

Effects of Automatic Detection on Dynamic Decision Making

Cleotilde Gonzalez

Carnegie Mellon University

Rickey P. Thomas

University of Oklahoma

ABSTRACT: The usefulness of the dual-process theory of automaticity to dynamic decision-making tasks is unclear. Dynamic decision making is characterized by multiple, diverse, and interrelated decisions that are often constrained by time limitations and workload. We investigated the relevance of this theory in a compound task consisting of multiple, dynamic components. In the first experiment, we reproduced the original findings that shaped the theory in a dynamic visual search task. In the second experiment, we added a decision-making component. The results replicated the original findings and extended them to decision-making components. Working under consistent-mapping conditions in the visual search component led to more accurate decision making despite variability in the decision-making conditions. Likewise, working under varied-mapping conditions in the visual search component led to poorer decision making, particularly under high workload. The implications of these results to real-world dynamic tasks are discussed.

Introduction

AUTOMATICITY HAS A LONG-STANDING HISTORY IN PSYCHOLOGY WITH PROPOSED THEORIES such as the modal view of automaticity (Kahneman, 1973), the dual-process theory (Schneider & Shiffrin, 1977), and the instance-based theory of automatization (Logan, 1988), among others. One of the most widely cited theoretical developments in automaticity research is the *dual-process theory* (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). According to this theory, automatic processing develops with extended practice when targets are consistently mapped (CM; targets are always targets and never appear as distractors). Such automatic processing is fast and parallel, requiring little attention or awareness. Thus, performance under automatic processing is unaffected by workload (Schneider & Shiffrin, 1977; Shiffrin & Schneider, 1977). On the other hand, under varied-mapping (VM) conditions, wherein stimuli may be targets in one instance but distractors in another, performance occurs under controlled processing, which is voluntary, is serial, and requires attention.

Researchers who followed the original work of Schneider and Shiffrin (1977) extended their basic model to include workload and dual-task performance (Fisk

ADDRESS CORRESPONDENCE TO: Cleotilde Gonzalez, Dynamic Decision Making Laboratory, Carnegie Mellon University, 5000 Forbes Ave., Porter Hall 223-C, Pittsburgh, PA 15213, coty@cmu.edu. Visit the *JCEDM* Online Companion at <http://cedm.webexone.com>.

Journal of Cognitive Engineering and Decision Making, Volume 2, Number 4, Winter 2008, pp. 328–348. DOI 10.1518/155534308X377810. © 2008 Human Factors and Ergonomics Society. All rights reserved.

& Schneider, 1983; Schneider & Fisk, 1984), automatic process development (Schneider & Fisk, 1984), skill acquisition (Fisk & Hodge, 1992), and training (Schneider, 1985). Recent reviews present a summary of the most up-to-date developments on automaticity research (Bargh & Ferguson, 2000; Moors & De Houwer, 2006). For example, Moors and De Houwer reviewed the various views of automaticity theory and advocated the gradual approach of automaticity development, under which clear processes and features are essential parts of automatic processes.

The development of automaticity is particularly important in complex tasks characterized by multiple target stimuli and distractors, time pressure, and workload, which make the tasks extremely difficult to perform in the absence of a practiced skill set. For example, in real-world contexts, ranging across fields such as aviation, military, and health care, highly skilled and automated levels of performance are needed to achieve fast and accurate responses to critical stimuli in the environment. These automated detection processes are particularly relevant to situations such as that of a pilot detecting the presence of an enemy among “friendly” aircraft, a physician detecting the presence of a tumor in X-rays, or a luggage screener detecting the presence of a hidden weapon among objects in passenger luggage.

The dual-process theory of automaticity suggests that CM is the most important factor for developing automatic processing (Kramer, Strayer, & Buckley, 1991; Schneider, 1985). Keeping CM in simple visual search tasks might be clear, but maintaining CM in complex real-world tasks can be a challenge. Real-world decision-making tasks are replete with multiple and heterogeneous stimuli often demanding decision making on varied or inconsistent conditions. For example, in air traffic control, controllers make multiple decisions about the allocation of runways and airspace depending on particular and often unique environmental conditions, as well as the interdependency of decisions made regarding arrivals, departures, and weather conditions. Likewise, in the field of military aviation, pilots must be able to distinguish the attributes of friendly aircraft from those of their enemy so they can select the appropriate response, which may further vary dramatically by situation and environment.

In contrast to these real-world examples, most research supporting theories of automaticity use simple stimuli and simple task paradigms. For example, the dual-process theory relies on a visual recognition task in which individuals determine whether static visual inputs such as letters or numbers are part of a memory set (Schneider & Shiffrin, 1977). The instance theory of automatization (Logan, 1988) uses linguistic stimuli in symbolic tasks and a counting task in which respondents are presented with visual and static patterns of discrete elements and asked to determine the number of patterns (Lassaline & Logan, 1993). The modal view of automaticity uses the well-known Stroop effect task, again with static stimuli (Kahneman & Chajczyk, 1983). Most tasks used in studies of automaticity involve static environments in which stimuli do not change while decisions are made.

Intuitively, it seems that CM training would be more beneficial than VM training in improving perception, attention, and detection skills, but it seems less clear

how CM would facilitate decision accuracy. From the foregoing examples, it is evident that CM could improve some real-world target detection tasks, but it is unclear how improved detection would influence the decisions made. For example, a military bomber may need to be identified as potentially hostile before a decision can be made about whether to ignore (if it is a friendly) or destroy (if it is an enemy), a physician would need to identify a potential problem in an X-ray before deciding whether to operate, and so on. Thus, one would expect that automating lower-level skills, such as perception or attention, could influence performance in higher-level tasks, such as decision making; however, the exact utility of the dual-process theory of automaticity to decision making is unclear. Our focus is on assessing the effect of CM and VM in visual search with dynamic stimuli on the degree of dynamic decision accuracy. We attempt to determine the benefits that CM brings not only to the visual detection of dynamic targets but also to dynamic decision making executed after targets are detected.

In this article, we report results of two experiments that test the dual-process theory of automaticity in a dynamic decision-making task. The first experiment uses a dynamic visual search component to reproduce the classical finding of the dual-process theory that CM targets are detected more accurately and faster than VM targets, with VM showing a detrimental effect of workload and CM being resistant to increases of workload. This experiment establishes the benefit of CM in a dynamic target detection task, extending the findings from simple and static stimuli (e.g., Schneider & Shiffrin, 1977).

