the number of letters recalled in the correct order (maximum =75). Cronbach's alpha for the operation span task was .70.

Multitasking We used the SynWin paradigm (Elsmore, 1994) to measure multitasking ability. The task includes four subtasks—arithmetic, memory search, auditory monitoring, and visual monitoring—presented in the four quadrants of the computer screen for concurrent performance. In each subtask, points are awarded for a correct answer and deducted for incorrect answers; the total score is the sum of the subtask scores. Participants performed one 10-min session of SynWin. We only administered one block of SynWin, and so could not compute reliability, but Cronbach's alpha had been calculated as .89 in another study in which this task was used (Hambrick, Oswald, Darowski, Rench, & Brou, 2010).

Results

There was a considerable amount of variability in the overall UNRAVEL error rate: from 0 % to 10 % in Session 1 (M =3.1 %, SD =2.4 %), and from 0 % to 14 % in Session 3 (M =3.2 %, SD =2.6 %). To estimate internal-consistency reliability, we computed Cronbach's alpha, using the error rate from the four blocks (averaged across Sessions 1 and 3) as the variables. The alpha value was .84. To estimate test–retest reliability, we computed the correlation between the error rate for Session 1 and for Session 3. The correlation was .58 ($p < .01$) and was unchanged after statistically controlling for the number of days between Session 1 and Session 3 ($pr = .58$, $p < .01$). These reliability estimates are at least as high as those typically found for tests of executive functioning (e.g., Ettenhofer et al., 2005).

The correlations across tasks are displayed in Table 1 (values above the diagonal are corrected for measurement error). UNRAVEL error rate correlated significantly and negatively with Raven's score ($r = -.45$, $p < .01$). Thus, participants who performed accurately in UNRAVEL tended to also perform well in Raven's. Moreover, UNRAVEL error rate correlated more strongly with Raven's score than did

Table 1 Correlation matrix

Variable	1	2	3	4	5
1. UNRAVEL	(.84)	–.18	.02	–.09	–.61
2. WMC	–.14	(.70)	–.12	.28	.12
3. Task switching	.01	–.07	(.45)	–.25	–.02
4. Multitasking	–.08	.22*	–.16	(.89)	.43
5. Raven's APM	–.45**	.08	–.01	.33**	(.65)

$N = 132$. Correlations above the diagonal are corrected for measurement error by using the reliability estimates along the diagonal. The formula for the correction is $\hat{r}_{xy} = r_{xy}/\sqrt{r_{xx} * r_{yy}}$, where r_{xy} is the uncorrected correlation between two variables x and y, and r_{xx} and r_{yy} are their reliabilities (Hunter & Schmidt, 1990). * $p < .05$, ** $p < .01$.

WMC ($r = .08$), task switching ($r = –.01$), and multitasking ($r = .33$, $p < .01$). Although UNRAVEL correlated significantly with ACT score ($r = –.18$, $p < .05$), which is a measure of crystallized intelligence (Gc), this correlation was reduced to near zero after statistically controlling for Raven's score ($pr = –.02$). Thus, placekeeping ability appears to contribute to individual differences in Gf, but not in Gc.

The correlations among the other variables were generally consistent with those found in previous work. One example is the near-zero correlation ($r = –.07$) between operation span and task switching (e.g., Miyake et al., 2000, reported $–.10 < r < .10$ for correlations between operation span and measures of task switching). The correlation between operation span and Raven's score ($r = .08$) was lower than expected, perhaps because the participants were all college students at a moderately selective university (i.e., a restriction of range), and perhaps also because operation span and Raven's were administered on separate days.

Regression analyses

To more formally investigate the predictive validity of UNRAVEL, we regressed Raven's score onto UNRAVEL error rate before and after statistically controlling for the measures of WMC, task switching, and multitasking. These results are displayed in Table 2. UNRAVEL error rate accounted for 20.3 % ($\beta = –.451$, $p < .01$) of the variance in Raven's score when entered alone as a predictor variable, and for 18.3 % ($\beta = –.433$, $p < .01$) of the variance when entered after controlling for the other measures. Thus, we uncovered evidence for the incremental validity of UNRAVEL, since controlling for the other measures reduced its contribution to Raven's score by only about 10 %.

We also regressed Raven's score onto WMC, task switching, and multitasking before and after statistically controlling for UNRAVEL error rate. WMC, task switching, and multitasking together accounted for 11.1 % ($p < .01$) of the variance in Raven's score, and for 9.1 % of the variance when entered after statistically controlling for UNRAVEL. Thus, as

a set, these three predictors accounted for less variance in Raven's performance than did UNRAVEL. Individually, only multitasking ability made a statistically significant unique contribution to the prediction of Raven's score ($\beta = .334$ vs. .312 before vs. after controlling for UNRAVEL, $ps < .01$). Overall, then, UNRAVEL was a stronger predictor of Raven's score than were the other executive functions.

