Cognitive, Affective, & Behavioral Neuroscience
2009, 9 (1), 71-82
doi:10.3758/CABN.9.1.71

Electromyographic evidence for response conflict in the exclude recognition task

Travis L. Seymour
University of California, Santa Cruz, Santa Cruz, California

and

Eric H. Schumacher
Georgia Institute of Technology, Atlanta, Georgia

How do memory retrieval processes lead to overt responses in strategic recognition tasks (responding "old" to one class of familiar stimulus and "new" to another)? Many current theories of memory retrieval ignore the response requirements in such memory tasks, instead modeling them using memory processes (e.g., familiarity and recollection) alone (see Yonelinas, 2002). We argue that strategic recognition involves conflict in response processing similar to canonical conflict tasks (e.g., the Stroop task). The parallel task set (PTS) model (Seymour, 2001) accounts for performance in strategic recognition tasks (e.g., the exclude recognition task) by suggesting that motor response conflict occurs when one responds "new" to familiar stimuli. We tested this prediction using surface electromyography, a measure incontrovertibly related to motor execution. Overall, results are consistent with the PTS model's assumption that recognition, motor, and control processes interact in strategic retrieval tasks. The implications of these data for models of memory retrieval and response conflict are discussed.

Although we may often recognize many stimuli in our environment, it is not always appropriate to acknowledge this recognition. Whether feigning ignorance of confidential information or denying familiarity with a particular stock to maximize personal profit, we are often able to choose which recognitions to reveal. This leads to the following question: How do these familiarity and recollective memory processes interact with control and motor processes to help us produce only those recognition responses appropriate given our current goals?

One experimental task often used for studying these memory phenomena in the laboratory is the *exclude recognition* task (e.g., Jacoby, Kelley, Brown, & Jasechko, 1989). In it, participants study two lists of words, *List 1* and *List 2*, and then perform an old/new judgment task during which they must respond "old" to words from one list—usually, List 2 (e.g., Yonelinas & Jacoby, 1994)—and "new" to words from the other list (e.g., List 1), as well as responding to new *filler* words. Thus, in order to respond accurately, participants must consider the source of each word and acknowledge their recognition of List 2 words while rejecting words from List 1. Typically, results show that classification of *target* words (from List 2) and new filler words is more accurate than that for *probe* words (from List 1) (Jacoby, 1991; Jacoby, Kelley, et al., 1989). In addition, responses to probe stimuli in this procedure tend to be slower than those on target and filler trials (Jacoby, Kelley, et al., 1989; Seymour & Kerlin, 2008; Seymour, Seifert, Shafto, & Mosmann, 2000). This behavioral pattern is shown in Figure 1 (solid bars) for an exclude recognition experiment (Seymour, 2001; Seymour et al., 2000). It has been replicated using both verbal and visual stimuli in exclude tasks (Seymour & Kerlin, 2008; Seymour et al., 2000).

Memory Processing in the Exclude Recognition Task

Despite the importance of understanding how memory and response processes interact, most theories of exclude recognition performance focus mainly on the recognition memory components. For example, dual-process models of recognition memory posit a complementary system in which both fast familiarity and slower recollection (or search) processes can contribute to responses. Performance is explained as a result of either familiarity alone or a combination of familiarity and recollection. In some cases, the familiarity process is sufficient, but other times a recollection or search process using episodic information is necessary to identify previously studied items (e.g., Atkinson & Juola, 1974; Gillund & Shiffrin, 1984). In other dual-process variations, both processes are typically involved unless recollection fails, leaving familiarity to drive the response alone (Jacoby, 1991; Yonelinas, 2002).

As an example, consider how one influential dual-process model, developed by Jacoby and colleagues (Jacoby, 1991; Yonelinas, 2002), accounts for performance in the exclude recognition task. By employing a *process dissociation* procedure, Jacoby and colleagues have examined the relative contributions of recollection and fa-

T. L. Seymour, nogard@ucsc.edu; E. H. Schumacher, eschu@gatech.edu

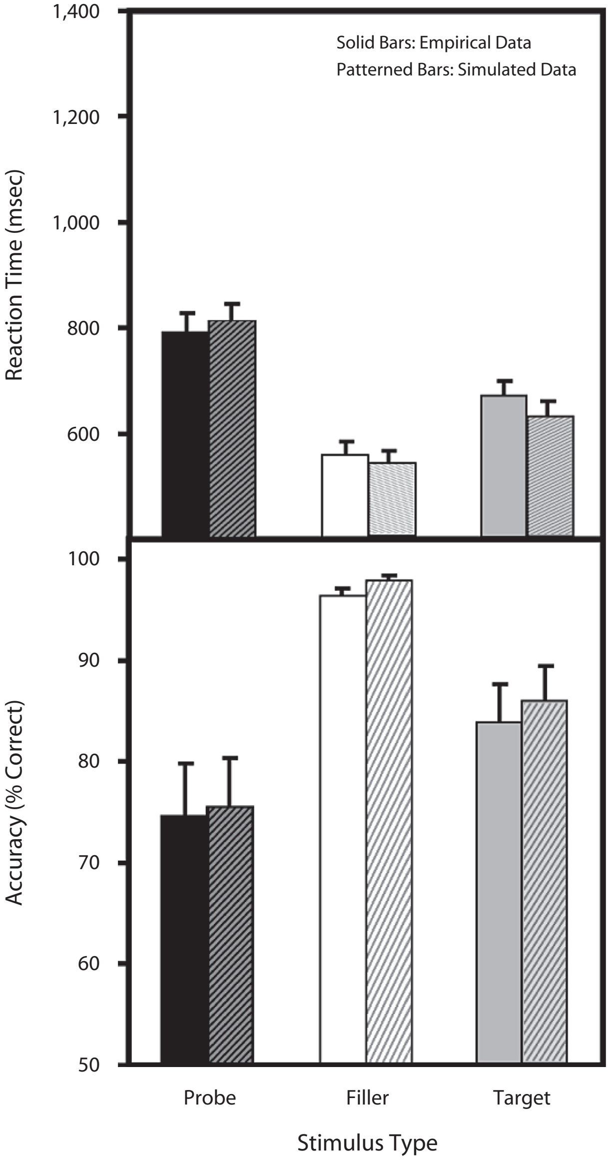

Figure 1. Reaction time and accuracy data from Seymour (2001). Solid bars represent empirical data, and patterned bars represent data simulated by the parallel task set model instantiated in the executive process interactive control architecture. Error bars represent the standard errors of the means.

miliarity memory processes in strategic recognition tasks. As in other models (Dosher, 1984; Gronlund & Ratcliff, 1989; Hintzman, Caulton, & Levitin, 1998; Hintzman & Curran, 1994; McElree, Dolan, & Jacoby, 1999; Ratcliff & McKoon, 1989), the familiarity process is fast and accurate and produces a positive result if the level of familiarity exceeds some threshold. Recollection, on the other hand, is slower and retrieves episodic information about the item (i.e., its source), if available; however recollection may fail completely, in which case no contribution is offered to the recognition decision. The two processes are independent and operate in parallel, with familiarity reaching a decision more quickly than recollection. Furthermore, whereas familiarity is automatic, the recollection process is assumed to require conscious control (Jacoby, Toth, & Yonelinas, 1993; Yonelinas, 2002).

