

Relating Individual Differences in Cognitive Ability and Strategy Consistency to Interruption
Recovery during Multitasking.

Hao Bai^a, Winston E. Jones^a, Jarrod Moss^a, and Stephanie M. Doane^a

^aMississippi State University

Correspondence concerning this article should be addressed to Hao Bai, Mississippi State University, P.O. Box 6161, Mississippi State, MS 39762. Email: hao.bai@skill-lab.org Phone number: 662-617-9067

Abstract

This research investigated the relationship between individual differences in cognitive ability and the ability to recover from interruptions during multitasking performance. The Abstract Decision Making (ADM) task was used to examine multitasking. This task required participants to sort multiple objects into bins while being unexpectedly interrupted. Participants also completed a battery of cognitive measures from which two ability factors were extracted, referred to as general-ability and multiple-event-tracking. Performance was assessed using measures of speed, errors, strategy consistency, and proportion of interruption resumption. The general-ability factor, which was correlated with working memory (WM), and the multiple-event-tracking factor affected both general ADM performance and interruption recovery differently. Moreover, consistent strategies were found to facilitate interruption recovery, an effect that was greater for lower general-ability individuals. The findings suggest that training individuals to use consistent strategies facilitates interruption recovery by alleviating WM load when interrupted, especially for low ability individuals.

Keywords: Multitasking; Individual Differences; Strategy Consistency; Interruptions

Relating Individual Differences in Cognitive Ability and Strategy Consistency to Interruption Recovery during Multitasking.

1. Introduction

Multitasking is a situation in which multiple tasks are performed under time pressure with the possibility of being interrupted by another task. For example, in an office-work environment, people have to work on multiple tasks such as writing documents, managing emails, and answering phone calls. Writing in a document might be interrupted by a phone call. When the phone call occurs, details about the writing task have to be held in memory if one is to resume the task quickly after dealing with the phone call. Depending on the duration and content of the phone call, information about the writing task may be maintained successfully in working memory (WM) or require encoding in long-term memory (LTM).

Large individual differences in the ability to handle multiple tasks have been observed in complex job environments such as air-traffic control or emergency dispatch (e.g., Joslyn & Hunt, 1998; Seamster, Redding, Cannon, Ryder, & Purcell, 1993). Understanding the cognitive basis of individual differences in multitasking ability has therefore become an important part of training and selection in these high-stakes jobs. However, knowledge about individual differences in multitasking performance remains limited, and therefore the present study was designed to examine the relationship between individual differences in cognitive ability and the ability to recover from interruptions.

This focus was chosen because interruptions are one of the key characteristics of multitasking, causing detriments to task performance when they occur (Burgess, 2000; Edwards & Gronlund, 1998; Monk, 2004). Previous studies have found disruptive effects of interruptions that increase task completion time and decrease performance accuracy (Edwards & Gronlund,

1998; Monk, 2004). Studies have also found a relationship between individual differences in cognitive ability and multitasking performance (e.g., Hambrick, Oswald, Darowski, Rench, & Brou, 2010; König, Bühner, & Mürling, 2005). However, none of these studies examined the relationship between cognitive ability differences and the ability to recover from interruptions.

One way in which individual differences might affect interruption recovery is through the adoption of different task strategies. Previous studies have found a relationship between strategy use and multitasking performance (e.g., Hambrick et al., 2010; Logie, Trawley, & Law, 2011), but these studies did not examine interruption recovery. A strategy refers to a sequence of actions in a context in which different sequences of actions are possible. The present study examines the relationship between individual differences in cognitive ability, strategy consistency, and interruption recovery in order to understand the role of cognitive resources in interruption recovery in a multitasking environment.

1.1 Interruption disruption and interruption recovery

The cognitive processes involved in interruption recovery have been examined in prior research (Altmann & Trafton, 2002; Cades, Boehm-Davis, Trafton, & Monk, 2011; Monk, Trafton, & Boehm-Davis, 2008; Speier, Valacich, & Vessey, 1999; Trafton, Altmann, Brock, & Mintz, 2003). A fundamental aspect of interrupted task performance is the suspension and resumption of goals (Monk et al., 2008). To explain how goals are suspended and resumed, Altmann and Trafton (2002) proposed the memory for goals model that fits reaction time and errors from a task that required frequent suspension and resumption due to interruptions. According to the memory for goals model, interruption disruptions arise from memory processes for encoding, maintaining, and retrieving both the goal and task state just prior to, during, and

after an interruption (Altmann & Trafton, 2002; Monk et al., 2008; Trafton et al., 2003). The goal and task states of the interrupted task can be viewed as a problem state (Borst, Taatgen, & Van Rijn, 2010; Salvucci & Taatgen, 2011). A problem state is the information about a task and the status of the task that needs to be maintained to resume work on the task. When switching between interrupted and interrupting tasks, a person must replace the problem state of one task with that of the other.

The use of WM and retrieval in memory for goals model is consistent with the ACT-R cognitive architecture (Anderson, 2007). In ACT-R, the immediately accessible contents of WM consist of information held in buffers. The most relevant buffer for resuming from an interruption is the imaginal buffer, which contains the current problem state. Individual differences in WM capacity can be interpreted as differences in the amount of source activation that can be spread to activate chunks in LTM (Daily, Lovett, & Reder, 2001). The more source activation that can be spread, the more quickly and successfully chunks from LTM can be brought into a buffer to be acted upon. According to the memory for goals model implemented in ACT-R, interruptions require people to encode the current task state in LTM and then retrieve it upon task resumption. Thus, interruptions disrupt performance by causing problem state suspension and resumption, and individuals with more source activation to spread to LTM will be more likely to retrieve suspended problem states. In this way, the memory for goals model would predict that individuals with greater WM resources resume interrupted tasks faster and more accurately.

Previous findings concerning interruption recovery may be explained by the demands an interrupting task makes on WM. This explanation is consistent with a more general finding in multitasking that indicates WM capacity as a potential predictor of multitasking performance

(Colom, Martinez-Molina, Shih, & Santacreu, 2010; Hambrick et al., 2010; König et al., 2005; Bühner, König, Pick, & Krumm, 2006; Logie et al., 2011). There are additional reasons to suspect that WM plays a role. For example, more complex interruptions were found to require longer recovery times (Cades, Werner, Trafton, Boehm-Davis, & Monk, 2008). Because more complex interruptions require more time and potentially more task state information to complete, the interrupted task state in LTM experiences a longer delay before attempted retrieval, leading to a lower activation for that task state in memory and making it more difficult to recover from the interruption. Thus, previous findings using the memory for goals model and ACT-R architecture suggest a potential relationship between individual differences in cognitive resources (e.g., WM) and interruption recovery. However, WM may not be the only individual difference that impacts interruption recovery. Therefore, the present study examined a number of potential sources of individual differences in multitasking and interruption recovery.

1.2 Past research on individual differences in multitasking

Prior research has explored individual differences in multitasking ability such as how experience with media multitasking relates to task switching (Alzahabi & Becker, 2013; Ophir, Nass, & Wagner, 2009) or how individual differences relate to dual-task performance decrements (Strayer & Drews, 2007; Watson & Strayer, 2010). However, prior research has not specifically addressed the relationship between individual differences in cognitive ability and interruption recovery in multitasking situations.

Prior studies of multitasking have predominantly used tasks that involved switching frequently between different tasks. However, switching between different tasks can require updating task goals, rules and, potentially, cognitive processes in addition to retrieving

suspended problem state information from memory. For example, in a task used by Hambrick and colleagues (2010), participants had to frequently switch between four subtasks (arithmetic, memory search, auditory monitoring, and visual monitoring tasks) in order to achieve a higher overall score on the task. Each of the four subtasks required different sets of actions. Thus, experimental paradigms that involve frequent switching between distinct tasks often make it difficult to isolate performance changes due only to interruption recovery (i.e., memory retrieval of problem state) from performance changes due to switching between distinct task rules and cognitive processes.

Although tasks used in prior research have made it difficult to isolate the impact of interruptions specifically, these tasks have provided potential predictors of overall multitasking performance. One potential predictor is fluid intelligence (König et al., 2005). Fluid intelligence, usually measured by Raven's Standard Progressive Matrices (Raven & Court, 2003), is the ability to solve novel problems that cannot be solved directly by referring to a store of long-term knowledge, instead requiring analytic or reasoning processes (Prabhakaran, Smith, Desmond, Glover, & Gabrieli, 1997). Given the correlation often observed between Raven's Matrices and WM measures (Conway, Cowan, Bunting, Theriault, & Minkoff, 2002; Gray, Chabris, & Braver, 2003; Kyllonen & Christal, 1990; Prabhakaran et al., 1997; Stauffer, Ree, & Carretta, 1996), it is perhaps not surprising that both measures have been found to relate to multitasking performance (Hambrick et al., 2010).

Other potential predictors of multitasking performance found in previous studies include spatial problem solving ability and perceptual speed. Spatial problem solving is the ability to reason about visual displays (McGee, 1979). Measures of spatial problem solving ability were found to be associated with success in multitasking environments such as air traffic control and

piloting (Alderton, 1989). Importantly, these measures have been found to explain variance in multitasking performance beyond that accounted for by fluid intelligence or WM in previous studies (e.g., Morgan et al., 2011). Therefore Alderton's measure of spatial ability was included in the present study. Perceptual speed is the ability to quickly and accurately attend to specific details in the environment (Thurstone & Jeffrey, 1984), and research has suggested that perceptual speed is correlated with multitasking performance (Oberlander, Hambrick, Oswald, & Jones, 2007). The process of finding the interrupted task in order to resume after an interruption might involve the process of quickly and accurately attending to details in the environment.

Given that multitasking often requires one to allocate attention to more than one thing at a time, measures of the ability to track multiple objects or events are also likely to be related to multitasking. One such measure is called Multi-Threading (Brou & Cotton, 2011), which requires participants to track the current status of multiple task-relevant objects and to make quick decisions about what to do next in response to target events, processes likely to occur in multitasking situations. Previous findings have shown that individual differences in Multi-Threading task performance explained unique variance in multitasking ability while operating work stations in simulated helicopters (Brou & Cotton, 2011). The ability to track multiple events might rely more on one's ability to share attention across those events, and this ability might not be measured spatial reasoning or perceptual speed. Therefore, a multiple-event tracking measure was also included in the present study.