Structural equation modeling

We performed structural equation modeling (SEM) to estimate the relationship between placekeeping ability and Raven's performance at the level of latent variables. For placekeeping ability, the indicators were two measures that contribute to the overall error rate: the error rate on trials immediately following an interruption (*postinterruption error rate*), and the error rate on all remaining trials (*baseline error rate*). If placekeeping ability is a general ability, then these error measures should correlate moderately with each other and contribute to a latent factor, even though the processing contexts are somewhat different, as reflected in the large difference in the frequencies of the two types of errors ($M_{baseline} = .01$, $SD = .01$, vs. $M_{postinter} = .12$, $SD = .10$; see also Altmann et al., 2014). For Raven's, the indicators were three item parcels, reflecting the numbers of problems correctly solved among Items 1, 4, 7, 10, 13, and 16 (Parcel 1); Items 2, 5, 8, 11, 14, and 17 (Parcel 2); and Items 3, 6, 9, 12, 15, and 18 (Parcel 3).

The baseline and postinterruption error rates correlated moderately with each other ($r = .46$, $p < .001$), and, as is shown in Fig. 2, both had strong positive loadings on a Placekeeping Ability factor. This pattern is consistent with a placekeeping mechanism that plays roles both in advancing from one step to the next generally and in resuming the task at the correct location after an interruption. The pattern is actually quite robust, given that the correlation between the two indicators would presumably be attenuated by individual differences in strategies for resuming a task after an interruption (Trafton, Altmann, Brock, & Mintz, 2003).

The baseline error rate and postinterruption error rate also correlated very similarly with Raven's score ($r_{baseline} = –.38$, $r_{postinter} = –.39$, $ps < .001$). As is shown in Fig. 2, the Placekeeping Ability factor strongly predicted a Raven's factor ($\beta = –.68$), accounting for 46.6 % of the variance. The overall model fit was excellent, $\chi^2(4) = 3.32$, $p = .51$, CFI = 1.00, NFI = .98, RMSEA = .00.

Discussion

Placekeeping is an important form of executive function that plays a role in any kind of task performance that involves

Table 2 Results of regression analyses predicting Raven's Advanced Progressive Matrices scores

Predictor variable(s)	R^2	F	df
UNRAVEL			
Alone	.203	33.17**	1, 130
After control of WMC, TS, and MT	.183	32.99**	1, 127
WMC, TS, and MT			
Alone	.111	5.33**	3, 128
After control of UNRAVEL	.091	5.46**	3, 127

$N = 132$. WMC, working memory capacity; TS, task switching; MT, multitasking ability. $^* p < .05$, $^{**} p < .01$.

sequential constraints. And yet, placekeeping has received relatively little attention from cognitive researchers and, to our knowledge, has been entirely neglected in individual-differences research.

The present results establish the UNRAVEL task as a good candidate for measuring placekeeping ability. Error rate showed both internal-consistency reliability and test–retest reliability, indicating that performance on this task reveals stable individual differences. Moreover, the task shows predictive validity for the criterion measure of Raven's Advanced Progressive Matrices, the gold standard of general fluid intelligence tests. Finally, the relationship between UNRAVEL error rate and Raven's score could not be accounted for by the other factors we tested that have been hypothesized to explain variation in cognitive ability.

What cognitive mechanisms might be shared between UNRAVEL and Raven's performance? Carpenter, Just, and Shell (1990) developed cognitive simulations of Raven's performance that help us address this question, including one with a "goal management" module that carried out basic placekeeping operations. Behavioral evidence collected by Carpenter et al. (1990) suggested that people perform Raven's by serially evaluating hypotheses for rules that relate the different entries of a Raven's matrix. Testing a given hypothesis involves examining visual features and comparing them across matrix entries, to see whether they follow the hypothesized pattern. The model evaluated hypotheses

serially, which allowed it to focus only on those visual features relevant to the current goal. A hypothesis in Raven's maps well to a step in UNRAVEL, in that it governs which visual features are relevant to attend to. For both UNRAVEL and Raven's, then, being able to focus on one goal or step at a time out of several alternatives helps constrain the attentional selection of perceptual features, and may play a similar role in other supporting operations such as response selection.

The goal management module also implemented a backtracking function similar to resuming a task sequence after an interruption. If a hypothesis turned out to be incorrect because some feature did not match the pattern, the system had to select the next hypothesis in line, without returning to ones that it had already rejected, and without skipping any that might have turned out to be correct. Hypotheses take on the order of tens of seconds to evaluate (Carpenter et al., 1990), which is about as long as an interruption in the UNRAVEL task (in our data, $M = 20.3$ s, $SD = 6.0$ s, across sessions). Thus, the operations involved in testing a hypothesis served as an interruption, in the sense that they shifted the focus of attention away from the larger plan for exploring the space to the details of a given hypothesis. In this sense, problem space search involves an interruption anytime the system descends to a lower level of the goal hierarchy and, after achieving some number of subgoals, has to return to the higher level and set the next goal. Indeed, Carpenter et al. (1990) reported a high correlation [$r(43) = .77$] between accuracy on Raven's and accuracy on the Tower of Hanoi, a puzzle task often used to study the processing of goals and subgoals in problem solving (see, e.g., Patsenko & Altmann, 2010). Thus, interruptions in the UNRAVEL task arguably capture a general aspect of cognitive processing under novel conditions.