According to this model, responses can be based on a combination of recollection and familiarity or on familiarity alone if the recollection process fails. In the exclude recognition task, both probe (i.e., responded to with a "new" response) and target (i.e., responded to with an "old" response) words are familiar but require different responses. Thus, in order to yield correct responses, the recollection process (or both recollection and familiarity) must return accurate information. Thus, mistakenly responding "old" on probe trials can occur when the response is driven by familiarity (which would indicated an incorrect "old" response) unopposed by any contribution from recollection or if both processes fail. If the probe word's source (e.g., "List 1") is successfully recollected, the opposing familiarity information is discounted (i.e., has less influence on the recognition decision than does recollection information) and results in the correct "new" response (Dodson & Johnson, 1996; Yonelinas & Jacoby, 1994). Although not intended as a complete model of recognition, Jacoby's model is similar to other successful dual-process models of recognition that focus on the relative contributions of fast familiarity and slower recollective memory processes (for a review of recognition models, see Yonelinas, 2002).

In spite of the success of the Jacoby model's account of exclude recognition performance, its focus on the contribution of memory processes has limited its account to accuracy data. Although some work has examined the time course and relative contributions of familiarity and recollection processes on fast and slow trials (e.g., McElree et al., 1999; Yonelinas & Jacoby, 1994), the model is currently without a precise account of the reaction time (RT) data. However, a complete understanding of memorial behavior requires understanding how memory and motor processes interact (for an example of this in working memory, see Kieras, Meyer, Mueller, & Seymour, 1999).

This is a critical gap in our current knowledge of retrieval of information from long-term memory. It is important to understand how final response decisions ("old" or "new") are translated into overt behavior, the degree to which the motor system may constrain probe responses, and the influence of probe responses on subsequent trial responses. Some information about how these additional motor processes may interact with memorial ones may come from the literature investigating response processing in relatively easy perceptual–motor tasks.

One experimental procedure that embodies an integration of performance/evaluation and response processes is the Stroop task (Stroop, 1935). In this task, participants are sometimes asked to ignore the orthography of color words and name an incongruent ink color instead (e.g., the word RED written in green ink). In this case, it is necessary to inhibit the prepotent tendency to read the color word, which competes with the ink-naming response. In particular, theorists argue that the vocal response activated by the word (e.g., the vocal command to say "red") interferes with the vocal response simultaneously activated by the ink color (e.g., "green") (see, e.g., Cohen, Dunbar, & McClelland, 1990).

Response conflict in the Stroop task is well established in the literature (for a review, see MacLeod, 1991), yet the existence of similar conflict dynamics in the exclude recognition task has been the subject of much less investigation. We propose that responses on probe trials may be slower than those for targets and fillers because, on these trials, "old" and "new" motor responses compete with one another.

Response Conflict in Exclude Recognition

If response conflict exists in the exclude recognition task, theories of response conflict may offer some insight into how memory and motor processes interact. One influential theory, the *conflict-monitoring* theory, models performance in Stroop and other perceptual–motor tasks as a dynamic interaction between response selection processes and those that detect and manage response conflict (Botvinick, Braver, Barch, Carter, & Cohen, 2001). Response conflict involves the simultaneous activation of incompatible response tendencies that occurs when a prepotent tendency must be overridden in order to make a correct response. For example, in the Stroop task, the prepotent reading response must be overridden in favor of the color-naming response.

Conflict-monitoring theory posits that conflict-monitoring processes (likely mediated by the anterior cingulate cortex [ACC] or adjacent medial frontal regions) signal the presence of multiple conflicting responses to an environmental stimulus. This signal induces control processes (in other areas of the association cortex) to facilitate strategic control when necessary (e.g., to emit the correct response on incongruent Stroop trials; see, e.g., Botvinick et al., 2001; Botvinick, Cohen, & Carter, 2004). The computational model engendering this theory has been quite successful in modeling response conflict in a wide range of tasks (see Botvinick et al., 2001).

This theory, however, has *not* been used to model behavior in the exclude recognition task. A different computational model, the parallel task set (PTS) model, however, *has* been applied to the exclude recognition task (Seymour, 2001). Using the production rule formalism of the executive process interactive control (EPIC) architecture (Meyer & Kieras, 1997), the PTS model combines the independent familiarity and recollection processes from the Jacoby and related recognition models (Yonelinas, 2002) with the interaction of conflict monitoring and related control processes from conflict-monitoring theory (Botvinick et al., 2001). This model was designed to be a general framework for explaining how memory processes interact with task strategy when retrieval of different subsets of information is mapped to different response outcomes (i.e., strategic recognition tasks). Its goal is to complement the conflict-monitoring theory's accounts of dynamical system interaction and recruitment of control processes with a stage-by-stage information-processing account of the individual processes involved in task performance (Seymour, 2001).

An overview of the PTS model is depicted in Figure 2. The PTS model combines memory, response selection, response preparation, and response execution processes into a *task set* (see Monsell, 1996) that describes each processing stage from stimulus encoding to the execution of an overt response. The familiarity task set (Figure 2, light gray boxes) and the recollection task set (Figure 2, dark gray boxes) operate independently and in parallel. Although the retrieval of some initial recollective information is assumed to be automatic following presentation of a familiar stimulus, the additional stages in the recollection task set are under conscious control. The familiarity process, on the other hand, as well as subsequent response selection, motor programming, and response execution, is assumed to be automatic. Whereas, on average, the recollection process is completed more slowly than the familiarity process, it may also be more variable. Despite the asynchronous operation of parallel task sets (both of which contain motor preparation and execution stages), only one overt response is made to each stimulus item. If two different motor preparation requests (such as "old" and "new") reach the manual motor processor simultaneously, or if one preparation request is sent to the manual motor processor while a different one is already being prepared, response conflict occurs. Because familiarity is faster, on average, than recollection, the first preparation to reach the manual motor processor will often (although not always) be the familiarity-based response. Thus, a model that merely waited until the manual motor processor was free to request the recollection-based response preparation would guarantee that an incorrect response would be made on most probe trials. Rather, the PTS model suggests that participants proceed with their conscious recollection-based strategy and a recollection-based response preparation request is sent to the manual motor processor without regard to whether it is already preparing a familiarity-based response. This suggests that on probe trials, response conflict will occur because the "old" response initiated by the familiarity task set is at odds with the "new" response suggested by the recollection task set.

The PTS model posits a task-invariant response conflict detection process that activates conflict resolution processes when response conflict occurs (see Botvinick et al., 2001; Carter et al., 1998; Gehring, Goss, Coles, Meyer, & Donchin, 1993). In this model, response conflict is explicitly defined as occurring when a motor preparation request is sent to the manual motor processor while a different one is already underway. The conflict resolution process attempts to avoid this state and promotes execution of the intended response (although this determination is not always accurate). When motor conflict is detected, the new preparation request is suspended while an attempt to abort the old preparation is made. The assumption of this process is simply that the newer preparation attempt is a correction and, thus, the "intended" response. If the abort fails because the preparation is too far along or already completed (e.g., Logan, 1994), the initial response continues to be prepared and executed (i.e., is uninterrupted), and the newer response preparation request remains suspended. If the abort of the old response succeeds, the newer response preparation is resumed and eventually executed.