In addition to these measures, individual differences in strategy use have been related to multitasking performance (Hambrick et al., 2010; Logie et al., 2011; McFarlane, 2002). Hambrick et al. (2010) measured strategy use by examining the proportion of time that participants either made successive responses in the same task or made successive responses in

different tasks, reflecting participants' tendency to stay on the same task or to make a transition from one task to another. These measures were found to account for variance in multitasking performance. Similarly, McFarlane (2002) found that different strategies for coordinating interruptions affected interruption performance. Therefore, effective strategies may help to manage the demands of multitasking environments, including interruptions. In particular, the present study focuses on strategy consistency as a potential influence on multitasking performance.

In the office-work example, people might adopt a strategy such as only responding to email alerts after completing a subtask that the email alert was interrupting. This strategy helps to minimize the amount of information that has to be remembered about a task's state. Supporting evidence for this kind of strategy use comes from a previous study that found that people tend to switch to an interrupting task when the amount of problem state information that had to be remembered about the interrupted task was minimal (Salvucci & Bogunovich, 2010). Another possible strategy might be to always check emails according to their arrival order and to always reply to an email immediately after it is read. With this strategy, a consistent sequence of steps would be followed, minimizing time spent deciding what to do next. The expected decrease in demand for processing resources that more consistent strategies might incur would likewise lead to better overall performance by helping to minimize the demands placed on different cognitive resources. Moreover, if a consistent strategy helps to minimize demands on cognitive resources, then using a consistent strategy may facilitate recovering from interruptions.

Based on the prior research concerning individual differences in cognitive ability, strategy use, and their limitations, the present study was designed to measure an array of individual differences in cognitive ability and strategy consistency to examine how these

measures impact multitasking performance, emphasizing their effect on interruption recovery.

1.3 Present research

The present study examined how interruption recovery performance is related to individual differences in cognitive ability. A battery of tasks was used to assess a wide range of cognitive abilities. Alderton's Integrating Details (Alderton, 1989) was used to measure spatial problem solving ability. Perceptual Speed (Thurstone & Jeffrey, 1984) was used to measure how fast people can do feature-mapping. Multi-Threading (Brou & Cotton, 2011) was used to measure the ability to track multiple objects and events. Raven's Standard Progressive Matrices (Raven & Court, 2003) was used to measure fluid intelligence. Although both fluid intelligence and WM measures have been found to be related to multitasking performance (Hambrick et al., 2010; König et al., 2005), the individual differences measures used in the present study included Raven's Standard Progressive Matrices, but no WM measures. Individual differences in WM were not measured because of the high correlation between WM and fluid intelligence (Gray et al., 2003; Hambrick et al., 2010; Kyllonen & Christal, 1990; Prabhakaran et al., 1997; Stauffer et al., 1996) and a desire to measure an array of abilities in a limited amount of time.

The Abstract Decision Making (ADM) task was used to examine multitasking in the present study, because it has multitasking features (i.e., time-pressure and interruptions) and has been shown to capture variance in multitasking performance that predicts performance in real-world multitasking settings such as emergency dispatch and air traffic control (Joslyn & Hunt, 1998). Each subtask of the ADM requires participants to sort an object into a bin by matching features (e.g., color, shape, and size) between the object and four possible bins. Participants first memorize a set of bin attributes before the start of a block of the ADM task (e.g. large, blue

squares belong in Bin 2). Participants use a text-based interface to select an object from the queue of available objects, query object features, and then place the object into the correct bin. For example, a participant might select Object 2 and then select the “shape” query and receive the response “square”. The feature disappears from the screen and the participant proceeds to query the color and size of the object before deciding that it can be placed into Bin 2. Figure 2 presents screenshots from the task showing each of these steps in sorting an object.

Participants could be interrupted at any time in this sorting process by the arrival of a new object. The arrival of the new object brought participants to an interface where they could select an object to sort. Each interruption therefore took a short time period during which participants could select any available object, including resuming querying and sorting the interrupted object or selecting any other object. Interruptions ended when participants selected an object. Because interruptions occur for some but not all objects, the design of the ADM task provides opportunities for comparing interrupted and non-interrupted task performance. It also allows for examining factors affecting interruption recovery by relating individual differences measures with differences between interrupted and non-interrupted task performance.

Although interruptions in the ADM task were shorter than those used in previous interruption studies (e.g., Monk et al., 2008; Brumby, Cox, Back, & Gould, 2013), costs associated with switching to subtasks with differing cognitive resource demands were eliminated. Both the interrupting and interrupted tasks involve a similar set of task rules and cognitive processes. Therefore, the ADM task provides a way to study the impact of interruptions without the confounding effects of switching between different tasks

The difficulty of the interrupting and interrupted tasks affects interruption disruption and recovery (Cades et al., 2008; Speier et al. 1999). Therefore, the present study included three

ADM blocks with increasing task difficulty, with higher difficulty requiring that more object information be queried and retained in order to make the correct bin-assignment decision. By manipulating block difficulty, the present study manipulated WM demands, allowing for further examination of the relationship between WM and interruption disruption. This method of manipulating WM demands is similar to that of previous studies in which the role of WM in performance errors has been examined by not only including individuals with different WM capacity, but also manipulating WM demands (Byrne & Bovair, 1997; Ament, 2011; Ament, Cox, Blandford, & Brumby, 2013).

With regard to strategy consistency, the ADM task allows for some flexibility in deciding the order in which participants query object features, how many features they need to query before sorting, and whether or not they should return to an object that was interrupted. Flexibility of possible actions that participants can take in sequences in the ADM task allowed for an examination of how strategy consistency relates to interruption recovery. Because strategy consistency may help compensate for low cognitive ability by reducing cognitive resource demands, interruption recovery improvements related to strategy consistency were expected to differ for individuals with different cognitive ability levels.

1.3.1 Hypotheses

Three hypotheses about the relationship of individual differences to interruption recovery in multitasking were proposed. The Interruption Disruption hypothesis states that interruptions will cause significant performance decrements, especially as the difficulty of the task increases. This hypothesis is consistent with previous studies suggesting that interruptions can cause powerful disruptions leading to performance decrements in task completion time and accuracy

(e.g., Edwards & Gronlund, 1998; Monk, 2004; Trafton et al., 2003). Therefore, participants were expected to perform slower and less accurately in interrupted trials than in non-interrupted trials and that these disruptive effects would increase with block difficulty.

The Cognitive Ability hypothesis examines the degree to which individual differences in cognitive abilities affect interruption recovery. As mentioned previously, interruption recovery involves retrieving the problem state. The problem state might contain a chunk indicating the value of the queried features (e.g., color is yellow, size is large, shape is unknown). After an interruption, one could either attempt to retrieve the current problem state and resume sorting the interrupted object, or one could simply decide to abandon the prior problem state and reconstruct it, possibly after failing to retrieve that state. Thus, higher WM individuals should resume an interrupted object more often and more quickly. However, other measures of cognitive ability might also be correlated with the ability to handle interruptions. Therefore, identifying how different measures vary in predicting the ability to handle interruptions is necessary. This hypothesis will be addressed by not only analyzing the speed and accuracy of resuming the task after an interruption but also by examining how the presence of an interruption affects errors and time spent sorting objects.

The Strategy Compensation hypothesis focuses on the role of strategy consistency in interruption recovery. Using a more consistent strategy is hypothesized to lead to more available processing resources, helping to compensate for a lack of available resources in those individuals low in cognitive abilities relevant to interruption recovery. Supporting evidence for this hypothesis would be that more consistent strategies are related to fewer performance decrements caused by interruptions. In addition, this relationship is expected to be stronger for people with lower cognitive abilities. In other words, cognitive abilities should moderate the relationship

between strategy consistency and interruption recovery.

2. Materials and Methods

2.1 Participants

Data were obtained for 229 participants recruited from Mississippi State University. Participants were 18 to 35 years old and had normal/corrected vision. Participants were either paid or compensated with course credit.

2.2 Materials and procedure

Two 1-hour sessions were performed separated by one to three days. In the first session, participants performed the ADM task (Joslyn & Hunt, 1998) followed by the Multi-Threading task (Brou & Cotton, 2011). In the second session, participants performed Raven's Standard Progressive Matrices (Raven & Court, 2003), Integrating Details (Alderton, 1989), and Thurstone's Perceptual Speed (Thurstone & Jeffrey, 1984) in that order. All tasks were presented on a computer that recorded response times and accuracy.

2.2.1 Raven's Standard Progressive Matrices

Raven's Standard Progressive Matrices was designed to measure fluid intelligence (Raven & Court, 2003). In each trial, participants saw a figure with a missing piece and six possible pieces below it. All figures and pieces were presented in black on a white background. Each figure was a shaded rectangle or a matrix of geometric figures. Participants had to click to select the missing piece that best matched the pattern. The dependent measure was the number of correct trials minus the number of incorrect trials.

2.2.2 Thurstone's Perceptual Speed

Perceptual speed was measured using a rapid feature-matching task originally developed by Thurstone and Jeffrey (1984). In each trial, a row of six figures was presented in black on a white background. The left figure of each row was the target figure. Only one of the five subsequent figures in the same row was identical to the target figure, whereas the other figures differed from the target figure slightly. Participants were instructed to click to select an identical match for each target figure. 145 rows were presented with a total task time limit of 5 minutes. The dependent measure was the number of correct trials minus 0.25 multiplied by the number of incorrect trials.

2.2.3 Integrating Details

Integrating Details was designed to measure spatial problem solving ability (Alderton, 1989; Hunt, Pellegrino, Frick, Farr, & Alderton, 1988). In each trial, a set of component shapes was presented on the left of the display and a target shape was presented on the right. All shapes were presented as black lines on a white background, and each side of a component shape was labeled with a lower-case letter. Participants were asked to press one of two keys to indicate whether or not the presented component shapes could be connected on sides with matching labels to create the composite target shape. The dependent measure was the number of correct trials.

2.2.4 Multi-Threading

The Multi-Threading task was designed to measure the ability to keep track of multiple events (Brou & Cotton, 2011). This task lasted 10 minutes, during which participants tracked and

responded to an increasing number of balls bouncing around inside a square on a computer screen (shown in Figure 1). The test began with a single ball moving inside the square, and a new ball was introduced every minute until a maximum of 10 balls were on screen for the last minute of the test. Participants pressed the space bar whenever an “event” occurred. These “events” occurred when: 1) an even numbered ball touched a wall labeled ‘even’, 2) an odd numbered ball touched a wall labeled ‘odd’, or 3) a ball flashed blue. Distractor events also occurred, including “odd” balls hitting “even” walls or vice versa and balls flashing orange instead of blue. The dependent measure was the number of events identified within 1 second of the event minus the number of false alarms.

[Insert Figure 1 here.]