The second experiment tests the influence of stimulus mapping (CM or VM) of the visual search component on decision-making accuracy. This experiment demonstrates that decision making can benefit from the automatic processing in the dynamic visual search task. The results from this research establish the benefits of CM for decision making in situations involving dynamic visual search and other workload conditions characteristic of real-world tasks.

The Radar Task

The radar task is a single-user control task in which the goal is to detect and eliminate a hostile enemy aircraft by selecting an appropriate weapon system (a screenshot of the radar task is shown in Figure 1). Radar is similar to military target visual detection devices, in which a moving target needs to be identified as a potential threat and a decision is made on how to best destroy the target. The development of this task was inspired by the theater defense program created by Bolstad and Endsley (2000), which was an adaptation of the radar monitoring and target elimination task developed by Kaber and Endsley (1997).

The radar task has two components: (a) visual and memory search and (b) decision making. The visual and memory search component requires the user to memorize a set of targets and then look for the presence of one or more targets on a radar grid. These visual and memory search activities are often considered two separate components in the original task by Schneider and Shiffrin (1977). The

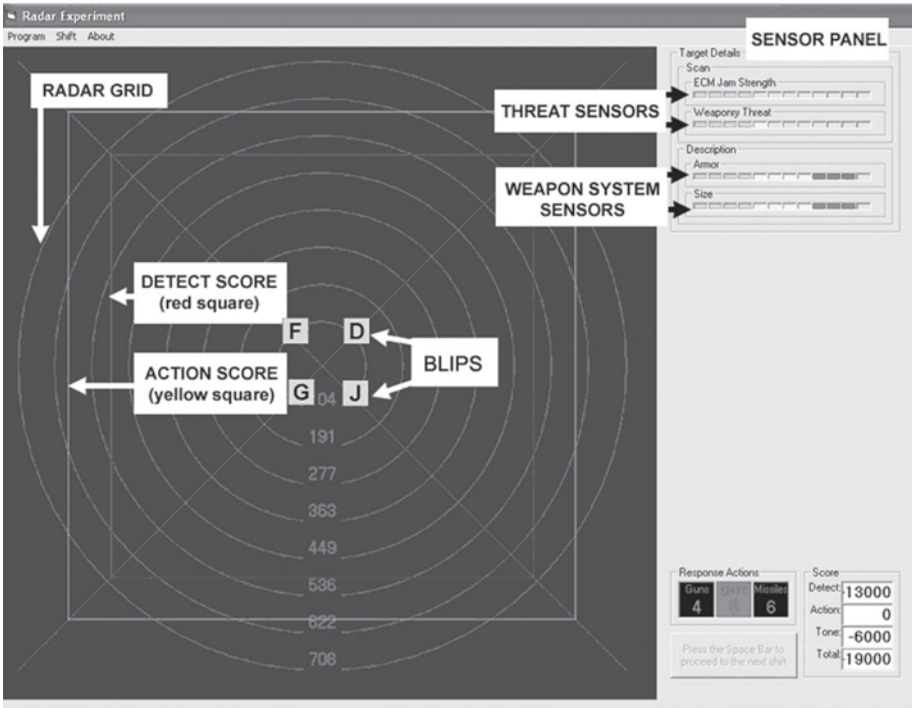


Figure 1. Illustration of the radar task: visual and memory search component, sensor reading and weapon selection, and feedback components.

visual and memory search components of the radar task essentially reproduce the goals of the original task, except that the visual elements in the radar task are not static but are dynamic. In Schneider and Shiffrin's task, all stimuli appeared within foveal vision, whereas in the radar task, the stimuli move on the screen, and thus eye movements and visual scanning are required to find a target. A target threat may or may not be present among a set of moving blips that represent incoming aircraft. The blips—in the form of symbols, digits, consonants, or blank masks—begin at the four corners of the radar grid and approach the center at a uniform rate. The detection of an enemy aircraft must occur before the blips collapse in the middle of the grid.

The second component of the radar task is decision making. This component is activated after a target has been selected, either correctly or incorrectly. The decision-making component requires the operator to select the correct weapon system to destroy a selected target according to the information provided by the "sensor meters." The sensors indicate probabilities on the hostility of the target, type, and size. Accurate decisions (most effective weapons) are determined by reading four different sensors. Two threat scan sensors, *ECM jam strength* and *weaponry threat*, indicate the probability that the selected target is hostile. Sensors

present green, yellow, or red light levels indicating low, medium, or high probability of the target being a threat, respectively. Two weapon system sensors, *armor* and *size*, represent the probability a target is of three different types and is small, medium, or large. Readings from the four sensors indicate which weapon system will be more effective to destroy the target. Generally, if the scan sensors' readings are low, the correct response action is to press the 5 key on the number pad (ignore). When ignore is the correct action, the two scan readings will be less than or equal to 4 (green). If the scan sensors' readings are high (6–12, red), then the weapon system sensors must be evaluated for the action of guns (pressing the 4 key) or missiles (pressing the 6 key). Generally, guns are used for low armor and size readings, whereas missiles are used for high armor and size readings.

The four sensors' values are randomly assigned to each target during the simulation run; thus, the users are only informed of the levels of the sensor in real time. The user must select the appropriate weapon system using dynamic information as targets are detected in each trial. If the operator detects no targets, he or she submits a quiet airspace report at the end of the trial by pressing the Shift key.

By using the radar task, one can manipulate the consistency of mapping and the workload. In CM, the targets that appear in the memory set before the visual search remain targets throughout the visual search task. In VM, an item can be a target in some trials but a distractor in other trials. The exact manipulations of consistent and varied mapping in the task will be explained in the experimental methods section. The workload is manipulated by increasing the number of items memorized (one item to memorize is low workload and four items is high workload). The number of nonblank items on the radar screen during the search also determines the workload (three blank items is low workload and no blank items is high workload). Thus, it is possible to measure automatic detection in the memory and visual detection components. Moreover, the radar task parallels many real-world tasks in which operators (such as AWACS and Aegis operators) must detect targets and make decisions about the appropriate action under extreme time pressure.

Experiment 1

Schneider and Shiffrin's (1977) first experiment reported that under VM, performance decreased monotonically as task workload increased. Task workload was defined as the product of memory set size (the number of characters individuals were asked to memorize) and frame size (the number of symbols per frame). Under CM, however, the effects of memory set size and frame size on performance were small, *not* statistically significant, and *not* monotonic with task workload. Schneider and Shiffrin's second experiment reported almost no change in the mean reaction time as a function of memory set size and frame size under CM but found large effects of both memory set size and frame size under VM. We expected to replicate these findings using the visual and memory search components of the radar task.