Although Jacoby's (1991) model requires that recollection either contribute accurate source information to the response decision or contribute no information at all, the PTS model assumes that recollection can sometimes retrieve incorrect source information. Several studies have reported evidence for incorrect source attributions (called *misrecol-*

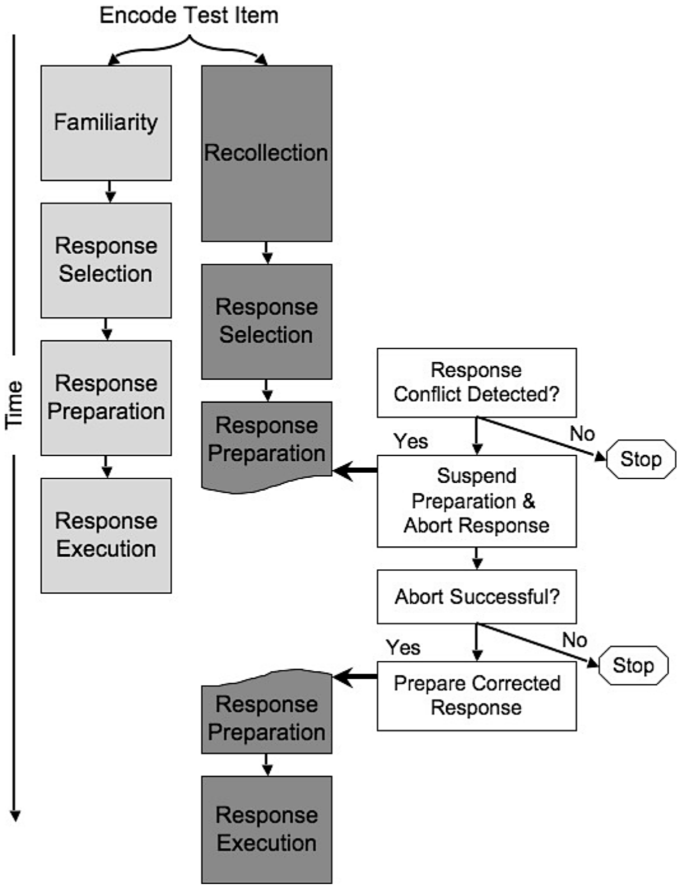

Figure 2. Diagram of the parallel task set model (Seymour, 2001). Light gray boxes indicate the processing stages within the familiarity task set. Dark gray boxes depict processing stages in the recollection task set. White boxes represent task-invariant control processes that monitor and attempt to overcome response conflict.

lections) in strategic recognition tasks; this is particularly likely when target and probe study contexts are similar (Dodson, 2007; Dodson, Bawa, & Slotnick, 2007; Dodson, Holland, & Shimamura, 1998; Dodson & Johnson, 1996; Simons, Dodson, Bell, & Schacter, 2004). For example, Dodson and Johnson used an exclude recognition procedure in which the similarity of the target and probe study contexts was manipulated and found that participants sometimes misrecollected probes as targets. False alarms (responding "old" to probes) were more prevalent when study contexts were similar (probes encountered in a word fragment task and targets encountered in an anagram-solving task) than when they were dissimilar (probes merely heard and targets encountered in an anagram-solving task).

The PTS model incorporates such source attribution errors and uses the potential for the abort command to fail, along with occasional misrecollections in which probe items are mistakenly recognized as targets (or vice versa), to account for exclude recognition accuracy effects. The suspend, abort, and resume processes will result in these responses being slower on correct probe trials than on filler trials on which no conflict is predicted. These delays give rise to the RT effects in the exclude recognition task.

The PTS model has been used to successfully simulate behavior in the exclude recognition task; Figure 1 shows both empirical (solid bars) and simulated (patterned bars) data from one such effort. This simulation shows that the model is sufficient to account for human RT and accuracy data with only the durations of the familiarity and recollection processes as free parameters. Indeed, the average root-mean squared error of the fit was 23 msec ($R^2 = .99$) for RT and 2% ($R^2 = .97$) for accuracy. For further details on the PTS model's fit of both exclude and include recognition data, see Seymour (2001).

Measuring Conflict in the Motor System

Despite the successful model fit, the PTS model's assumption that exclude recognition procedures involve response conflict in the motor system has yet to be tested. Thus, the goal of the present study was to precisely test for the existence of motor conflict posited by the PTS model in the exclude recognition task. To achieve this, we used electromyography (EMG) to measure motor neuron recruitment in the arms—a dependent measure clearly related to motor processing. This is particularly important because conflict could occur at multiple stages (e.g., source/context resolution, response selection, or response execution) in the exclude recognition paradigm.

Previous research in which EMG has been examined in response conflict tasks has used the presence of subthreshold muscle activity associated with the incorrect responses (i.e., *partial errors*) prior to suprathreshold activity associated with the correct response as evidence for response override type dynamics (e.g., Burle, Possamaï, Vidal, Bonnet, & Hasbroucq, 2002; Burle, Vidal, Tandonnet, & Hasbroucq, 2004; Coles, Gratton, Bashore, Eriksen, & Donchin, 1985; Coles, Scheffers, & Fournier, 1995). These EMG patterns, averaged across trials, typically overlap (e.g., Burle, Allain, Vidal, & Hasbroucq, 2005; Burle et al., 2002; Carbonnell & Falkenstein, 2006). This pattern is used as evidence for the coactivation of competing responses predicted by conflict-monitoring theory (Botvinick et al., 2001; Burle, Roger, Allain, Vidal, & Hasbroucq, 2008).

The PTS model predicts that we should observe a similar pattern of EMG partial errors on correct probe trials. However the specific prediction that conflict occurs at the response preparation stage, after response selection, suggests a nonoverlapping pattern—that is, fleeting manual motor activity associated with preparation of the familiarity-based "old" response, followed by subsequent, temporally distinct activity associated with the recollection-based "new" response. Although less likely, partial errors are also possible on correct target trials ("new" then "old") that involve response conflict (viz., those on which below threshold familiarity leads to an initial "new" response, followed by a recollection-based "old" response).

We predicted that partial errors would occur most often on probe trials, followed by target trials, and would occur rarely on filler trials. Therefore, finding either no partial errors or no difference in partial errors by stimulus type would undermine the assumptions of the PTS model and would undermine our suggestion that exclude recognition tasks involve response conflict.

METHOD

Participants

Twenty-three right-handed volunteers (ages, 18–23; 13 of them female) participated in this experiment. All the participants were recruited from the University of California Santa Cruz community and gave their informed consent.