2.2.5 Abstract Decision Making

The ADM is a task in which participants sort objects into bins based on their features. Time pressure was included by instructing participants to sort as quickly and accurately as possible. Objects consisted of three features: color, shape, and size. Using key presses to navigate the task interface shown in Figure 2, participants made queries about object features in order to receive a text-based description of that feature. Each bin could only accept one object type. For example, one bin would only accept large, red squares. The features of bins changed between blocks. Participants were only able to see and memorize the features of the bins at the beginning of each block (not while sorting).

[Insert Figure 2 here.]

The task started with the arrival of the first object. Each object could be queried for its features as needed. Once participants felt they knew enough to accurately match the object to a

bin, they selected a bin (bottom left of Figure 2). For each sort attempt, participants received a score based on the number of features shared between the object and the selected bin. One point was awarded for each feature that matched and one point was removed for each feature that did not match the bin's features. For example, a score of 3 was received if the selected bin matched the object completely, but an object that matched two features of the bin resulted in a score of 1 (i.e., two points for the matching features minus one point for the mismatching feature). This scoring system is the same system originally used by Joslyn and Hunt (1998) and was not used in the analyses of errors described later. The score was only used as feedback to participants.

Objects arrived with a 0.5 probability every five seconds. Objects arriving during the sorting process would interrupt the task by returning to the queue screen (bottom right of Figure 2). Interruptions ended after an object was selected to query (top right of Figure 2). Each object remained in the queue until it was correctly sorted. Each block lasted until all objects were correctly sorted.

An initial practice block with nine objects was followed by four blocks of 20 objects each. Bin feature overlap was manipulated to change the difficulty of sorting across blocks. The overlap was calculated by counting the number of unique pairs of bins that shared at least one feature. Supposing Bin 1 and Bin 2 shared one feature (e.g., color), then this pair received a bin feature overlap value of 1. Greater bin feature overlap meant more object attributes had to be known in order to sort an object correctly. For example, for a bin feature overlap value of 0, only one feature needed to be known to sort an object correctly, whereas the bins with an overlap value of 1 would require knowledge of at least two features.

Given four bins, six pairings of the bins exist (e.g., 1-2, 1-3, 1-4, 2-3, 2-4, 3-4). For each of these six pairings, the number of pairs that had overlapping features varied across blocks.

Across the non-practice blocks, the number of pairs that had at least one overlapping feature was 3, 4, 6, and 3 respectively. With this design, bin overlap increased from the first to the third block, and the first and last blocks had matched bin overlap. See Supplementary Material for specific feature values assigned to each bin in each block.

2.2.6 ADM measures

Five dependent measures were used. The accuracy measure was the number of bin-assignment errors. A bin-assignment error occurred whenever an object was sorted into the wrong bin. The analyses of bin-assignment errors used a count of the errors and did not use the scoring system that provided feedback to participants. A measure called selected time was used for speed of task completion. Selected time was the amount of time spent actively querying and sorting an object after selecting it from the queue. Therefore, selected time excluded the time that an object was present in the queue of objects but not being worked on.

Two measures were used for interruption recovery performance. The first measure, resumption lag, is a common measure from the interruption literature (e.g. Altmann & Trafton, 2002) and is the amount of time required to re-initiate task progress following an interruption event. Resumption lag was measured from the time an interruption occurred to the time at which a participant made the next action in the task interface.

The second measure, interruption return proportion, was the probability of resuming interrupted tasks following the interruption. This measure was calculated as the total number of times a participant selected the interrupted object from the queue immediately after the interruption divided by the total number of interruptions per block. For example, if a participant working on object 17 was interrupted by object 16, then re-selected object 17 from the queue

first would be counted as immediately returning to the interrupted object, whereas selecting any other object would not.

Strategy consistency measured the degree to which participants' sorting behavior followed the same sequence of actions for each object. Participants had not received any instruction on strategy. Using a formula (entropy) borrowed from information theory (Shannon, 2001), the likelihood that the next action could be predicted based on previous sequences of actions was calculated. For each block, consistency values were calculated and averaged across six initial actions: selecting an object from the queue for the first time, querying color, querying size, querying shape, returning from an interruption, and making an incorrect sort. Consistency values following a correct bin-assignment as no further actions can be taken on an object after it has been correctly sorted. Separate counts were recorded of each different action made following each initial action. Using these counts, consistency was calculated with the entropy formula:

$$consistency_i = 1 - entropy = 1 - \left(- \sum_{j=1}^{j=n} p_{ij} \log(p_{ij}) \right)$$

where each p_{ij} is the proportion of times that action j followed action i . Higher entropy indicates less consistency (i.e., it is harder to predict what that participant would do next). For clarity, the formula $(1 - entropy)$ was used to arrive at a measure of consistency for action i .

3. Results

The primary hypotheses were concerned with the impact that cognitive ability and strategy consistency have on the ability to recover from interruptions in a multitasking environment. In order to examine these hypotheses, participants were grouped into four different cognitive ability groups. Initial analyses focused on how each ADM performance measure was

affected by cognitive ability, interruptions, and task difficulty (i.e., the degree of bin overlap). Strategy consistency was then examined to see if it accounted for variance in interruption disruptions in addition to individual differences in cognitive ability.

While the study was designed such that the fourth block had the same bin overlap as the first block in order to investigate effects of skill acquisition, none of the individual difference measures were related to improvement from the first to fourth block. Therefore, block 4 data is not reported in the present paper because the focus of the hypotheses is on the effect of individual differences in handling interruptions.

3.1 Individual differences factors and groups

Cognitive ability groups were derived from the four ability measures through the following process. First, the cognitive ability task scores were standardized by computing z-scores and then entered into a principal component factor analysis using varimax rotation. Task scores were found to load onto two distinct factors with eigenvalues equal to or above 1. Table 1 shows the loadings of each individual task on the two factors. Scores from Raven's Standard Progressive Matrices, Integrating Details, and Thurstone's Perceptual Speed loaded primarily on the first factor (eigenvalue of 2). The first factor is referred to as the *general-ability* factor and accounted for 45.3% of the variance in the data. This factor was conceptualized as being composed of a variety of cognitive mechanisms including WM capacity, visuospatial ability, and perceptual speed. Only Multi-Threading task scores loaded primarily on the second factor (eigenvalue of 1). The second factor was conceptualized as being related to the ability to spread attention to track multiple events across dimensions of visual information (e.g., color and spatial location) in order to make a simple judgment. This factor was referred to as *multiple-event-*

tracking and accounted for 25.2% of the variance in the data.

[Insert Table 1 here.]

Discriminant analyses were then conducted to classify participants into four cognitive ability groups using factor scores from both factors as the basis for the groupings. Each group was labeled as being low or high in each factor (e.g., the low-high group was low in general-ability and high in multiple-event-tracking). Resulting participant counts in each factor group were: low-low, 57; low-high, 38; high-low, 77; and high-high, 54. The accuracy of these groupings was verified by finding that the low-low and low-high groups had significantly lower general-ability scores ($M = -0.98$, $SD = 0.60$) than the high-low and high-high groups ($M = 0.71$, $SD = 0.51$), $F(1, 224) = 516.83$, $p < .001$. Also, the low-low and high-low groups had significantly lower multiple-event-tracking scores ($M = -0.66$, $SD = 0.52$) than the low-high and high-high groups ($M = 0.96$, $SD = 0.71$), $F(1, 224) = 396.34$, $p < .001$.

3.2 Interruption recovery and individual differences factors

The following analyses focus on the Cognitive Ability hypothesis in order to better understand how the measured individual differences factors are associated with measures of task performance and the ability to effectively deal with interruptions. In particular, the results will provide an understanding of the role of the general-ability or the multiple-event-tracking factors in predicting the ability to handle interruptions inherent in multitasking.

3.2.1 Interruption and individual difference effects on speed and errors

In order to assess the hypothesis that interruption disruption would be moderated by cognitive ability, two measures of ADM performance (bin-assignment errors and selected time)

were analyzed in two separate 3 (block: first, second, third) x 2 (interruption: interrupted, non-interrupted) x 2 (general-ability group: low, high) x 2 (multiple-event-tracking group: low, high) ANOVAs with block and interruption as within-subject factors and general-ability grouping and multiple-event-tracking grouping as between-subject factors. The interruption factor was quasi-experimental because interruptions occurred randomly throughout a block, leading to a variable number of interrupted objects in each block for each participant. For all analyses, in cases where violations of sphericity were likely as determined by Mauchly's criterion, Greenhouse-Geisser corrected p-values are reported.

First, bin-assignment errors and selected time were examined for evidence of the Interruption Disruption hypothesis. The hypothesis states that interruptions disrupt performance, especially in more difficult blocks. Supporting evidence for interruption disruptions would be shown by a main effect of the interruption factor or an interaction between the block and interruption factors. The mean number of bin-assignment errors per block is shown in Figure 3 for each of the individual differences groups, and the mean selected time for each block and group is shown in Figure 4. Interruptions did disrupt task performance by increasing the number of bin-assignment errors made to the interrupted objects, $F(1, 196) = 319.72, p < .001, \eta_p^2 = .620$, and by increasing the selected time for interrupted objects compared to non-interrupted objects, $F(1, 196) = 979.60, p < .001, \eta_p^2 = .833$. These findings support the hypothesis that interruptions caused performance decrements. In addition, an interaction between block and interruption was found for bin-assignment errors, $F(2, 392) = 3.12, p = .046, \eta_p^2 = .016$, showing that errors increased in later blocks in the interrupted trials, $F(1, 196) = 4.57, p = .034, \eta_p^2 = .023$, but not in the non-interrupted trials, $F < 1$. This result supports the Interruption Disruption hypothesis that interruptions cause performance decrements, especially as the difficulty of the task increases.

The interruption disruption effect on the selected time measure was also found to be affected by block, $F(1.92, 376.55) = 8.17, p < .001, \eta_p^2 = .040$. Further analysis of the interruption by block interaction showed that the increase of selected time caused by interruptions in the first block was significantly larger than that in the second, $F(1, 196) = 7.31, p = .008, \eta_p^2 = .036$, and third blocks, $F(1, 196) = 14.67, p < .001, \eta_p^2 = .070$. This finding is inconsistent with the hypothesis that interruption detriments are larger in more difficult blocks.

[Insert Figure 3 here.]

[Insert Figure 4 here.]