Method

Participants. Nine participants (6 women and 3 men), all graduate and undergraduate students, were recruited from local universities through bulletin boards. All of the participants were right-handed, had normal color vision, and had corrected or normal visual acuity. The average age was 20.11 years ($SD = .70$). The study consisted of six sessions of approximately 3 hr each during 6 consecutive days, for a total of 18 hr of practice. Participants were paid \$11.00 per hour for their participation. To avoid any fatigue effects, we encouraged individuals to take up to three breaks of 5 min each in each session.

Experimental Design. This experiment included stimulus mapping (CM and VM), frame size set at the two most extreme levels (1 and 4) in Schneider and Shiffrin's (1977) experiment, and memory set size set at the two most extreme levels (1 and 4). All variables were manipulated within subjects for a total of eight different conditions. A participant's response was target present (space bar) or absent (no response). The dependent variables were hits, false alarms, and detection time. Hits denote the average share of trials in which participants correctly detected a target. False alarms denote the average share of trials in which participants responded to a distractor. Detection time is the time it took participants to respond to the presence of a target by pressing the space bar.

In the CM condition, the memory set items (numbers or letters) never appeared in the visual display as distractors. In the VM condition, the memory set items appeared as targets in some trials and as distractors in other trials. Half of the participants were assigned to search for digits among letters in the CM condition and letters among letters in the VM condition; the other half were assigned to search for letters among digits in the CM condition and digits among digits in the VM condition. This design followed the original design by Schneider and Shiffrin (1977), which avoided the possible confounded effect because of the chosen categories.

The memory set size was either one or four items, and the frame size was also either one or four. That is, in each frame, there was either one symbol (one blip was a letter or a number, and the other blips were blank) or all four symbols (no blank blips).

This experiment had eight blocks of 16 frames each. Each block corresponded to one of the eight conditions. In each of six sessions (one session per day), participants were randomly presented with eight blocks of 16 frames (without replacement).

Procedure. Each block began with the presentation of the memory set (either one or four targets). Then individuals were presented with a sequence of frames. Each frame consisted of four moving blips beginning outside the foveal view. The operator's task was to detect any item from the memory set in the visual display. Targets could appear in any frame except the first and the last of the trial. The operators were instructed to press the space bar when they detected a target. A target was present in 75% of the trials. A different memory set was presented at the beginning of each block. When the frame size was 1, all noncharacter blips were blank. The character positions for each frame were selected randomly.

The time between the presentation of the blips and their collapse—called the *frame time*—remained constant (1,450 ms) across all frames in each trial. The radar grid was presented to the operators for 1,000 ms before each frame. The interframe interval was 1,500 ms.

Participants received feedback on their performance at the end of each block. If the operator had made an error, a tone sounded at the end of the frame. Operators also received text feedback at the end of each frame detailing the type of error they had committed (false alarms, misses, etc.).

Statistical Model. For each dependent variable—hits, false alarms, and detection time (for correct detections)—we ran three-way, repeated-measures analyses of variance (ANOVAs) using stimulus mapping (CM or VM), frame size (1 or 4), and memory set size (1 or 4) as within-subjects factors.

Results

The detection accuracy across participants (hits and false alarms) and the detection time (for correct detections) are plotted in Figure 2, averaged by conditions. Statistical analyses are shown in Table 1.

Participants performed significantly better at detecting targets under CM than under VM—that is, they had more hits, $F(1, 8) = 26.92, p < .01$, and fewer false alarms, $F(1, 8) = 23.23, p < .01$. Participants also performed better when maintaining one rather than four items in memory (hits, $F(1, 8) = 44.73, p < .01$; false alarms, $F(1, 8) = 11.08, p < .01$) and when searching for one rather than four potential targets (hits, $F(1, 8) = 17.55, p < .01$; false alarms, $F(1, 8) = 15.67, p < .01$). There were significant two-way interactions between stimulus mapping and frame size (hits, $F(1, 8) = 17.69, p < .01$; false alarms, $F(1, 8) = 6.89, p < .01$), between stimulus mapping and memory set size (hits, $F(1, 8) = 36.30, p < .01$; false alarms, $F(1, 8) = 17.67, p < .01$), and between frame size and memory set size (hits, $F(1, 8) = 32.91, p < .01$; false alarms, $F(1, 8) = 27.04, p < .01$).

There was a significant three-way interaction between frame size, memory set size, and stimulus mapping (hits, $F(1, 8) = 26.31, p < .01$; false alarms, $F(1, 8) = 12.35, p < .01$), suggesting, as shown in Figure 2, that increasing the task workload decreases the number of hits and increases the false alarms, particularly under VM.

In addition, participants responded faster to targets under CM than under VM, $F(1, 8) = 34.75, p < .01$; when maintaining one rather than four items in memory, $F(1, 8) = 206.91, p < .01$; and when searching for one rather than four potential targets, $F(1, 8) = 148.20, p < .01$. There were significant two-way interactions between stimulus mapping and frame size, $F(1, 8) = 38.92, p < .01$, and between frame size and memory set size, $F(1, 8) = 7.32, p < .05$. There was a statistically significant three-way interaction between stimulus mapping, frame size, and memory set size, $F(1, 8) = 41.31, p < .01$, indicating that, overall, increasing the task workload affects the VM condition more than the CM condition.

The triple interactions were analyzed with a set of planned comparisons and simple effects analyses. Planned comparisons indicated that under VM, performance