Behavioral Procedure

The experiment consisted of three phases and a distractor task (shown in Figure 3). In the first two phases, the participants studied separate sets of 6 two-word phrases. These study phases were separated by 10 min, during which the participants performed a distractor math task. The phrases from the first study phase would later serve

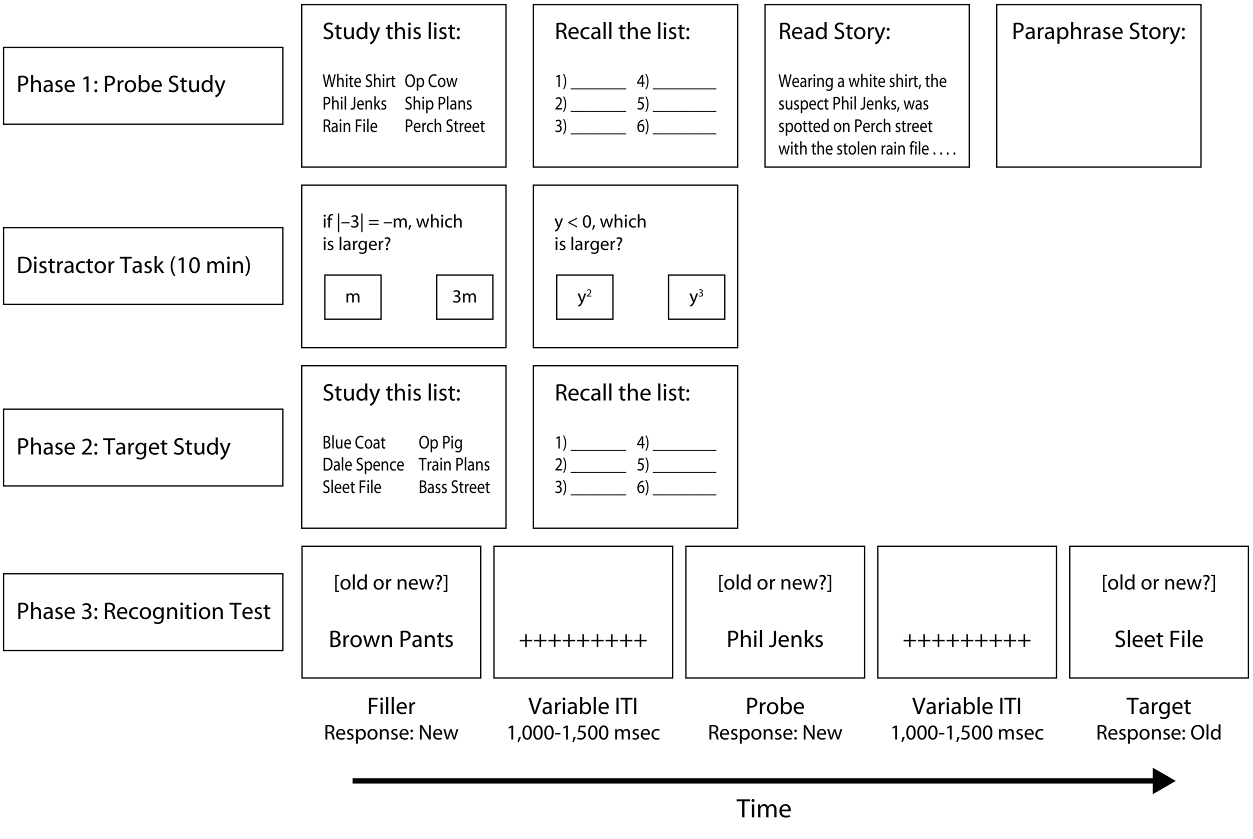

Figure 3. Overview of the experimental paradigm. During probe study, six phrases were studied and recalled three times and then viewed as part of a mock newspaper story that had to be paraphrased later. After a 10-min distractor task in which participants answered a series of math questions, six target phrases were memorized and recalled three times. The final recognition task displayed a random series of phrases including probes, targets, and new fillers. Participants responded "old" to targets and "new" to both novel fillers and familiar probes within a 1,500-msec response deadline.

as probe items. Phrases from the second phase would later serve as target items. This categorization of stimuli was not revealed to the participants, who were simply asked to study one set of phrases and then the other. In addition to probe and target items, 24 unstudied filler phrases were randomly selected for each participant, for a total stimulus set of 36 phrases. In general, this procedure was similar to previous exclude recognition paradigms using verbal stimuli (Jacoby, 1991; Jacoby, Kelley, et al., 1989; Jacoby, Woloshyn, & Kelley, 1989; Seymour & Kerlin, 2008; Seymour et al., 2000). In particular, we replicated the procedure previously reported in Seymour et al. but added the EMG measure.

The stimulus set consisted of 36 two-word verbal phrases sampled from a set of 72 phrases taken from Farwell and Donchin (1991). For each participant, 6 probe phrases (e.g., "Phil Jenks," "Blue Coat," "Op Cow," "Rain File," "Perch Street," and "Brass Plans") and 6 target phrases (e.g., "Wayne Bryant," "White Shirt," "Op Pig," "Sleet File," "Shark Street," and "Steel Plans") were randomly sampled from the full set, along with 24 corresponding filler phrases.

Each participant (1) studied the six probe items, (2) completed a 10-min distractor task (consisting of SAT-type math problems), (3) studied six target items, and (4) completed the exclude recognition task. During both the probe and target study tasks, the participants were shown all six phrases and were asked to commit them to memory. When ready to proceed, the participants pressed the space bar and were then asked to recall the list in the order previously presented. They were given accuracy feedback following each recalled phrase, and the entire study–recall sequence was repeated three times. This was sufficient for most of the participants to correctly recall at least 5/6 (83%) of the probe and target phrases. To help differentiate the probe and target lists in memory, following the probe study task we asked the participants to read and paraphrase a mock newspaper story in which their six probe phrases had been embedded. For example (embedded critical phrases have been underlined),

> Campus police reported confronting a student attempting to purchase an altered copy of his academic grade report. Last night he apparently dialed up the university computer system from his home, logged in using the alias blue coat, and entered the password brass plans to gain access to his email account. The student then proceeded to write an email to Phil Jenks who is suspected of running a grade-change operation code named op cow. In the email, the student scheduled a meeting at a San Jose establishment called the Shark Street Café and asked the perpetrator to bring his rain file (which apparently indicates a doctored copy of his grade report). Police staked out the scene and questioned both the parties.

The participants were given up to 5 min to paraphrase the story, and no accuracy feedback was given.

Following the 10 min distractor task, the participants completed the target study phase. Target study was identical to probe study, except that no story paraphrase was used. Thus, whereas the story paraphrase task was used for the probe study phase, no additional task was used during the target study phase. After studying the target items, the participants began the exclude recognition task. First, they were shown how to use the response apparatus to respond either "old" or "new." For this task, we created a special response device that would immediately measure the participant's RT to each stimulus, as well as facilitate the measurement of EMG signals from the left and right triceps. The participants were asked to sit so that their arms hung straight at their sides with their elbows bent at a 90º angle. The participants gripped two vertical cylinders on the table in front of them and pressed downward toward the surface to make each response. Electric switches at the base of each cylinder allowed the measurement of this response, and electrodes attached to the triceps on each arm were used to measure EMG signals concomitant with each response. Because elbows and forearms were supported by the chair's armrest and the table (isometric isolation of the triceps), a moderate downward force was required to make each response. After the participants had been connected to the physiological recorder and moved into position, they received several practice trials on each hand in which they were asked to quickly produce either the right ("old") or the left ("new") response.