Given the results indicating that interruptions in the ADM were disruptive to performance, the following analyses explored the possibility that individual cognitive ability differences might account for differences in the ability to handle interruption disruption, as predicted by the Cognitive Ability hypothesis. For bin-assignment errors, participants in the low general-ability groups made more errors than those in the high general-ability groups, $F(1, 196) = 23.82, p < .001, \eta_p^2 = .108$, and participants in the low multiple-event-tracking groups made more errors than those in the high groups, $F(1, 196) = 4.95, p = .027, \eta_p^2 = .025$. While these two main effects indicate that both factors affected the number of errors, only the general-ability factor was found to interact with the interruption factor. Participants in the high general-ability groups showed a smaller increase in errors when interrupted (from their non-interrupted error baseline) than did those in the low general-ability groups, $F(1, 196) = 8.08, p = .005, \eta_p^2 = .040$. This result supports the Cognitive Ability hypothesis by showing that individual differences in the general-ability factor modulate the effect of interruptions. In addition, Figure 3 suggests that the interaction between block and interruption mainly happened to participants from the low-low group. Further simple effects analyses show that only the low-low group had a significant

interaction between block and interruption, $F(2, 106) = 3.73, p = .028, \eta_p^2 = .066$. Participants from groups other than the low-low group did not have this interaction, $F_s < 2$. No other main effects or interactions were significant in the bin-assignment errors ANOVA.

For selected time, similar to the bin-assignment error results, individual differences in both cognitive ability factors affected selected time as higher general-ability and higher multiple-event-tracking scores were both associated with a reduction in the amount of time an object was selected, $F(1, 196) = 30.62, p < .001, \eta_p^2 = .135$ and $F(1, 196) = 4.34, p = .039, \eta_p^2 = .022$, respectively. Again, only the general-ability factor was found to interact with interruptions, with participants in the high general-ability groups showing a smaller increase in selected time when interrupted than did those in the low general-ability groups, $F(1, 196) = 16.58, p < .001, \eta_p^2 = .078$. This result is consistent with the Cognitive Ability hypothesis, and indicates that the general-ability factor is more associated with successfully handling interruptions than the multiple-event-tracking factor. In addition, consistent with a possible practice effect, selected time decreased across blocks, $F(1.88, 368.40) = 23.35, p < .001, \eta_p^2 = .106$. No other main effects or interactions were significant.

3.2.2 Effects of individual differences on interruption resumption

Resumption lag and interruption return proportion were the two interruption resumption metrics examined. Because these metrics are only defined for interrupted objects, they were analyzed using two separate 3 (block) x 2 (general-ability group) x 2 (multiple-event-tracking group) ANOVAs. In cases where violations of sphericity were likely, as determined by Mauchly's criterion, Greenhouse-Geisser corrected p-values are reported.

The mean resumption lag for each block and group is shown in Figure 5. Resumption lag

decreased across blocks, $F(1.75, 342.34) = 42.61, p < .001, \eta_p^2 = .179$, suggesting a practice effect. Of the two individual differences factors, only the general-ability factor had a main effect on resumption lag, $F(1, 196) = 7.02, p = .009, \eta_p^2 = .035$. In addition, a three-way interaction between block and the two individual differences factors was marginally significant, $F(1.75, 342.34) = 3.02, p = .057, \eta_p^2 = .015$. Given the relevance to the Cognitive Ability hypothesis, this interaction was examined by using a t-test to compare each pair of the four ability groups' mean resumption lags for each block. The low-low group was significantly worse than the high-low and high-high groups in block 1, $t(112) = 3.19, p = .002, d = .598, t(100) = 3.01, p = .003, d = .599$, and in block 3, $t(102) = 2.62, p = .010, d = .500$, and $t(98) = 2.43, p = .017, d = .488$. As depicted in Figure 5, the low-low group improved their ability to recover from interruptions, reaching the level of other groups in block 2. At the same time, this group was more affected by block difficulty and did not reach the level of the other groups in block 3. No other main effects or interactions were significant.

[Insert Figure 5 here.]

The mean interruption return proportion for each block and group is shown in Figure 6. Interruption return proportion increased across blocks, $F(2, 392) = 3.97, p = .020, \eta_p^2 = .020$, suggesting a practice effect. Participants in the high general-ability group returned to interrupted objects more frequently than those in the low group, $F(1, 196) = 16.35, p < .001, \eta_p^2 = .077$. This finding of an effect of general-ability but not multiple-event-tracking is consistent with results of the other measures showing that the Cognitive Ability hypothesis mainly pertains to the individual differences indexed by the general-ability factor. No other main effects or interactions were significant for the interruption return proportion.

[Insert Figure 6 here.]

To examine if all of the individual difference measures (i.e., Raven's Standard Progressive Matrices, Integrating Details, and Perceptual Speed) loading on the general-ability factor equally accounted for the effects of resuming from interruptions, each measure was correlated with both resumption lag and interruption return proportion. For resumption lag, the correlations for were Raven's, $r(198) = -.26, p < .001$, Integrating Details, $r(198) = -.19, p = .008$, and Perceptual Speed, $r(198) = -.27, p < .001$. Similar to the resumption lag result, individual differences measures loading on the general-ability factor are all correlated with the measure of interruption return proportion, Raven's, $r(198) = .36, p < .001$, Integrating Details, $r(198) = .36, p < .001$, and Perceptual Speed, $r(198) = .24, p < .001$. There is some indication that Raven's and Perceptual Speed are most associated with resumption lag, and that Raven's and Integrating Details may be more associated with returning to the object that was interrupted.

3.3 The role of strategy consistency

Consistent strategies were hypothesized to demand fewer cognitive resources, helping people to perform better in multitasking. Thus, individuals with lower scores on the individual differences factors might benefit the most from a consistent strategy that minimizes demands on cognitive resources. To examine this hypothesis, hierarchical regression analyses were conducted using the general-ability factor, the strategy consistency measure, and their interaction as predictors of performance. All four of the dependent measures examined above (errors, selected time, resumption lag, and interruption return proportion) were examined in four separate regression analyses.

The average consistency in the first two blocks was used to predict performance (e.g., errors) in the third block. This way the predictor and the predicted measures were not from the

same block. For errors and selected time, the difference between interrupted and non-interrupted trials in block 3 was the dependent measure in order to examine changes in performance caused by interruptions (i.e., interruption-induced errors).

The regression analyses were conducted in three steps. The general-ability factor was entered in the first step, the strategy consistency measure was entered in the second step, and their interaction was entered in the third step. This hierarchical order for entry reflected what was assumed to be the direction of influence in the relation of general-ability and consistency (i.e., an increase in general-ability would decrease the relationship between an individual's strategy consistency and the performance decrement caused by interruptions). For each variable entered in each step, a significant increment in variance (increment in R^2) accounted for by the variable would indicate its unique contribution.

The results of the hierarchical regression analysis for bin-assignment errors are summarized in Table 2. They show that the impact of interruptions was reduced by increases in both general-ability and strategy consistency. In Figure 7, a significant interaction shows that the relationship between consistency and interruption-induced errors is greatest in the low ability group. One participant had a lower strategy consistency and a greater increase in the amount of bin-assignment errors made due to interruptions than other participants. After removing this data point, a second analysis showed that general-ability ($\Delta R^2 = .04, F = 9.13, p = .002$) and strategy consistency ($\Delta R^2 = .05, F = 10.44, p = .001$) still contributed to unique variance of increases in bin-assignment errors. The interaction between general-ability and strategy consistency was not quite significant at the .05 level, $F = 3.19, p = .075$. Given that there was nothing else about this participant's data to indicate that s/he was not performing the task normally, one possibility is simply that this is the only participant in our sample that fell on the tail of the distribution.

[Insert Table 2 here.]

[Insert Figure 7 here.]

To examine the interaction, further analyses were conducted to examine the relationship between strategy consistency and interruption disruptions in terms of errors for each of the two general-ability groups. Individuals low on general-ability had a significant relationship between strategy consistency and interruption-induced errors, $B = -2.51$, $R^2 = .09$, $F = 8.43$, $p = .005$, whereas individuals high on general-ability did not have this significant relationship, $B = -1.61$, $R^2 = .02$, $F = 2.61$, $p = .109$. This result supports the hypothesis that a consistent strategy may reduce demands on limited cognitive resources shown as a moderating effect of strategy consistency on performance.

The results of the hierarchical regression of the selected time and resumption lag measures suggested that general-ability primarily accounted for unique variance of interruption-induced increases in selected time, $\Delta R^2 = .06$, $F = 13.56$, $p < .001$, and resumption lag, $\Delta R^2 = .03$, $F = 5.31$, $p = .022$. Neither strategy consistency nor the interaction accounted for unique variance, $\Delta R^2 < .01$, $F < 1$, respectively. Both general-ability ($\Delta R^2 = .11$, $F = 25.26$, $p < .001$) and strategy consistency ($\Delta R^2 = .02$, $F = 5.26$, $p = .023$) accounted for unique interruption return proportion variance, but the interaction between the general-ability factor and strategy consistency did not account for unique variance, $\Delta R^2 < .01$, $F < 1$.

4. Discussion

The ADM task served as a complex multitasking environment that allowed examination of three hypotheses, each of which addresses the impact of specific cognitive abilities or strategy consistency on interruption recovery. Evidence was found for each of the three hypotheses

proposed: the interruption disruption hypothesis, the cognitive ability hypothesis, and the strategy consistency hypothesis.

First, the Interruption Disruption hypothesis states that interruptions result in performance decrements especially as the difficulty of the task increases. Consistent with this hypothesis, interruptions led to increased bin-assignment errors and selected time needed to sort the interrupted objects. These findings not only confirm previous findings that interruptions have a disruptive effect on multitasking performance (e.g., Edwards & Gronlund, 1998; Monk, 2004), but also confirm that the short interruptions used in this study were significantly disruptive. Likewise, finding performance decrements in interrupted trials confirms that the similarity between the interrupting task and the interrupted task did not eliminate the effects of interruptions (Gould, Brumby, & Cox, 2013). The effect of interruptions on bin-assignment errors also increased as task difficulty increased, further supporting the Interruption Disruption hypothesis. In addition, across blocks, the decrease in selected time and resumption lag and the increase in the interruption return proportion across blocks suggest a practice effect.

4.1 Cognitive ability hypothesis

The Cognitive Ability hypothesis predicted that individual differences would affect the ability to recover from interruptions. The present study adopted a set of four individual difference measures of cognitive ability and extracted two ability factors (general-ability and multiple-event-tracking) from them. The fact that common variance for three of the measures loaded on a single general-ability factor is consistent with previous findings that indicate a relationship between measures of fluid intelligence, perceptual speed, and spatial ability (Miyake, Friedman, Rettinger, Shah, & Hegarty, 2001; Redick, Unsworth, Kelly, & Engle, 2012). The

Cognitive Ability hypothesis was concerned with how different factors would differ in accounting for different measures of individual differences in interruption recovery and minimizing interruption disruption.