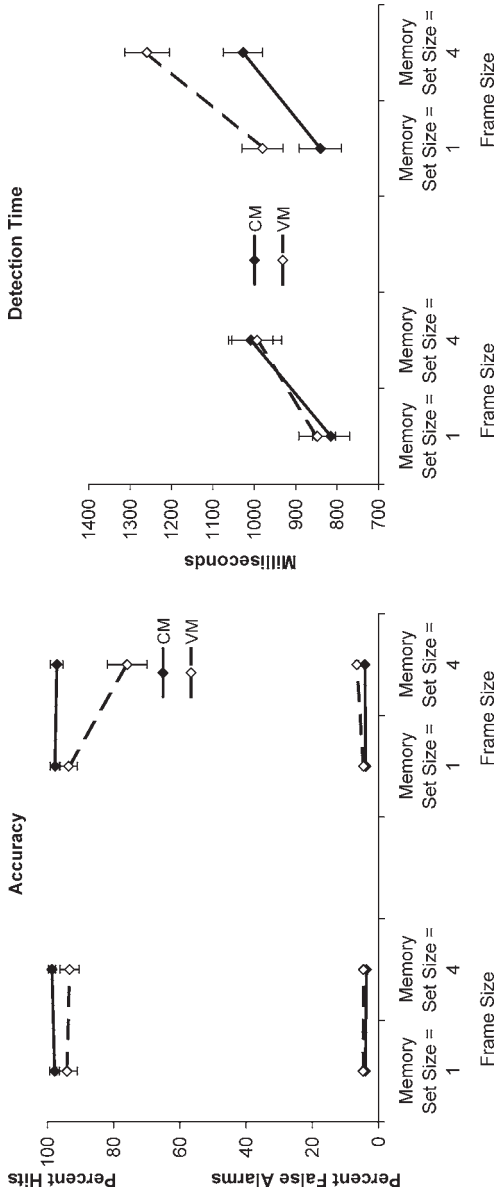


Figure 2. Results of Experiment 1. Average detection accuracy: hits and false alarms and average detection time plotted as a function of SM (CM or VM), FS (1 or 4), and MSS (1 or 4). CS = consistent mapping; FM = frame size; MSS = memory set size; SM = stimulus mapping; VM = varied mapping.

TABLE 1. Statistical Results From Experiment 1

| Variable | Hits | | False Alarms | | Detection Time (ms) | |
|--|---------------------------------|-------------|---------------------------------|-------------|---------------------------------|-------------|
| | <i>F</i> (<i>df</i> = 1, 8) | Effect Size | <i>F</i> (<i>df</i> = 1, 8) | Effect Size | <i>F</i> (<i>df</i> = 1, 8) | Effect Size |
| Stimulus mapping | 26.92** | .77 | 23.23** | .74 | 34.75** | .81 |
| Memory set size | 44.73** | .85 | 11.08** | .58 | 206.91** | .96 |
| Frame size | 17.55** | .69 | 15.67** | .66 | 148.20** | .95 |
| Stimulus mapping × Frame size | 17.69** | .69 | 6.89** | .46 | 38.92** | .83 |
| Stimulus mapping × Memory set size | 36.30** | .82 | 17.67** | .69 | 0.77 | .09 |
| Frame size × Memory set size | 32.91** | .80 | 27.04** | .77 | 7.32* | .48 |
| Frame size × Memory set size × Stimulus mapping | 26.31** | .77 | 12.35** | .61 | 41.31** | .84 |

p* < .05; *p* < .01.

TABLE 2. Summary of Simple Effects Analyses for Experiment 1

| CM and VM Contrasts | Means | | | |
|-------------------------------------|-------|-------|-----------------|-------------|
| | CM | VM | <i>F</i> (1, 8) | Effect Size |
| <i>Hits</i> | | | | |
| Frame size = 1, memory set size = 1 | .979 | .942 | 10.13* | .56 |
| Frame size = 1, memory set size = 4 | .987 | .934 | 12.72** | .61 |
| Frame size = 4, memory set size = 1 | .977 | .938 | 8.85* | .53 |
| Frame size = 4, memory set size = 4 | .973 | .759 | 31.86** | .80 |
| <i>False alarms</i> | | | | |
| Frame size = 1, memory set size = 1 | .037 | .043 | 1.385 | .29 |
| Frame size = 1, memory set size = 4 | .034 | .043 | 5.89* | .42 |
| Frame size = 4, memory set size = 1 | .036 | .042 | 18.49** | .69 |
| Frame size = 4, memory set size = 4 | .038 | .064 | 32.29** | .80 |
| <i>Detection time (ms)</i> | | | | |
| Frame size = 1, memory set size = 1 | 814 | 849 | 11.90** | .59 |
| Frame size = 1, memory set size = 4 | 1,008 | 993 | 0.34 | .04 |
| Frame size = 4, memory set size = 1 | 840 | 979 | 25.23** | .76 |
| Frame size = 4, memory set size = 4 | 1,027 | 1,258 | 46.43** | .85 |

Note. CM = consistent mapping; VM = varied mapping.

p* < .05; *p* < .01.

decreases monotonically as the product of memory set size and frame size increases (hits, $F(1, 8) = 36.85, p < .01$; false alarms, $F(1, 8) = 43.15, p < .01$; detection time, $F(1, 8) = 26.58, p < .01$). Under CM, however, these effects were not statistically significant (hits, $F(1, 8) = .778, ns$; false alarms, $F(1, 8) = .285, ns$; detection time, $F(1, 8) = .795, ns$). The results of simple effects analyses appear in Table 2. As these results show, the difference in performance between the CM and VM conditions is strongest under the high task workload (high frame size and high memory set size).

Summary

Experiment 1 replicated the results from Schneider and Shiffrin's (1977) first and second experiments. This replication demonstrated that the automaticity effect under CM conditions extends to dynamic visual search tasks. Unlike Schneider and Shiffrin's experiment, the task we used included targets that moved in the visual display under time constraints.

The fact that hits and false alarms are monotonically related to the product of memory set size and frame size in VM but not in CM suggests that two different search processes—*controlled search* (serial search process) and *automatic detection* (parallel search), respectively—are at work (Schneider & Shiffrin, 1977). The low rate of false alarms indicates accuracy in the processing of the individual characters' features.

These results demonstrate automaticity development in the visual and memory search components of the radar task, which involves dynamic targets and distractors. Our primary goal, however, was to demonstrate the effects of automatic and controlled processing on accuracy in the decision-making component of the radar task. Experiment 2 addressed this goal.

Experiment 2

We expected that the use of controlled or automatic processing in the visual and memory search components of the radar task would influence the accuracy of the decision-making component of the task. More concretely, we expected that automatic processing in the visual and memory search task would benefit the decision-making task. We predicted that individuals would make the poorest decisions under controlled search in the target detection task—that is, when they were exposed to VM and high task workload. We also predicted that they would make the best decisions under automatic search—that is, when they were exposed to CM in the visual and memory search task.

Furthermore, we expected to demonstrate that better decision making under CM would be attributable to the release of cognitive resources from the visual search task and extra time for performance of the decision task. Because the visual and memory search task and decision-making components of the radar task are sequential and decisions can be made only after the target is detected, it is possible that an improvement in decision-making accuracy is caused by the time released from the visual search task under CM. Similarly, decision accuracy may be degraded by limited time for performance when the visual search task is under VM because participants may take longer to detect targets under VM and high workload conditions.