Following practice, the participants were given instructions for the exclude recognition task. On each trial, the participants saw a two-word phrase and made a speeded old/new judgment. They were asked to respond "old" (with the right cylinder) if the phrase had been studied during the second study phase (target phrases, referred to as "one of the phrases you *just* studied"). Otherwise, they were to respond "new" (with the left cylinder). This included not only new filler items, but also familiar phrases studied in the first study phase (probe phrases). The participants were encouraged to respond both quickly and accurately. This task consisted of three blocks of 36 trials each. On each trial, a fixation cross was displayed for 1,000 msec, followed by a mask (++++++++++++) presented for 1,000 msec. Next, a randomly selected stimulus from the 6 probe, 6 target, or 24 filler phrases was presented in the center of the display until one of the two response buttons was depressed. Responses were to be made before a 1,500-msec deadline had elapsed. If any response exceeded the deadline, the message TOO SLOW was displayed for 800 msec. The intertrial interval was randomly varied between 1,000 and 1,500 msec. The presentation and randomization of stimuli, as well as the recording of manual response data (RT and accuracy), were handled by E-Prime experimental presentation software (Schneider, Eschman, & Zuccolotto, 2002) running on a desktop computer.

Psychophysiological Procedure and Analysis

EMG was recorded from electrodes placed on the skin covering the medial head of the triceps brachii of each arm. A ground electrode was placed over the skin just adjacent to the palmar carpal ligament of the left or right wrist (counterbalanced). EMG was recorded with an ADInstruments physiological recorder and was amplified using an ADInstruments dual-channel biopotential amplifier. EMG was recorded from -5 mV to $+5$ mV and was digitized at 1000 Hz. The data were also high-pass filtered with a 50-Hz cutoff, were rectified, and were low-pass filtered with a 20-Hz cutoff (Van Boxtel, Geraats, Van den Berg-Lenssen, & Brunia, 1993). In addition, the data were smoothed using a 6-Hz low-pass filter to aid in visual identification of partial-error onsets. Figure 4 shows representative example EMG waveforms for non-partial-error probe (Figure 4A), filler (Figure 4B), and target (Figure 4C) trials. These waveforms depict EMG amplitude in microvolts and are plotted relative to stimulus onset.

To test the prediction that the exclude recognition task involves response conflict, we tallied the number of partial errors for each trial type. Because very few errors were made on target and filler trials (fewer than 7% overall), we restricted our analysis to trials on which correct responses were made.

Following the filtering and data conditioning, an EMG baseline was taken during the fixation period of each trial. Significant subthreshold deviation from this baseline ($\geq$5% of maximum amplitude) occurring on one arm, followed by suprathreshold deviation from baseline occurring on the other arm, was coded as a partial error. Figure 5A shows a representative example of a probe trial during which an old–new partial error was recorded and is plotted relative to partial-error onset. An initial "old" response (dashed waveform) is initiated, inhibited (note the lack of overlap in the waveforms), and subsequently replaced by a "new" response (solid waveform), using the other arm. Old–new partial errors for correct probe and filler trials and new–old partial errors on correct target trials (e.g., Figure 5B) accounted for over 99% of the partial errors in the data; therefore, we focused on these partial errors exclusively (Figure 5C shows an average of all correct probe and target partial errors). We also ignored trials on which concurrent bilateral EMG signals occurred (fewer than 1% of trials on which bilateral activation was observed). Because these trials featured concomitant peaks, they were believed to reflect bilateral movement artifacts. Figure 5D shows an example bilateral activation pattern; note that the morphologies of these distributions are similar and do not appear to reflect the partially overlapping EMG pattern evident in Figures 5A and 5B. Finally, because there were more filler (72) than probe or target (18

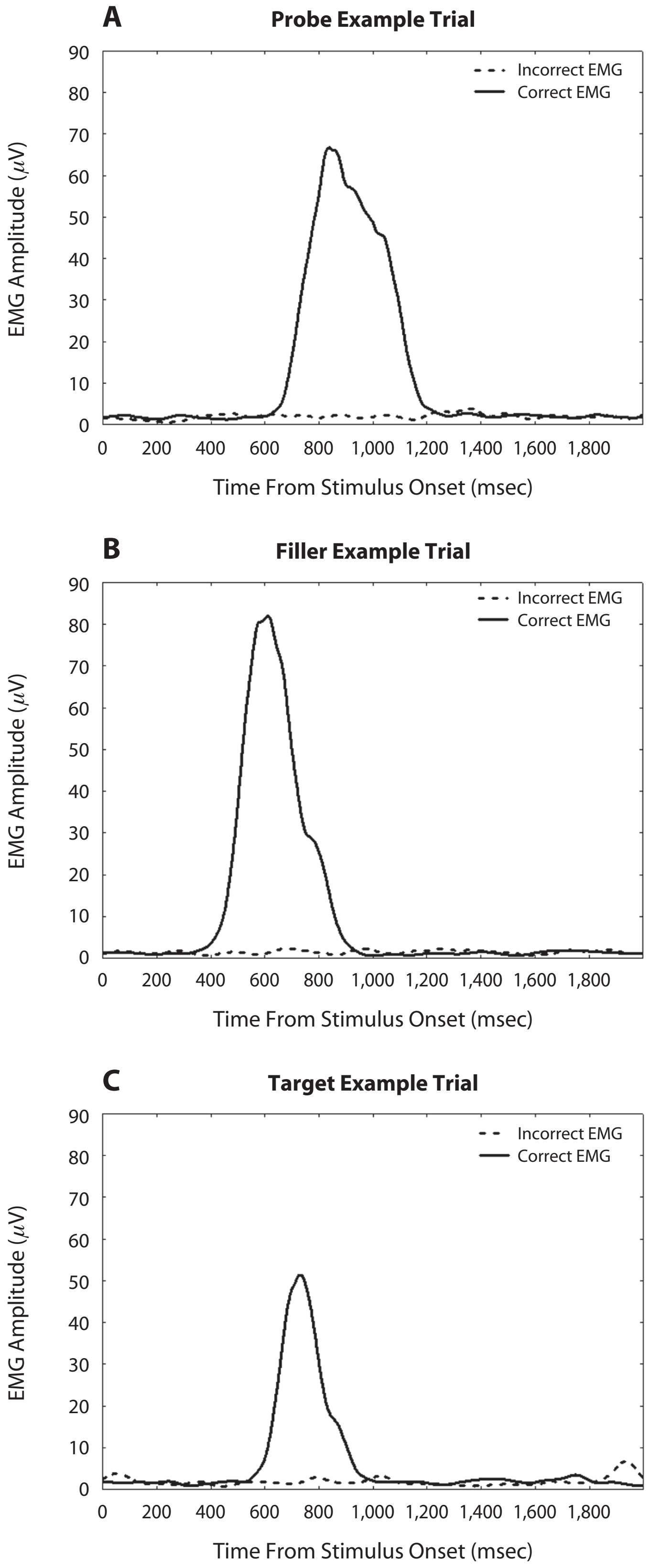

Figure 4. Filtered and rectified electromyograph (EMG) data (non-partial-error trials): suprathreshold EMG waveforms relative to stimulus onset for representative correct probe (A), filler (B), and target (C) trials.