Both individual differences factors affected the number of bin-assignment errors and the duration of selected time per object during ADM performance. This is consistent with previous findings that a number of individual differences measures account for variance in multitasking performance (e.g., Hambrick et al., 2010; König et al., 2005). However, the two individual differences factors played different roles in explaining performance related to interruption disruption and recovery.

4.1.1 The role of the general-ability factor in interruption recovery

The general-ability factor accounted for some variance in interruption recovery performance that the multiple-event-tracking factor did not. For example, only the general-ability factor was associated with interruption disruption for selected time. Although participants selected time increased when interrupted, the increase was smaller for participants with higher general-ability. In terms of interruption recovery, only the general-ability factor was related to interruption return proportion such that participants high in general-ability were more likely to resume partially completed work on an interrupted object. The general-ability factor was also related to resumption lag and errors associated with interruptions. Participants with lower general-ability had a larger resumption lag and a larger increase in errors associated with interruptions.

The relationship between the general-ability factor and interruption recovery performance might be explained by individual differences in WM resources. Even though the current study

did not include a specific WM measure, the general-ability factor is assumed to have a WM component, because measures such as Raven's, which is correlated with WM capacity, loaded high on the general-ability factor (e.g., Hambrick et al., 2010; Redick et al., 2012). As mentioned earlier, The ACT-R architecture (Anderson, 2007) along with the memory for goals theory (Altmann & Trafton, 2002) and the concept of a problem state (Borst et al., 2010; Salvucci & Taatgen, 2011) provide an account of how differences in WM resources might affect interruption recovery. According to these accounts, interruptions in the present study would be expected to cause problem state information (e.g., color is red, size is large, shape is unknown) of the currently selected object to be replaced by information needed to make an object selection. Problem state information encoded before the interruption occurred would need to be retrieved from LTM when resuming the interrupted task. Individual differences in WM (i.e., source activation differences) would affect the speed and likelihood of successful retrieval (Daily et al., 2001). Thus, the finding that higher general-ability individuals exhibited shorter resumption lags and a higher interruption return proportions could be interpreted as being a product of greater WM resources allowing more spreading activation for faster problem state retrieval during interruption recovery. This explanation of the general-ability measure is therefore consistent with the memory for goals theory and other existing literature on interruption resumption and problem state interference (e.g., Brumby et al., 2013; Salvucci & Bogunovich, 2010). Although beyond the scope of the current study, a detailed model could help assess these claims.

The present findings also highlight the importance of examining how the different cognitive resources measured by the general-ability factor might be utilized when handling interruptions. Individual difference measures loading high on the general-ability factor (i.e., Raven's, Integrating Details, Perceptual Speed) differentially correlated with the two interruption

recovery measures (i.e., resumption lag and interruption return proportion). The results show that Raven's and Perceptual Speed were more associated with resumption lag than Integrating Details, whereas Raven's and Integrating Details were more associated with returning to the interrupted task than Perceptual Speed. It is possible that the underlying cognitive processes for re-initiating a task after an interruption share cognitive resources with those for mentally mapping shape features in Perceptual Speed. Similarly, the cognitive processes enabling people to return to an interrupted task might share resources with those that enable spatial problem solving. These findings indicate that cognitive resources that enable people to quickly re-initiate a task after an interruption are different from those enable people to accurately return to the interrupted task.

The general-ability factor's effect on interruption recovery needs to be examined to see whether it generalizes to other tasks with different interruption characteristics. The interruptions in the present study were all unexpected and short. In contrast, expected interruptions (e.g., Monk, 2004) and interruptions that participants have the option to schedule (e.g., McFarlane, 2002) might not demonstrate the same performance changes related to general-ability factor differences. Further work will need to identify exactly which cognitive resources are critical for interruption recovery or if there are a number of resources, including WM, that are important for recovery from different types of interruptions.

4.1.2 The role of the multiple-event-tracking factor in interruption recovery

Although the multiple-event-tracking factor did not account for most interruption recovery performance, low-high group participants did not show the same pattern of resumption lag and error performance as did low-low group participants even though low-high participants had an equally low general-ability factor score. As shown in Figure 3, low-high participants did

not show as much of an increase in interruption errors as low-low participants did as block difficulty increased. In addition, as shown in Figure 5, the resumption lags of low-high participants did not increase as much as those of low-low participants in response to changes in block difficulty from block 2 to block 3 (the most difficult block). These findings suggest that higher multiple-event-tracking ability compensated for low-high participants' lack of general-ability resources, enabling them to be less affected by the increase in block difficulty than low-low participants during interruption recovery.

One way of explaining this potential compensation effect for low-high participants is by examining the potential overlapping cognitive processes for interruption recovery and the Multi-Threading task, the primary task loading on this factor. Participants who performed well in the Multi-Threading task might have adopted a proactive strategy and transferred it to the ADM task. The distinction between proactive and reactive strategies was introduced by Braver, Gary, and Burgess (2008). The transfer of a proactive strategy is consistent with Taatgen (2013)'s theory of the nature and transfer of cognitive skills. A proactive strategy can enable participants to prepare for the upcoming target events (e.g., a ball flashed blue or an even numbered ball touched an even labeled wall) so that they responded to target events more accurately and earned higher task scores than those who adopted a more reactive strategy to wait for an event without proactive searching for the event in the Multi-Threading task. When the proactive strategy was transferred to the ADM task, participants could prepare for handling interruptions so that they resumed from an interruption faster and were more accurate after resuming. For example, a proactive strategy for handling interruptions might be to constantly rehearse object features to make the problem state more retrievable after an interruption. The proactive strategy might have also compensated more for participants with low WM as block difficulty increased (i.e. in block 3). Using the

example of rehearsal, increased problem state activation in LTM due to many rehearsals would mean that low-WM individuals with less source activation to spread would not have been affected as much by interruptions as those who did not proactively rehearse. Additional studies are needed to further examine this explanation about the role of the multiple-event-tracking factor in interruption recovery.

4.2 Strategy compensation hypothesis

Analyses to explore the role of strategy consistency in multitasking performance were motivated by previous findings about a relationship between strategy use and overall multitasking performance (Hambrick et al., 2010). The present study extended these previous findings by focusing on how consistency impacts interruption disruption rather than overall multitasking. The Strategy Compensation hypothesis proposed that more consistent strategies require fewer decisions on what to do next, which means fewer problem states to be encoded and retrieved and fewer cognitive resources needed to perform the task. In this way, consistent strategies were expected to be useful for reducing interruption disruption in all individuals, but would specifically help to moderate increases in interruption disruption observed in low cognitive ability individuals. Supporting this hypothesis, interruption disruption was lower when strategies were more consistent as indicated by a moderation of the influence of cognitive ability on increases in bin-assignment errors during interruptions and unique variance in interruption return proportion accounted for by strategy consistency.

The compensatory role of strategy consistency provides a potential means for training low-ability individuals to improve interruption recovery. One implication of the Strategy Compensation hypothesis suggests that consistent strategies might improve interruption recovery

for individuals with limited cognitive resources. Individual differences in strategy consistency might be due to factors such as prior experience. In addition, it is unclear how strategy consistency effects might extend to tasks from previous studies that involve longer, dissimilar interruptions (e.g., Monk et al., 2008). Consistent strategy use may result in even greater benefits in such tasks due to the increased disruption of longer and more complex interruptions.

4.3 Limitations

Limitations of the present study suggest future directions for studying interruptions in multitasking. First, the present study did not include a direct measure of WM. Although many individual differences measures adopted in the present study were found to be related to WM (e.g., Gray et al., 2003; Kyllonen & Christal, 1990), the lack of a direct measure of WM makes it hard to directly examine the relationship between WM resources and interruption. The relationship found between the general-ability factor and interruption recovery performance might be explained by other cognitive resources. Future studies may adopt some commonly used measures of WM capacity to examine the relationship between WM capacity and interruption recovery performance. With these further studies, it should be possible to determine whether the relationship found between the general-ability factor and interruption recovery performance might be explained by WM or other cognitive resources.

Also, interpretation of the present findings is constrained due to a feature of the task design. Because interruptions appeared with a 50% chance every five seconds, faster performance led to fewer interruptions. This artifact might have exaggerated the relationship between individual differences in cognitive ability and interruption disruption because individuals low in cognitive ability typically took longer to complete the task. Future studies

would need to better control the amount of interruptions in order to confirm the present findings. However, findings from a study that uses a modified version of the ADM task that better controls interruptions are consistent with results of the present study (Jones, Bai, Moss, & Doane, 2014), suggesting that this artifact did not significantly impact the present findings.

The interruption design also did not examine the interference potentially caused by using similar subtasks in the ADM. Edwards and Gronlund (1998) have suggested that interference might occur in WM for interrupting tasks that share associated elements with their interrupted task. Their results indicated that non-associated and associated interruptions were equally disruptive. However, Gould et al. (2013) found an interference effect for associated interruptions, but they noted that this effect was primarily observed when switching between subtasks and therefore might simply be switch costs not specific to interruptions. These conflicting results make the impact that associated interruptions have on performance difficult to determine in the present study and more work is needed in this area.

In addition, the role of strategy in interruption recovery was primarily explored via measures of consistency rather than assigning strategies to participants. It is possible that participants who used more consistent strategies were also different in other ways that enabled them to handle interruptions better. Future studies will need to further explore the role of strategies and the potential for training strategies for interruption recovery and multitasking as well as whether the current results generalize to tasks with longer interruptions.

4.4 Conclusions

The current study examined the role of individual differences in cognitive ability and strategy consistency in interruption recovery during multitasking. Results were interpreted under

the view that interruptions disrupt performance by competing for cognitive resources, such as WM. Both of the cognitive factors extracted from four individual differences measures impacted general multitasking performance, but the general-ability factor affected interruption recovery to a greater degree than multiple-event-tracking. However, the multiple-event-tracking factor allowed for participants with low general-ability to compensate when recovering from interruptions in the most difficult block. These findings suggest that multiple cognitive resources impact interruption recovery in different ways. In addition, the results indicate that the use of consistent strategies can compensate for low cognitive ability in multitasking situations, particularly when recovering from interruptions. The present study provides unique contributions to understanding interruption effects in multitasking by adopting a task in which factors other than interruptions were better controlled than in previous studies.

The findings also provide new directions for studying the cognitive mechanisms underlying multitasking. Building on these findings, future research might aim to examine questions about the mechanisms underlying multitasking and why large differences in strategy consistency and interruption recovery occur between individuals. Also, the current findings indicate potential avenues for multitasking training (e.g., using consistent strategies) that could improve multitasking performance for individuals with differing abilities.