Following the capacity view of automaticity (Moors & De Houwer, 2006) and original developments of dual-process theory of automaticity (Schneider & Shiffrin, 1977), we expected that a single attention pool would be constrained by multitasking. That is, if the benefit of CM over decision making simply is caused by the time released from the visual search task, then decision-making performance will be affected when that time is used to perform another task in parallel.

In contrast, if the benefit of CM on decision making is caused by a release of cognitive resources, then the addition of a parallel task will not affect decision accuracy. We tested this possibility by manipulating the presence or absence of the performance of another task in parallel after the visual search task. Under VM and high task workload, we expected poor decision making regardless of dual-task conditions; under CM, we expected task workload to have no effect on decision performance regardless of dual-task conditions.

Method

Participants. Seven participants (3 women and 4 men), all graduate and undergraduate students, were recruited from local universities through bulletin boards. All participants were right-handed, had normal color vision, and had corrected or normal visual acuity. Average age was 23 years ($SD = .93$). The study consisted of six experimental sessions that took place on 6 consecutive days. Each session lasted approximately 3.5 hr. Participants were paid \$11.00 per hour for their participation. To avoid any fatigue effects, we encouraged individuals to take up to three breaks of 5 min each in each session.

Experimental Design. This experiment employed the full version of the radar task, including both the visual and memory search tasks and the decision-making components. The experimental design of the visual search component was the same as in Experiment 1. We manipulated stimulus mapping (CM and VM), frame size (1 and 4), and memory set size (1 and 4) variables within subjects. Each of the six sessions comprised eight blocks of 20 trials each. We increased the number of trials from 16 to 20 to allow individuals more practice than they were afforded in Experiment 1 because the participants would be required to perform the decision-making component of the task in addition to the visual search component. Blocks corresponding to one of the eight conditions were randomly presented in each session.

All participants were asked to accomplish both the visual search and the decision-making components of the task in each trial. In addition, during the last three sessions, participants were asked to accomplish an extra auditory task, called the *tone task*, involving the detection and counting of nonstandard tones that are higher or lower in pitch than a standard tone. The tone task was performed in parallel with the visual search and decision-making components of the radar task.

In addition to hits, false alarms, and detection time (for correct detections), we also analyzed *decision accuracy* for the decision-making task and *tone accuracy* for the tone task. Decision accuracy refers to the percentage of correct decisions made (use of guns, missiles, or ignore) given the sensors' readings in each particular trial. Tone accuracy denotes the percentage of correct counting of nonstandard tones. The tone and decision variables are elaborated in the next section.

Procedure. The process of the target detection portion of the task was the same as in Experiment 1. When a target was detected, the operator pressed the space bar,

and the sensors' readings were displayed on the right side of the screen (see Figure 1). Participants were asked to assess the top two threat scan sensors described in the radar task, ECM jam strength and weaponry threat. If the threat sensors added to a score of 4 or fewer tick indicators, the target was considered nonhostile, and the correct answer was to ignore the target (press 5 on the numeric keypad). If the scan readings added to 5 or more, then the next two weapon system sensors, armor and size, had to be evaluated before an action could be determined. If the sum of the two weapon system sensors was low (less than or equal to 9), the correct response was a gun shift; if the sum of both sensors was high (greater than 9), the correct response was an active missile shift.

Participants received feedback on their performance at the end of each trial, indicating how well they used the scoring rules. Participants were instructed on how to apply the sensor rules and practiced detecting targets before the first session began.

The tone task involved the detection and counting of nonstandard tones that are higher or lower in pitch than a standard tone. The difference in pitch was determined by calibrating each participant individually before the experiment. The standard tone was played during every presentation of the target list (memory set). The standard tone was constant throughout the experiment. The tones were presented for 200 ms, with an average interstimulus interval of 1,000 ms (50% deviation). At the end of each trial, participants were asked to enter the total number of nonstandard tones. They were told to pay equal attention to the tone task and to the radar task. Participants received feedback on their performance.

Statistical Models. We averaged the dependent variables per condition and separated the average into the first three sessions (no tone task) and the last three sessions (tone task). For the dependent variables—hits, false alarms, and detection time (for correct detections)—we ran three-way, repeated-measures ANOVAs using stimulus mapping (CM or VM), frame size (1 or 4), and memory set size (1 or 4) as within-subjects variables. To determine the effects on decision accuracy, we used the same model with stimulus mapping, frame size, and memory set size over decision accuracy as the dependent variable. To determine the effect of the tone task on the visual search and decision-making components, we ran the same model using the last three sessions, in which participants accomplished this extra auditory task. We ran the same analyses for hits, false alarms, detection time, decision accuracy, and tone accuracy for those sessions.

Results

The detection accuracy across participants (hits and false alarms) and the detection time (for correct detections) for the first three sessions, averaged per condition, are plotted in Figure 3; the corresponding statistical analyses are shown in Table 3. These results essentially replicated those of Experiment 1.

Participants performed significantly better at detecting targets under CM than under VM—that is, they had more hits, $F(1, 6) = 43.99, p < .01$; fewer false alarms, $F(1, 6) = 35.51, p < .01$; and faster response times, $F(1, 6) = 69.15,$

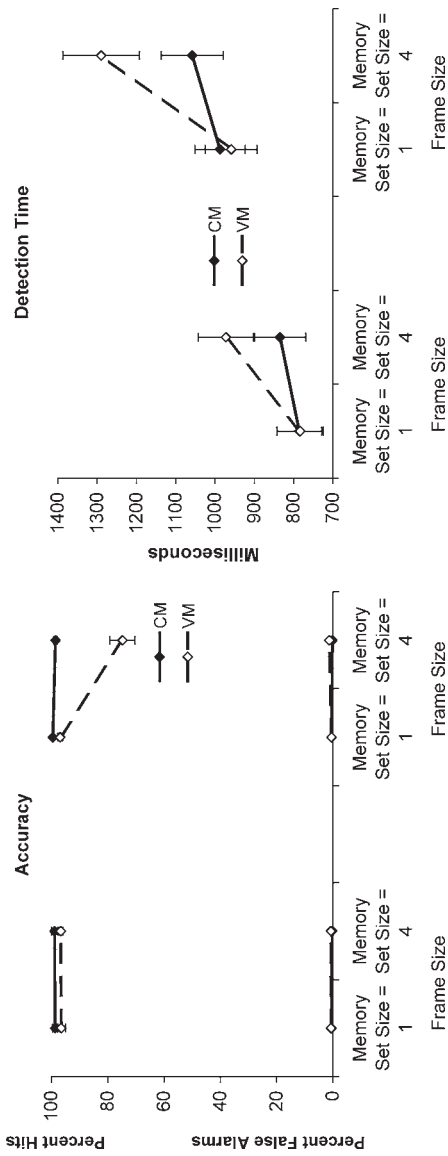


Figure 3. Results of Experiment 2. Average detection accuracy: hits and false alarms and average detection time plotted as a function of SM (CM or VM), FS (1 or 4), and MSS (1 or 4). CS = consistent mapping; FM = frame size; MSS = memory set size; SM = stimulus mapping; VM = varied mapping.