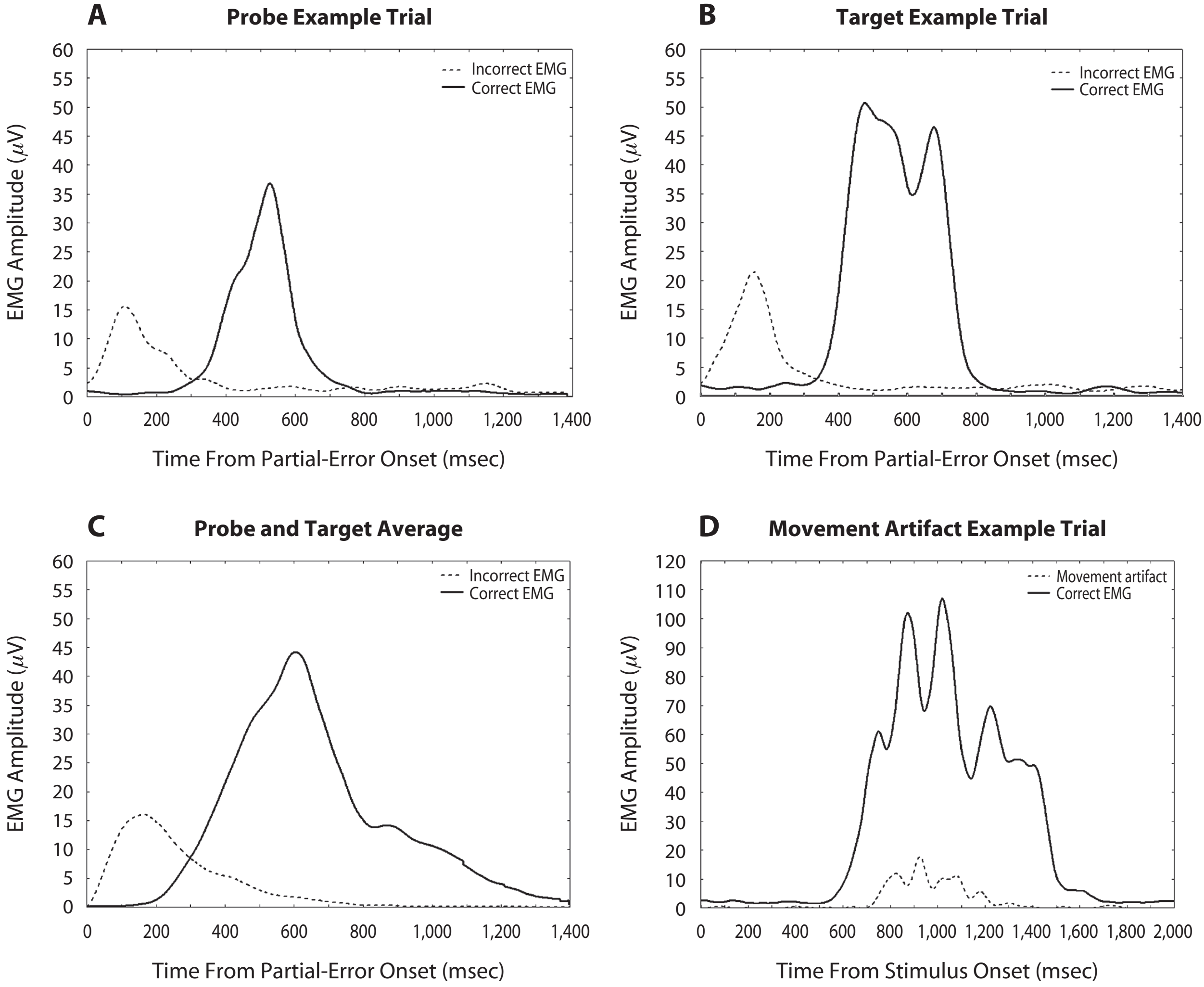

Figure 5. Filtered and rectified electromyograph (EMG) data (partial-error trials). Dashed lines represent subthreshold responses, and solid lines represent suprathreshold responses. Panels A–C show EMG waveforms relative to partial-error onset for representative individual partial errors on a probe (A) and a target (B) trial. Panel C is an average of all correct probe (old–new) and target (new–old) partial errors. Panel D shows an example of concomitant bilateral activation patterns ($<$1% of all bilateral EMGs) excluded from the analysis.

each) trials, we converted number of partial errors into percentage of partial errors by trial type for the analysis.

RESULTS

Mean RT and accuracy data (shown in Figure 6) from the exclude recognition task were analyzed with separate within-subjects ANOVAs with stimulus type (target, probe, or filler) as a factor. Stimulus type had a significant effect on both correct RT [$F(2,44) = 43.38, p < .001$] and accuracy [$F(2,44) = 26.94, p < .001$]. As in previous reports using similar procedures (Jacoby, Woloshyn, & Kelley, 1989; Seymour & Kerlin, 2008; Seymour et al., 2000), probe responses were slower and less accurate on average than both target and filler responses. Figure 6 and the pairwise *t* tests shown in Table 1 reveal that probe responses were significantly slower than both target and filler responses. In addition, target RTs were longer than filler RTs. Figure 6 and Table 1 also show that all pairwise *t* tests for the accuracy data were statistically significant. That is, accuracy

Table 1
Statistics for Planned Comparisons

	Probe vs. Filler	Target vs. Filler	Probe vs. Target
Reaction time	$t(22) = 9.46, p < .001^*$	$t(22) = 6.49, p < .001^*$	$t(22) = 3.09, p < .01^*$
Accuracy	$t(22) = -5.93, p < .001^*$	$t(22) = -4.41, p < .001^*$	$t(22) = -4.39, p < .001^*$
Percentage of partial errors	$t(22) = 4.12, p < .001^*$	$t(22) = 4.39, p < .001^*$	$t(22) = 2.84, p < .01^*$

Note—All *t* tests are one-tailed except those comparing probe versus target, which are compared using two-tailed tests. Asterisks indicate statistically significant results.