Acknowledgements

This research was supported in part by a grant from the Office of Naval Research (N00014-10-1-0491) awarded to the fourth and third authors. We are grateful to Chad Stewart, Paul Ladney, Devin Busha, Blake Edwards, Jennifer Kueven, Delta Boyles, Rachel Clarke, Aaron Wong, and Skylar Swindle for their assistance in collecting, analyzing, and preparing the data for publication.

References

- Alderton, D. L. (1989). *Development and Evaluation of Integrating Details: A Complex Spatial Problem Solving Test* (NPRDC Technical Report 89-6) (p. 46). San Diego, CA: Navy Personnel Research and Development Center.
- Altmann, E. M., & Trafton, J. G. (2002). Memory for goals: An activation-based model. *Cognitive Science: A Multidisciplinary Journal*, *26*, 39–83. doi:10.1016/S0364-0213(01)00058-1
- Alzahabi, R. & Becker, M. W. (2013). The association between media multitasking, task-switching, and dual-task performance. *Journal of Experimental Psychology: Human Perception and Performance*, *39*, 1485 – 1495. doi: 10.1037/a0031208
- Ament, M. G. A. (2011). The role of goal relevance in the occurrence of systematic slip errors in routine procedural tasks. (Unpublished doctoral dissertation). University College London, United Kingdom.
- Ament, M. G. A., Cox, A. L., Blandford, A., & Brumby, D. P. (2013). Making a task difficult: evidence that device-oriented steps are effortful and error-prone. *Journal of Experimental Psychology: Applied*, *19*, 195-204. doi: 10.1037/a0034397.
- Borst, J.P., Taatgen, N.A., & Van Rijn, H. (2010). The problem state: A cognitive bottleneck in multitasking. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *36*, 363-382. doi: 10.1037/a0018106
- Braver, T. S., Gray, J. R., & Burgess, G. C. (2008). Explaining the Many Varieties of Working Memory Variation: Dual Mechanisms of Cognitive Control. In A. Conway, C. Jarrold, M. Kane, A. Miyake, & J. Towse (Eds.), *Variation in Working Memory* (pp. 76–106). Oxford University Press. Retrieved from

<http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780195168648.001.0001/acprof-9780195168648-chapter-4>

- Brou, R. & Cotton, J. (2011). *Assessing Multitasking Ability: Tracking and Responding to Multiple Threads of Information* (NPRST-TN-2011). Millington, TN: Navy Personnel Research, Studies, and Technology.
- Brumby, D. P., Cox, A. L., Back, J., Gould, S. J. J. (2013). Recovering from an interruption: Investigating speed-accuracy tradeoffs in task resumption strategy. *Journal of Experimental Psychology: Applied*, 19, 95-107. doi: 10.1037/a0032696
- Burgess, P. W. (2000). Real-world multitasking from a cognitive neuroscience perspective. In S. Monsell & J. Driver (Eds.), *Control of Cognitive Processes: Attention and Performance VIII* (pp. 465–72). Cambridge, MA: The MIT Press.
- Bühner, M., König, C. J., Pick, M., & Krumm, S. (2006). Working Memory Dimensions as Differential Predictors of the Speed and Error Aspect of Multitasking Performance. *Human Performance*, 19, 253–275. doi:10.1207/s15327043hup1903_4
- Byrne, M. D., & Bovair, S. (1997). A working memory model of a common procedural error. *Cognitive Science*, 21, 31-61. doi: 10.1207/s15516709cog2101_2
- Cades, D. M., Werner, N., Boehm-Davis, D. A., Trafton, J. G., & Monk, C. A. (2008, September). *Dealing with Interruptions can be Complex, but does Interruption Complexity Matter: A Mental Resources approach to Quantifying Disruptions*. Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting, New York, NY (Vol. 52, pp. 398–402). Thousand Oaks, CA: Sage Publications. doi:10.1177/154193120805200442
- Cades, D. M., Boehm-Davis, D. A., Trafton, J. G., & Monk, C. A. (2011). Mitigating disruptive

- effects of interruptions through training: What needs to be practiced? *Journal of Experimental Psychology: Applied*, 17(2), 97–109. doi:10.1037/a0023497
- Colom, R., Martinez-Molina, A., Shih, P. C., & Santacreu, J. (2010). Intelligence, Working Memory, and Multitasking Performance. *Intelligence*, 38, 543–551.
doi:10.1016/j.intell.2010.08.002
- Conway, A. R., Cowan, N., Bunting, M. F., Theriault, D. J., & Minkoff, S. R. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, 30, 163–183. doi:10.1016/S0160-2896(01)00096-4
- Daily, L. Z., Lovett, M. C., & Reder, L. M. (2001). Modeling individual differences in working memory performance: A source activation account in ACT-R. *Cognitive Science*, 25, 315-353.
- Edwards, M. B., & Gronlund, S. D. (1998). Task Interruption and Its Effects on Memory. *Memory*, 6, 665–687.
- Gould, S. J. J., Brumby, D. P., Cox, A. L. (2013, September). What does it mean for an interruption to be relevant? An investigation of relevance as a memory effect. Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting, San Diego, CA (Vol. 57, pp. 149-153). Thousand Oaks, CA: Sage Publications.
doi:10.1177/1541931213571034
- Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience*, 6, 316–322. doi:10.1038/nn1014
- 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

Hunt, E., Pellegrino, J. W., Frick, R. W., Farr, S. A., & Alderton, D. (1988). The ability to reason about movement in the visual field. *Intelligence, 12*, 77–100. doi:10.1016/0160-2896(88)90024-4

Jones, W.E., Bai, H., Moss, J., & Doane, S.M. (2014). Many tasks, one mind: An fMRI study exploring the neural correlates of multitasking ability. Manuscript in preparation.

Joslyn, S., & Hunt, E. (1998). Evaluating individual differences in response to time-pressure situations. *Journal of Experimental Psychology: Applied, 4*, 16–43. doi:10.1037//1076-898X.4.1.16

König, C. J., Bühner, M., & Mürling, G. (2005). Working Memory, Fluid Intelligence, and Attention Are Predictors of Multitasking Performance, but Polychronicity and Extraversion Are Not. *Human Performance, 18*, 243–266.
doi:10.1207/s15327043hup1803_3

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

Logie, R. H., Trawley, S., & Law, A. (2011). Multitasking: Multiple, domain-specific cognitive functions in a virtual environment. *Memory & Cognition, 39*, 1561–1574.
doi:10.3758/s13421-011-0120-1

McGee, M. G. (1979). Human spatial abilities: Psychometric studies and environmental, genetic, hormonal, and neurological influences. *Psychological Bulletin, 86*, 889–918.
doi:10.1037/0033-2909.86.5.889

Miyake, A., Friedman, N. P., Rettinger, D. A., Shah, P., & Hegarty, M. (2001). How are visuospatial working memory, executive functioning, and spatial abilities related? A

- latent-variable analysis. *Journal of Experimental Psychology: General*, *130*, 621–640.
doi:10.1037//0096-3445.130.4.621
- Monk, C. A. (2004, September). *The effect of frequent versus infrequent interruptions on primary task resumption*. Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting, New Orleans, Louisiana (Vol. 48, pp. 295–299). Thousand Oaks, CA: Sage Publications. Retrieved from <http://pro.sagepub.com/content/48/3/295.short>
- Monk, C. A., Trafton, J. G., & Boehm-Davis, D. A. (2008). The Effect of Interruption Duration and Demand On Resuming Suspended Goals. *Journal of Experimental Psychology: Applied*, *14*, 299.
- Morgan, B., D'Mello, S., Fike, K., Abbott, R., Haass, M., Tamplin, A., Radvansky, G., Forsythe, C. (2011, September). *Individual Differences in Multitasking Ability and Adaptability*. Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting, Las Vegas, Nevada (Vol. 55, pp. 919–923). Thousand Oaks, CA: Sage Publications. doi:10.1177/1071181311551191
- Oberlander, E. M., Oswald, F. L., Hambrick, D. Z., & Jones, L. A. (2007). *Individual Difference Variables as Predictors of Error during Multitasking* (NPRST Technical Note 07-9). Millington, TN: Navy Personnel Research, Studies, and Technology.
- Ophir, E., Nass, C., & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences*, *106*, 15583-15587. doi: 10.1073/pnas.0903620106
- Prabhakaran, V., Smith, J. A., Desmond, J. E., Glover, G. H., & Gabrieli, J. D. (1997). Neural substrates of fluid reasoning: An fMRI study of neocortical activation during

- performance of the Raven's Progressive Matrices Test. *Cognitive Psychology*, 33, 43–63.
- Raven, J. C., & Court, J. H. (2003). *Manual for Raven's Progressive Matrices and Vocabulary Scales: General Overview*. Harcourt Assessment.
- Redick, T. S., Unsworth, N., Kelly, A. J., & Engle, R. W. (2012). Faster, smarter? Working memory capacity and perceptual speed in relation to fluid intelligence. *Journal of Cognitive Psychology*, 24, 844–854. doi:10.1080/20445911.2012.704359
- Salvucci, D. D., & Bogunovich, P. (2010, April). *Multitasking and Monotasking: The Effects of Mental Workload on Deferred Task Interruptions*. Paper presented at the Proceedings of the 28th International Conference on Human Factors in Computing Systems, Atlanta, GA (pp. 85–88). New York, NY: ACM. doi:10.1145/1753326.1753340
- Salvucci, D. D., & Taatgen, N. A. (2011). *The Multitasking Mind*. Oxford University Press.
- Seamster, T. L., Redding, R. E., Cannon, J. R., Ryder, J. M., & Purcell, J. A. (1993). Cognitive Task Analysis of Expertise in Air Traffic Control. *The International Journal of Aviation Psychology*, 3, 257–283. doi:10.1207/s15327108ijap0304_2
- Shannon, C. E. (2001). A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, 5, 3–55. doi:10.1145/584091.584093
- Speier, C., Valacich, J. S., & Vessey, I. (1999). The influence of task interruption on individual decision making: An information overload perspective. *Decision Sciences*, 30, 337–360.
- Stauffer, J. M., Ree, M. J., & Carretta, T. R. (1996). Cognitive-Components Tests Are Not Much More than g: An Extension of Kyllonen's Analyses. *The Journal of General Psychology*, 123, 193–205. doi:10.1080/00221309.1996.9921272
- Strayer, D. L., & Drews, F. A. (2007). Cell-Phone-Induced Driver Distraction. *Current Directions in Psychological Science*, 16(3), 128–131. doi:10.1111/j.1467-

8721.2007.00489.x

Taatgen, N. A. (2013). The nature and transfer of cognitive skills. *Psychological Review*, *120*(3), 439–471. doi:10.1037/a0033138.