TABLE 3. Statistical Results From Experiment 2

| | Hits | | False Alarms | | Detection Time (ms) | | Decision Accuracy | |
|---|------------------|----------------|------------------|----------------|---------------------|----------------|-------------------|----------------|
| | F (df = 1, 6) | Effect Size | F (df = 1, 6) | Effect Size | F (df = 1, 6) | Effect Size | F (df = 1, 6) | Effect Size |
| Stimulus mapping | 43.99** | .88 | 35.51** | .86 | 69.15** | .92 | 73.06** | .92 |
| Memory set size | 32.18** | .84 | 0.39 | .06 | 86.99** | .94 | 9.39* | .61 |
| Frame size | 73.97** | .93 | 0.89 | .13 | 106.41** | .95 | 28.32** | .83 |
| Stimulus mapping × Frame size | 21.99** | .79 | 5.01 | .46 | 4.69 | .44 | 13.10* | .69 |
| Stimulus mapping × Memory set size | 47.57** | .89 | 25.00** | .81 | 179.94 | .97 | 26.98** | .82 |
| Frame size × Memory set size | 24.00** | .80 | 9.75* | .62 | 16.34** | .73 | 17.95** | .75 |
| Frame size × Memory set size × Stimulus mapping | 25.19** | .81 | 8.34* | .58 | 16.69** | .74 | 61.15** | .91 |

* $p < .05$; ** $p < .01$.

$p < .01$. Participants also had more hits and responded faster when maintaining one rather than four items in memory (hits, $F(1, 6) = 32.18$, $p < .01$; detection time, $F(1, 6) = 86.99$, $p < .01$) and when searching for one rather than four potential targets (hits, $F(1, 6) = 73.97$, $p < .01$; detection time, $F(1, 6) = 106.41$, $p < .01$). In addition, there were significant two-way interactions between stimulus mapping and frame size (hits, $F(1, 6) = 21.99$, $p < .01$), between stimulus mapping and memory set size (hits, $F(1, 6) = 47.57$, $p < .01$; false alarms, $F(1, 6) = 24.98$, $p < .01$), and between frame size and memory set size (hits, $F(1, 6) = 23.99$, $p < .01$; false alarms, $F(1, 6) = 9.75$, $p < .05$; detection time, $F(1, 6) = 16.34$, $p < .01$).

There was a significant three-way interaction between frame size, memory set size, and stimulus mapping, suggesting that, overall, increasing the task workload affects the VM condition more than the CM condition (hits, $F(1, 6) = 25.19$, $p < .01$; false alarms, $F(1, 6) = 8.34$, $p < .05$; detection time, $F(1, 6) = 16.69$, $p < .01$).

The triple interactions were examined using a set of planned comparisons and simple effects analyses. Planned comparisons indicated that under VM, performance decreases monotonically as the product of memory set size and frame size increases (hits, $F(1, 6) = 24.74$, $p < .01$; false alarms, $F(1, 6) = 16.98$, $p < .01$; detection time, $F(1, 6) = 22.08$, $p < .01$). Under CM, however, these effects were not statistically significant (hits, $F(1, 6) = 5.39$, ns ; false alarms, $F(1, 6) = 0.31$, ns ; detection time, $F(1, 6) = 1.39$, ns). The results of simple effects analyses are shown in Table 4. The results show that the contrast in performance between CM and VM conditions is strongest under high task workload (high frame size and high memory set size).

TABLE 4. Summary of Simple Effects Analyses for Experiment 2

| CM and VM Contrasts | Means | | | |
|-------------------------------------|-------|-------|----------|-------------|
| | CM | VM | F(1, 6) | Effect Size |
| <i>Hits</i> | | | | |
| Frame size = 1, memory set size = 1 | .987 | .965 | 10.04* | .63 |
| Frame size = 1, memory set size = 4 | .995 | .969 | 10.68* | .64 |
| Frame size = 4, memory set size = 1 | .988 | .966 | 5.32 | .47 |
| Frame size = 4, memory set size = 4 | .986 | .747 | 36.07** | .86 |
| <i>False alarms</i> | | | | |
| Frame size = 1, memory set size = 1 | .005 | .006 | 2.04 | .25 |
| Frame size = 1, memory set size = 4 | .005 | .006 | 0.125 | .02 |
| Frame size = 4, memory set size = 1 | .003 | .007 | 20.34** | .77 |
| Frame size = 4, memory set size = 4 | .002 | .013 | 44.02** | .88 |
| <i>Detection time (ms)</i> | | | | |
| Frame size = 1, memory set size = 1 | 785 | 783 | 0.06 | .00 |
| Frame size = 1, memory set size = 4 | 987 | 959 | 6.51* | .52 |
| Frame size = 4, memory set size = 1 | 834 | 972 | 47.18** | .89 |
| Frame size = 4, memory set size = 4 | 1,057 | 1,289 | 109.69** | .95 |
| <i>Decision accuracy</i> | | | | |
| Frame size = 1, memory set size = 1 | .815 | .757 | 10.49* | .64 |
| Frame size = 1, memory set size = 4 | .778 | .763 | 1.22 | .17 |
| Frame size = 4, memory set size = 1 | .807 | .760 | 4.02 | .40 |
| Frame size = 4, memory set size = 4 | .758 | .532 | 176.98** | .97 |

Note. CM = consistent mapping; VM = varied mapping.
 * $p < .05$; ** $p < .01$.

Average decision accuracy per condition is plotted in Figure 4; the corresponding statistical results are shown in Table 3. The results indicate that participants made more accurate decisions (correct use of guns, missiles, or ignore actions according to the sensors' readings in each trial) when the target detection task used CM rather than VM, $F(1, 6) = 73.06, p < .01$; when the target detection task required the participant to maintain one rather than four items in memory, $F(1, 6) = 9.39, p < .05$; and when the target detection task required the participant to search for one rather than four potential targets, $F(1, 6) = 28.32, p < .01$. In addition, there were significant two-way interactions between stimulus mapping and frame size, $F(1, 6) = 26.98, p < .01$; between stimulus mapping and memory set size, $F(1, 6) = 13.10, p < .05$; and between frame size and memory set size, $F(1, 6) = 17.95, p < .01$.