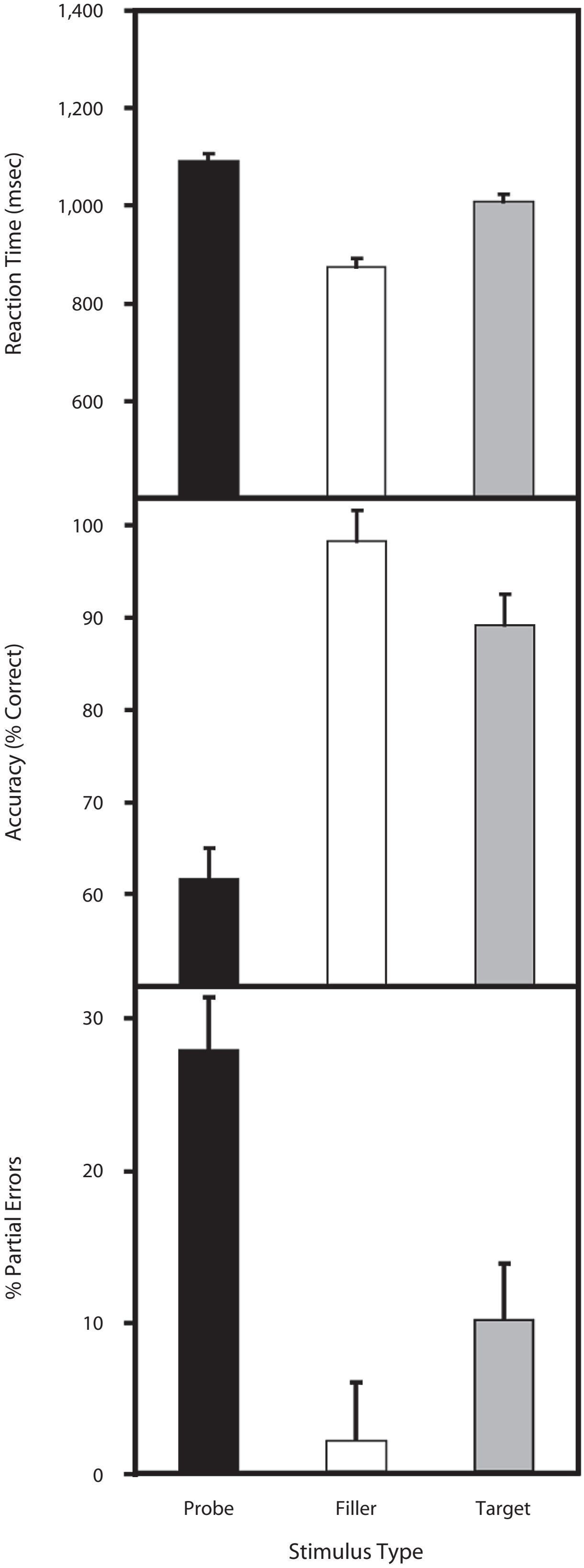

Figure 6. Mean reaction time, accuracy, and percentage of partial errors for probe, filler, and target trials. Error bars represent the standard errors of the means.

was lower for probe than for target and filler trials, and filler responses were more accurate than responses on target trials. In general, accuracy for target and filler trials were near ceiling, whereas the participants found it difficult to correctly reject familiar probe items as “new.” Thus, these results are typical for the exclude recognition paradigm.

A similar analysis was performed on the partial-error data. Figure 6 (bottom graph) shows mean percentage of partial errors for correct trials of each stimulus type. Supporting the response conflict prediction, an ANOVA on these data revealed a significant main effect of stimulus type [$F(2,44) = 12.81, p < .001$]. The largest mean proportion of partial errors occurred on probe trials followed by target trials and filler trials. Significant pairwise *t* tests, shown in Table 1, confirmed this pattern. Although the PTS model does not give a specific prediction of the proportion of conflict trials on which partial errors should be observed, we note that the 28% partial errors observed on probe trials is similar to the proportions of partial errors found on conflict trials in more canonical conflict tasks. For example, a recent study using an Eriksen flanker task reported 21.7% partial errors for conflict trials (Burle et al., 2008). However, due to differences between studies, a direct comparison between these proportions is not possible.

Individual Participant Analysis

The increase in partial errors on probe trials shown in the mean proportion data is reinforced by a comparison of the number of partial errors across the stimulus types within each participant. More participants produced partial errors on probe than on target or filler trials. In fact, only 2 (9%) participants failed to produce any partial errors on probe trials, as compared with 6 (26%) on target trials and 9 (39%) on filler trials. This comparison was significant for probe versus filler partial errors [$Z(22) = 2.42, p < .01$] and approached significance for the probe versus target comparison [$Z(22) = 1.56, p = .06$]. This suggests that, in addition to the differences reported in mean proportion of partial errors, each individual participant was significantly more likely to produce partial errors on probe trials than on target or filler trials.

Block Analysis

To investigate whether the stimulus type differences reported here changed over the course of the experiment, we replicated the analyses above, with block (1–3) as a factor. As in the previous analyses, significant main effects of stimulus type were found for RT [$F(2,42) = 28.07, p < .001$], accuracy [$F(2,42) = 30.30, p < .001$], and percentage of partial errors [$F(2,42) = 15.52, p < .001$]. However, no main effects or interactions involving block were statistically significant ($p > .1$ in all cases). Even after three repetitions of the stimulus set, probe responses were 145 msec slower and 33% less accurate, and led to 15% more partial errors than did filler responses. Thus, consistent with previous studies of response conflict, these effects appear to be quite robust across blocks (for a review, see MacLeod, 1991). They are also consistent with the results of a previous study using the present paradigm with RT, accuracy, and electrodermal response measures; robust stimulus type effects were found for all measures, but no block effects were observed (Seymour, 2009).

DISCUSSION

The present behavioral results substantially replicate previous published data (Jacoby, Kelley, et al., 1989; Jacoby, Woloshyn, & Kelley, 1989; Seymour & Kerlin, 2008; Seymour et al., 2000). Namely, filler responses were faster and more accurate than target and probe responses, and probe responses were significantly slower and less accurate than target responses, although the differences between these responses were smaller than when either stimulus type was compared with fillers.

Figure 4 depicts examples of trials during which a single suprathreshold EMG was associated with a correct manual response. Figure 5 shows examples of trials during which a suprathreshold EMG associated with the correct response were preceded briefly by a subthreshold EMG associated with the incorrect response. The critical measure in this study was the proportion of these *partial-error* responses. The results revealed that partial errors were significantly more likely on probe than on filler trials. This result is at least qualitatively similar to partial response data reported for conflict trials during more canonical response conflict tasks (viz., a flanker task, as in Burle et al., 2008). We also found significantly more partial errors on probe than on target trials, supporting a specific prediction made by the PTS model for the exclude recognition procedure. This pattern was highly consistent across participants; all but 2 produced more partial errors on probe than on target trials.

The present accuracy results can be explained by most dual-process theories of memory retrieval, which tend to account for such data as a function of the relative contributions of familiarity and recollective processes (for a review, see Yonelinas, 2002). The PTS model (Seymour, 2001) approaches the accuracy data in a similar fashion but assumes that both familiarity and recollection processes may lead to different response activations. This produces response conflict in the motor system, which must be resolved by control processes. The goal of the present study was to test whether such response conflict actually occurs in this task by using a dependent measure that clearly reflects response programming. Therefore, we used surface EMG to measure the proportions of partial errors (subthreshold response initiations followed by suprathreshold response activity) in the exclude recognition task. Our results revealed a significantly greater proportion of partial errors on probe and target trials than on filler trials. The EMG results clearly reflect conflicting response processing and, as such, offer a direct test for the existence of response conflict in the exclude recognition task. In particular, the present data are consistent with the PTS model's assumption that familiarity-based responses are sometimes explicitly initiated and, therefore, must be aborted in lieu of a corrective (although not necessarily accurate) recollection-based response.