Thurstone, L. L., & Jeffrey, T. E. (1984). *Space Thinking (Flags)*. Rosemont, IL: London House.

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

Watson, J. M., & Strayer, D. L. (2010). Supertaskers: Profiles in extraordinary multitasking ability. *Psychonomic Bulletin & Review*, *17*(4), 479–485. doi:10.3758/PBR.17.4.479

Figure Captions

Figure 1. Multi-Threading task example question as shown on computer screen.

Figure 2. ADM task sample screens.

Figure 3. Mean number of bin-assignment errors for each block and each group by interrupt condition. Error bars represent standard errors.

Figure 4. Mean selected time for each block and each group by interrupt condition. Error bars represent standard errors.

Figure 5. Mean resumption lag for each block and for each group. Error bars represent standard errors.

Figure 6. Mean interruption return proportion for each block and for each group. Error bars represent standard errors.

Figure 7. Relation of strategy consistency to the increase in bin-assignment errors due to interruptions in block 3 as a function of general-ability factor scores. The increase in bin-assignment errors due to interruptions is calculated by subtracting number of errors in interrupted trials by that in non-interrupted trials.

Table Captions

Table 1.

Loadings of Each Standardized Individual Differences Task Scores on Each Factor

Table 2.

Results of Hierarchical Regression Analyses Predicting Number of Errors Differences between Interrupted and Non-interrupted Cases

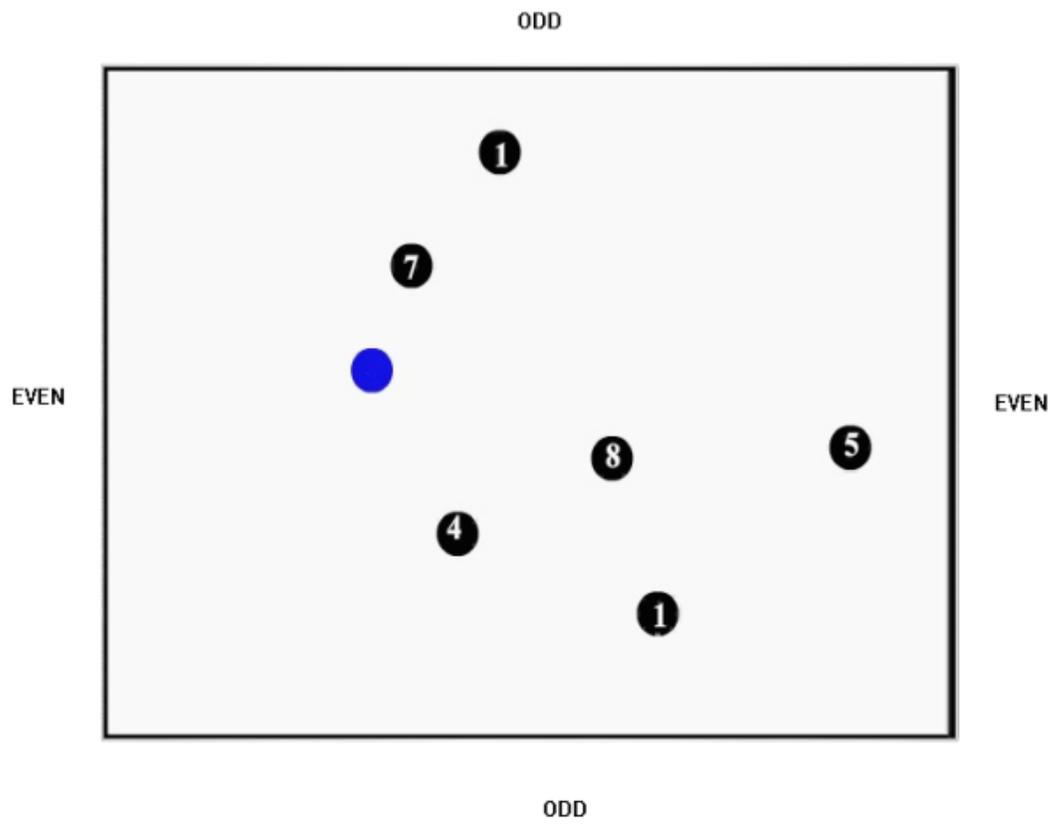


Figure 1. Multi-Threading task example question as shown on computer screen.

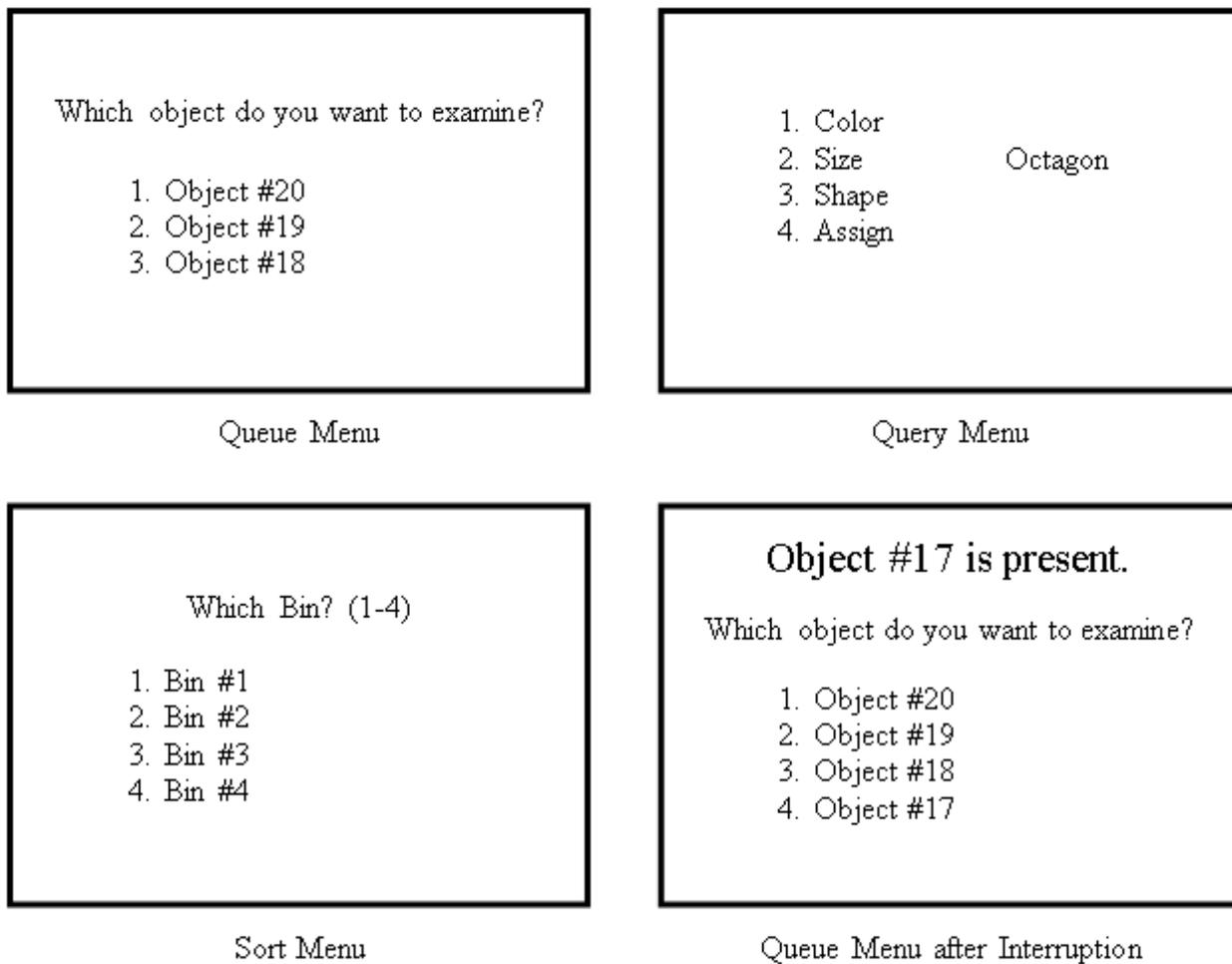


Figure 2. ADM task sample screens.

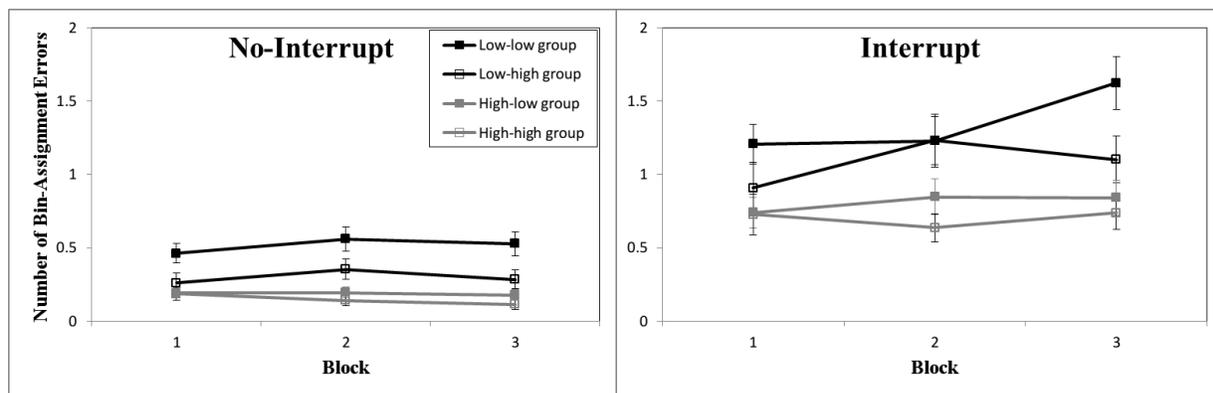


Figure 3. Mean number of bin-assignment errors for each block and each group by interrupt condition. Error bars represent standard errors.