There was a significant three-way interaction between frame size, memory set size, and stimulus mapping, $F(1, 6) = 61.15, p < .01$, suggesting that individuals make more accurate decisions when they are working under CM and low task workload than when they are working under VM and high task workload in the target detection task.

Once again, planned comparisons indicated that when individuals are under VM in the target detection task, their decision-making performance decreases

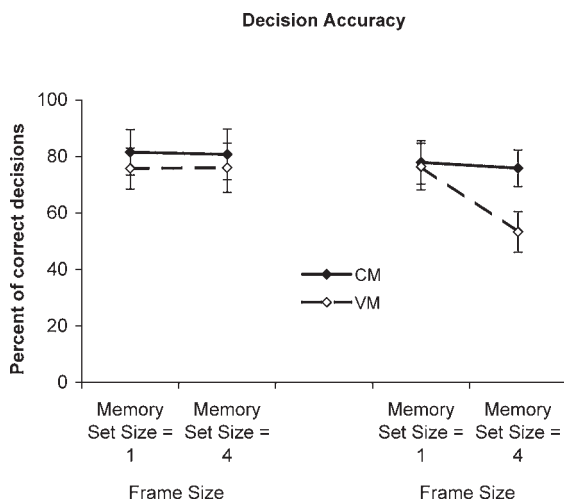


Figure 4. Decision accuracy per condition.

monotonically as the product of memory set size and frame size increases, $F(1, 6) = 35.89, p < .01$. Decision-making accuracy was not affected by task workload under CM in the target detection task, $F(1, 6) = 0.27, ns$. The results of simple effects analyses are also shown in Table 4. Once again, the strongest difference in decision accuracy between CM and VM conditions occurred under high task workload (high frame size and high memory set size).

Source of Improvement in Decision Accuracy

To test for the possibility that the improvement in decision accuracy under CM was not caused by only reduced time in the detection task, we analyzed the presence or absence of dual-task conditions as another within-subjects variable. The first three sessions were performed without dual-task conditions, the last three sessions with dual-task conditions. Decision-making performance in the CM condition was expected to show insensitivity to the presence of the tone task, whereas decision-making performance in the VM condition was not. If supported, these expectations would demonstrate that decision-making performance is related to not only the detection time reduction in a trial but also the cognitive resources available under CM conditions.

Decision accuracy, averaged by mapping and dual-task conditions, is plotted in Figure 5. As expected, statistical results indicate that the dual-task condition has a negative effect in the VM condition, $F(1, 6) = 24.80, p < .01$, but no effect in the CM condition, $F(1, 6) = 0.14, ns$.

The participants' tone accuracy was worst under VM and high workload (see Figure 6). The ANOVA results indicated that the number of correct identifications of tones was higher when participants were responding to targets under CM and when they were maintaining one rather than four items in memory. There was a

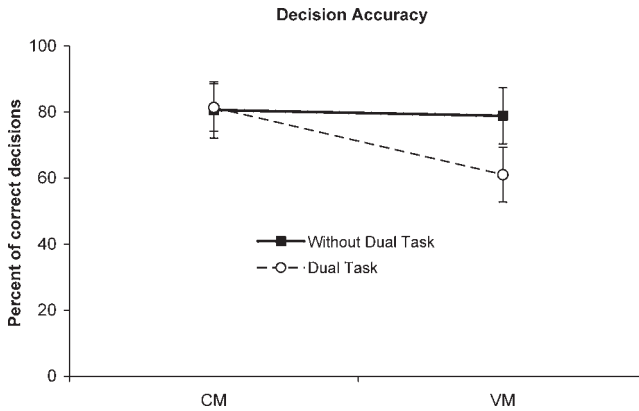


Figure 5. Decision accuracy in normal and dual-task conditions.

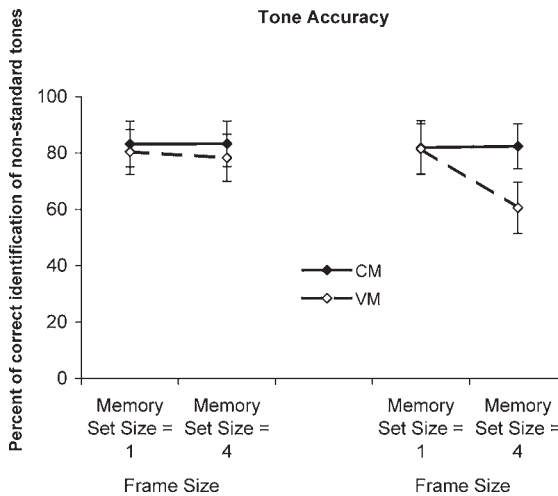


Figure 6. Tone accuracy.

significant two-way interaction between memory set size and stimulus mapping, $F(1, 6) = 9.43, p < .05$, and a significant three-way interaction between memory set size, frame size, and stimulus mapping, $F(1, 6) = 10.38, p < .05$. The analyses support the hypothesis that participants performed poorest under the condition with the highest workload and VM in the memory and visual search component of the radar task.

Summary

This study demonstrated that automatic detection in the visual search component of the radar task influenced decision accuracy in the decision-making component of

the task. This is an important test of the CM, VM, and workload principles of Schneider and Shiffrin's (1977) automaticity theory because multitasking with dynamic stimuli is a fundamental characteristic of complex tasks. We also demonstrated that the parallel performance of another task did not influence these results, indicating that a release of cognitive resources caused by CM in the visual search component of the task may also contribute to decision accuracy.

Discussion

In addition to demonstrating the dual-process theory of automaticity in a dynamic visual and memory search task, we demonstrated that the conditions that lead to controlled and automatic target detection directly influence the decision-making component of radar task performance in similar ways.

A main finding from this research is that, although the decision-making and auditory tasks were not consistently mapped in any way, they exhibited the benefits of the CM manipulation in the visual and memory search components of the radar task. Decisions were more accurate when the visual and memory search components were consistent rather than variably mapped. This benefit of CM over VM for decision accuracy was particularly strong under high workload. The second experiment demonstrates that in compound tasks, components that require controlled processing (e.g., decision making) may benefit from the automaticity developed under other components of the task.