One important advantage of the UNRAVEL task in relation to puzzle tasks such as Raven's and the Tower of Hanoi is that it is reusable, rather than "one-shot." Thus, for research or diagnostic purposes, UNRAVEL could be administered repeatedly to the same participants in order to assess aptitude–treatment interactions, effects of environmental stressors, or effects of cognitive aging. In future work, we plan to
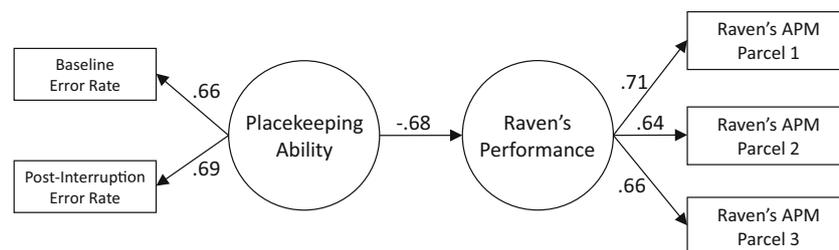

Fig. 2 Structural equation model with Placekeeping Ability factor predicting performance on Raven's Advanced Progressive Matrices (APM). Beta values are shown for each of the path arrows. All paths are statistically significant ($ps < .001$)

investigate other reusable tasks, such as event counting (Carlson & Cassenti, 2004), as candidate indicators of placekeeping, with the goal of further investigating the relationship between placekeeping ability and Gf at the latent-variable level.

References

Altmann, E. M. (2004). Advance preparation in task switching: What work is being done? *Psychological Science, 15,* 616–622. doi:10.1111/j.0956-7976.2004.00729.x

Altmann, E. M., Trafton, J. G., & Hambrick, D. Z. (2014). Momentary interruptions can derail the train of thought. *Journal of Experimental Psychology: General, 143,* 215–226. doi:10.1037/a0030986

Banich, M. T. (2004). *Cognitive neuroscience and neuropsychology* (2nd ed.). Boston, MA: Houghton Mifflin.

Botvinick, M. M., & Bylsma, L. M. (2005). Distraction and action slips in an everyday task: Evidence for a dynamic representation of task context. *Psychonomic Bulletin & Review, 12,* 1011–1017. doi:10.3758/BF03206436

Carlson, R. A., & Cassenti, D. N. (2004). Intentional control of event counting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1235–1251. doi:10.1037/0278-7393.30.6.1235

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

Cooper, R. P., & Shallice, T. (2000). Contention scheduling and the control of routine activities. *Cognitive Neuropsychology, 17,* 297–338. doi:10.1080/026432900380427

Duncan, J. (2010). *How intelligence happens*. New Haven, CT: Yale University Press.

Duncan, J., Emslie, P., Williams, R., Johnson, C., & Freer, C. (1996). Intelligence and the frontal lobe: The organization of goal-directed behavior. *Cognitive Psychology, 30,* 257–303. doi:10.1006/cogp.1996.0008

Elsmore, T. F. (1994). SYNWORK1: A PC-based tool for assessment of performance in a simulated work environment. *Behavior Research Methods, Instruments, & Computers, 26,* 421–426. doi:10.3758/BF03204659

Ettenhofer, M. L., Hambrick, D. Z., & Abeles, N. (2005). Reliability and stability of executive functioning in older adults. *Neuropsychology, 20,* 607–613. doi:10.1037/0894-4105.20.5.607

Hambrick, D. Z., Oswald, F. L., Darowski, E. S., Rench, T. A., & Brou, R. (2010). Predictors of multitasking performance in a synthetic work paradigm. *Applied Cognitive Psychology, 24,* 1149–1167. doi:10.1002/acp.1624

Hambrick, D. Z., Rench, T. A., Poposki, E. M., Darowski, E. S., Roland, D., Bearden, R. M., & Brou, R. (2011). The relationship between the ASVAB and multitasking in Navy sailors: A process-specific approach. *Military Psychology, 23,* 365–380. doi:10.1037/h0094762

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis*. Newbury Park, CA: Sage.

Jensen, A. (1999). *The g factor: The science of mental ability*. Westport, CT: Praeger.

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. doi:10.1006/cogp.1999.0734

Patsenko, E. G., & Altmann, E. M. (2010). How planful is routine behavior? A selective-attention model of performance in the Tower of Hanoi. *Journal of Experimental Psychology: General, 139,* 95–116. doi:10.1037/a0018268

Trafton, J. G., Altmann, E. M., Brock, D. P., & Mintz, F. E. (2003). Preparing to resume an interrupted task: Effects of prospective goal encoding and retrospective rehearsal. *International Journal of Human-Computer Studies, 58,* 583–603. doi:10.1016/S1071-5819(03)00023-5

Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods, 37,* 498–505. doi:10.3758/BF03192720