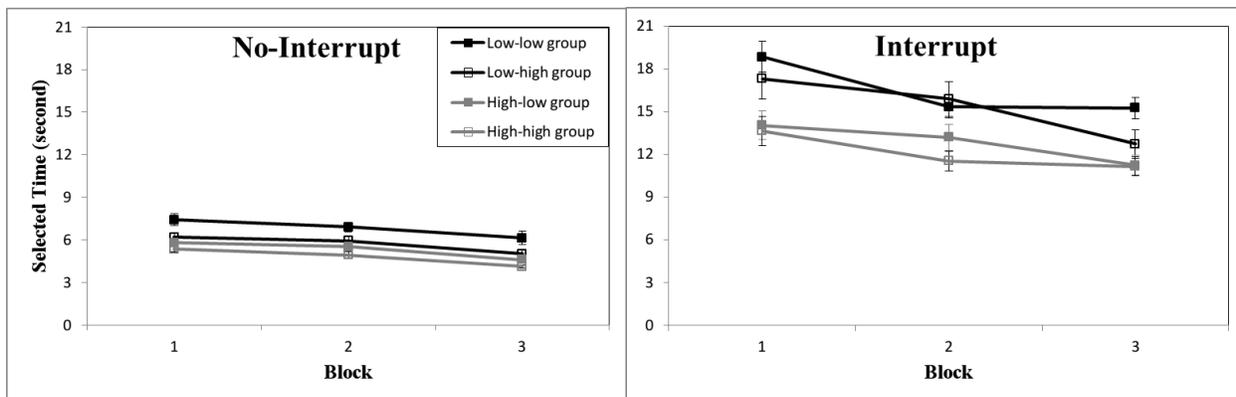


Figure 4. Mean selected time for each block and each group by interrupt condition. Error bars represent standard errors.

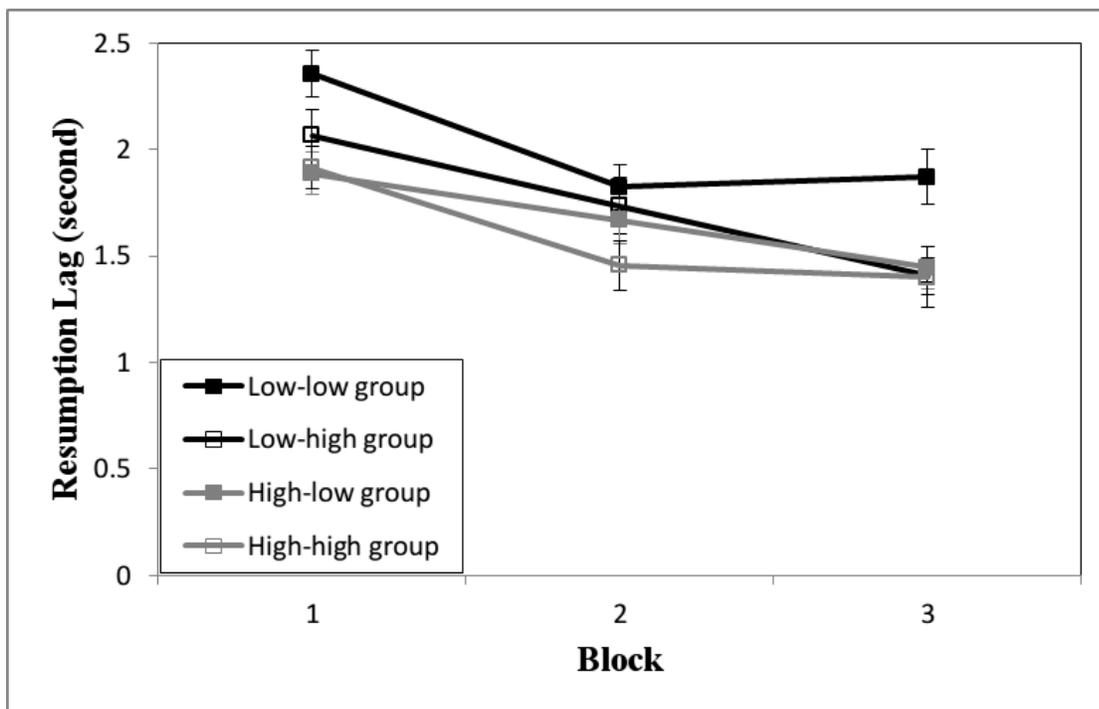


Figure 5. Mean resumption lag for each block and for each group. Error bars represent standard errors.

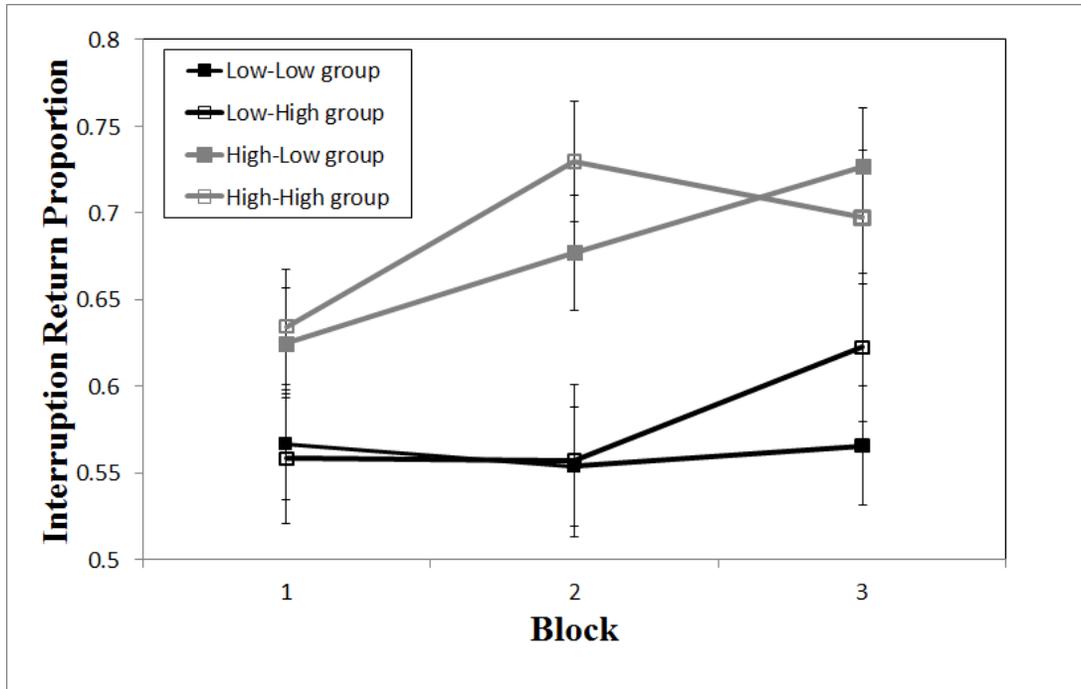


Figure 6. Mean interruption return proportion for each block and for each group. Error bars represent standard errors.

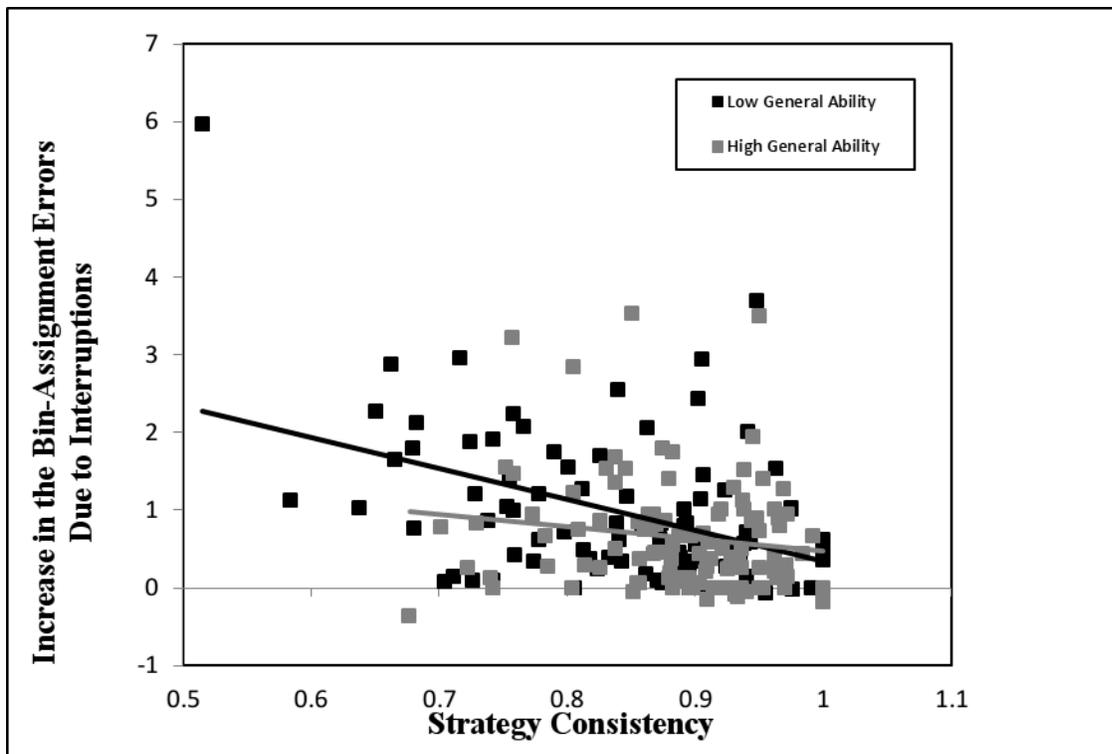


Figure 7. Relation of strategy consistency to the increase in bin-assignment errors due to interruptions in block 3 as a function of general-ability factor scores. The increase in bin-assignment errors due to interruptions is calculated by subtracting number of errors in interrupted trials by that in non-interrupted trials.

Table 1.

Loadings of Each Standardized Individual Differences Task Scores on Each Factor

Factor	Integrating Details	Progressive Matrices	Perceptual Speed	Multi-Threading
General-Ability	0.49	0.48	0.30	-0.08
Multiple-Event-Tracking	-0.18	-0.03	0.28	0.93

Table 2.

Results of Hierarchical Regression Analyses Predicting Number of Errors Differences between Interrupted and Non-interrupted Cases

Predictor Variable	R^2	ΔR^2	F change	p change
General-Ability	.06	.06	13.08	<.001
Strategy Consistency	.15	.09	19.71	< .001
General-Ability \times Strategy Consistency	.20	.05	13.38	< .001

Supplementary Material

Table S1

Bin Features in Each Block

Block	Bin	Bin Color	Bin Shape	Bin Size
1	1	Green	Triangle	Small
	2	Blue	Octagon	Small
	3	Yellow	Pentagon	Tall
	4	Blue	Octagon	Tall
2	1	Purple	Rectangle	Huge
	2	Red	Hexagon	Short
	3	Purple	Hexagon	Short
	4	Red	Rectangle	Huge
3	1	Yellow	Circle	Medium
	2	Orange	Circle	Tiny
	3	Orange	Square	Medium
	4	Orange	Circle	Medium
4	1	Red	Triangle	Tiny
	2	Blue	Square	Large
	3	Purple	Square	Tiny
	4	Red	Hexagon	Medium