We believe this is an important contribution to our current knowledge of automaticity in complex tasks. If some components of complex tasks operate in automatic mode, then other, more complex components may also benefit from the automaticity of simple subtasks. In this case, decisions were more accurate when a simple preceding visual search subtask was automatic, and they were less accurate when the preceding task was performed under controlled processing. It is important to note again that there was no CM training in the decision-making component of the task. The sensor values were randomized for each trial, and the correct response action (guns, missiles, or ignore) depended on the real-time readings on the sensors, which varied in each trial. Nevertheless, the accuracy of the decision-making component showed the same pattern of results as the task component in which CM and VM training took place.

One interpretation of these results indicates that because CM training resulted in faster responses in the visual search task, the extra time was directed to the decision-making component (the tasks were sequential in the radar simulation), resulting in more accurate decisions. However, the results based on testing with the additional auditory task suggest that extra time alone is unlikely the only reason for the higher decision-making accuracy. The pattern of decision-making results was the same with and without the additional tone task in the CM condition. Furthermore, the analysis of accuracy in the tone task indicates that all of the components of the radar task were performed together, with little interference among them. Participants did not use any compensatory strategies (Schneider, 1985)

and were able to perform both tasks under CM conditions, providing additional support for automaticity. The decrement in performance in the tone task was attributable only to the experimental manipulation—specifically, the VM condition and high workload.

Another interpretation of these results points to a transfer of automaticity. Schneider and Fisk (1984), for example, investigated the transfer of training in a search task and found that transfer was not specific to the elements of the training category but generalized to untrained members of the same category. Under high task workload, untrained members of the CM category presented substantial positive transfer. Related to this finding, our results suggest that the benefits of automaticity from the trained CM visual search tasks might extend to novel components of the task that are not mapped consistently. The decision-making component in the radar task is of a different structure compared with the search task, and our results suggest that CM training on a subtask can result in substantial benefits to other components of the same task, even when the category knowledge or task structure is different.

We interpret these results as an individual's ability to allocate his or her cognitive resources flexibly to different subtasks. Individuals may have been able to use the categories and rely on similar instances in the search task when the mapping was consistent. This consistency and similarity released cognitive resources that the individuals were able to use for efficient categorization of the sensors and actions. This explanation agrees with the capacity view of automaticity, which has been proposed recently and is supported by the major theories of automaticity (Moors & De Houwer, 2006).

The findings from this research have direct application to dynamic real-world tasks. For example, there are many cases in which patriot missile systems have targeted and actually shot down friendly fighters (Snook, 2000), and CM features likely differentiate enemy missiles from friendly fighter aircraft in theater defense. It seems plausible that radar characteristics are indicative of missiles but not friendly fighters. Thus, our findings may justify the development of radar technologies that are sensitive to such characteristics, the creation of symbols to display these characteristics to operators, and training and practice for operators to detect these symbols efficiently.

We are interested in other aspects of dynamic decision-making tasks that may influence the role of automaticity. From a practical perspective, it is important to understand whether a subtask is fundamental to the goal of the overall task. Task developers should map subtasks (i.e., make CM cues salient for subtasks) consistently if there is strong task-goal coupling (i.e., the subtask is highly relevant to the overall task goal) so that automatic detection can develop with experience. Operators who already have extensive practice differentiating two predefined categories may benefit from automatic detection when one category is mapped consistently to targets, even though they have no actual practice on the task.

Finally, in dynamic decision-making tasks, targets are often defined by the conjunction of several cues rather than a single cue. Although the mapping of

cues to target and distractor categories can change dynamically, the mappings probably do not shift on a trial-to-trial basis. Thus, it would be advantageous to investigate whether operators trained under CM and VM of conjunction search targets can employ appropriate strategies to efficiently detect targets that represent real-world dynamic tasks. These are possibilities we are currently investigating.

Acknowledgments

This research was supported by the Multidisciplinary University Research Initiative Program (MURI), Grant No. N00014-01-1-0677.

References

- Bargh, J. A., & Ferguson, M. J. (2000). Beyond behaviorism: On the automaticity of higher mental processes. *Psychological Bulletin*, *126*, 925–945.
- Bolstad, C. A., & Endsley, M. R. (2000, July–August). *The effect of task load and shared displays on team situation awareness*. Paper presented at the 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society, Marietta, GA.
- Fisk, A. D., & Hodge, K. A. (1992). Retention of trained performance in consistent mapping search after extended delay. *Human Factors*, *34*, 147–164.
- Fisk, A. D., & Schneider, W. (1983). Category and work search: Generalizing search principles to complex processing. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *9*, 177–195.
- Kaber, D. B., & Endsley, M. R. (1997). The combined effect of level of automation and adaptive automation on human performance with complex, dynamic control systems. In *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting* (pp. 205–209). Santa Monica, CA: Human Factors and Ergonomics Society.
- Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice Hall.
- Kahneman, D., & Chajczyk, D. (1983). Tests of the automaticity of reading: Dilution of Stroop effects by color-irrelevant stimuli. *Journal of Experimental Psychology: Human Perception & Performance*, *9*, 497–509.
- Kramer, A. F., Strayer, D. L., & Buckley, J. (1991). Task versus component consistency in the development of automatic processing: A psychophysiological assessment. *Psychophysiology*, *28*, 425–437.
- Lassaline, M. L., & Logan, G. D. (1993). Memory-based automaticity in the discrimination of visual numerosity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*, 561–581.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492–527.
- Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, *132*, 297–326.
- Schneider, W. (1985). Training high-performance skills: Fallacies and guidelines. *Human Factors*, *27*, 285–300.
- Schneider, W., & Fisk, A. D. (1984). Automatic category search and its transfer. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *10*, 1–15.
- Schneider, W., & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search and attention. *Psychological Review*, *84*, 1–66.

- Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, 84, 127–190.
- Snook, S. A. (2000). *Friendly fire: The accidental shootdown of U.S. Black Hawks over northern Iraq*. Princeton, NJ: Princeton University Press.

Cleotilde (Coty) Gonzalez is an associate research professor and director of the Dynamic Decision Making Laboratory at Carnegie Mellon University. Her research interests include the study of decision making in dynamic environments. She follows experimental methods using decision-making games and cognitive computational models to understand the decision-making process.

Rickey P. Thomas is an assistant professor whose primary research focus involves the application of memory models and experimental methods from memory paradigms to study the role of memory processes in judgment and decision-making phenomena.