Theoretical Implications of the Present Results

These data limit alternative theories that might be used to model RTs in the exclude recognition task. For example, if we assume the typical dual-process framework involving familiarity and recollection common to many recognition models (for a review, see Yonelinas, 2002), there are at least three possible ways to model response processes on correct probe trials: (1) On the basis of the combined (nonconflicting) consideration of familiarity and recollection information, a single response is selected and prepared; (2) both the recollection- and the familiarity-based responses become simultaneously activated (conflict) in the motor cortex, and following response selection, a single recollection-based response is initiated; or (3) both response mappings are activated (conflict), but the familiarity-based response is first initiated, subsequently inhibited, and replaced by a recollection-based response initiation. Models based on Option 1 would have trouble explaining the presence of any partial errors. A model based on Option 2 may predict biphasic motor activity if familiarity is faster than recollection and motor activation accrues gradually during the course of these processes. However, because both recollection- and familiarity-based responses would be simultaneously activated, this approach may have trouble accounting for the nonoverlapping motor activations found here (see, e.g., Figures 5A and 5B). Finally, Option 3, which is consistent with the PTS model, predicts nonoverlapping partial errors.

Although the present results are clearly related to response conflict and, thus, support our main hypothesis, some implications of these data remain unclear. For example, the averaged EMG data (Figure 5C) shows relatively nonoverlapping partial-error patterns, and the individual trial partial errors reveal an even greater degree of nonoverlap (see Figures 5A and 5B). However, conflict-monitoring theory (consistent with our Option 2 account above) predicts that trials with increased conflict will lead to increased partial-error overlap. Indeed, overlapping partial-error patterns are typically taken as support for that model. If exclude recognition involves response conflict, as we claim, why do we find nonoverlapping partial errors when others report ones that overlap (e.g., Burle et al., 2005; Burle et al., 2002; Carbonnell & Falkenstein, 2006; Coles et al., 1985)? One answer can be found in a recent report by Burle et al. (2008). Burle and colleagues present a detailed and thorough examination of the brain activation and EMG implications of conflict-monitoring theory and argue against the Botvinick et al. (2001) definition of response conflict as overlapping activation of competing response mappings. For example, they demonstrate that if one starts with a distribution of nonoverlapping partial errors (which would indicate a lack of conflict in conflict-monitoring theory) observed on individual trials, averaging these partial-error functions together to produce a single averaged graph leads to an artifactual degree of partial-error overlap (indicating conflict as defined by conflict-monitoring theory). This occurs because in order to create the averaged graph, individual EMG waveforms must be locked to either a trial event (e.g., the stimulus onset) or the partial-error onset, with relatively fast trials averaged together with relatively slow trials. This effect can be seen in our Figure 5C (an average) showing a greater degree of partial-error overlap than does either of the individual partial errors shown in Figures 5A and 5B.

Because our averaged data (Figure 5C) are similar to averaged data reported in studies using more canonical response conflict paradigms, it is possible that models used to account for those data also apply to our data. In particular, our EMG data may reflect gradual motor neuron recruitment associated with overlapping activations in the motor cortex (as conflict-monitoring theory implies) and not the motor initiation pattern proposed by the PTS model. However, Burle et al. (2008) also suggest that the pattern of nonoverlapping partial errors shown here (and presumably underlying previously reported averaged data) is not associated with evidence of bilateral response activations in the motor cortex during conflict tasks. Even without this evidence, the possibility that overlapping partial-error waveforms can be generated when nonoverlapping partial errors are averaged poses a problem for conflict-monitoring theory studies that rely on this overlap as an indication of conflict. Burle and colleagues propose that this theory's definition of conflict as overlapping response activations is problematic. The PTS model offers an answer to this puzzle by predicting that, rather than responses being simultaneously activated, one response is initiated and then aborted in favor of the final response. This predicts the nonoverlapping partial-error results presented here.

Although additional recent studies also call into question some of the assumptions underlying the conflict-monitoring theory (e.g., Burle et al., 2005; Carbonnell & Falkenstein, 2006), it is premature to discount a theory that so successfully accounts for such a substantial body of data. Despite Burle et al.'s (2008) report, it is possible that conflict occurs at multiple information-processing levels simultaneously and may manifest to different degrees at early (e.g., stimulus encoding), central (e.g., response selection, task representation), and response (e.g., motor programming) levels, depending on the task situation. Clearly, more research is needed to address this issue.

It is also possible that any differences between partial-error overlap in the present study and that reported using Stroop or flanker tasks may be influenced by the mnemonic processes involved in the exclude recognition task. For example, in the Stroop and flanker tasks, the two responses are mapped to relatively superficial aspects of the stimulus. Therefore, activation of the two responses can occur immediately and in parallel following stimulus onset and can lead to more overlapping partial errors. In the exclude recognition task, the response depends on the familiarity and source information not evident in the superficial stimulus; thus, episodic memory must be consulted before response selection can occur. Furthermore, these processes resolve at different times—with familiarity usually preceding recollection (see Yonelinas, 2002). This staggered availability of information may lead to more sequential response activations in the exclude recognition than in the Stroop or flanker task.

The PTS model suggests that differences between probe, target, and filler responses (whether measured using RT, accuracy, or EMG) in this procedure are caused by differences in the relative likelihood of response conflict on each trial type, with misrecollections playing a relatively minor role. However, one alternative explanation is that the effect on all measures is driven by source confusion. Yet, recent evidence suggests that source confusion is unlikely to cause these effects. Using an identical procedure (except that EMG was not measured), Seymour and Kerlin (2008) showed that participants were near ceiling on both nonspeeded recall and recognition posttests, missing less than one item per list on average. Because source confusion was not evident but the RT and accuracy results were similar to the ones reported here, it seems unlikely that source confusion played a dominant role in the present results.

Our present behavioral and electrophysiological results suggest that the exclude recognition task involves response conflict. One hallmark of canonical response conflict procedures is the activation of brain regions sensitive to response conflict (e.g., Carter et al., 1998; Gehring et al., 1993). Activation of medial frontal regions (e.g., the ACC) has been consistently associated with situations high in response conflict. Indeed, such activations have been used as evidence for the importance of this brain region in the conflict-monitoring theory (e.g., Botvinick et al., 2001; Botvinick et al., 2004; Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999). Therefore, if the exclude recognition task involves response conflict, future research should show a similar pattern of activation on probe, but not filler, trials. In fact, recent neuroimaging evidence from our laboratory largely confirms this prediction (Schumacher, Seymour, & Schwarb, 2009).

Conclusion

The data reported here show that strategic recognition procedures such as the exclude recognition task involve response conflict. As such, typical models that account for accuracy performance in these tasks by focusing on memory processes only (see Yonelinas, 2002) may need to adopt components of models typically used to account for canonical response conflict tasks (Botvinick et al., 2001; Botvinick et al., 2004; MacLeod, 1991) in order to account for both accuracy and RT. One attempt is the PTS model (Seymour, 2001), which suggests that performance on exclude recognition tasks (and presumably, other conflict tasks as well) is best explained as an integration between memory, response, and additional control processes responsible for mediating conflict. We show that the assumptions of this model can be tested using behavioral data such as RT and accuracy, as well as physiological measures such as EMG. These data provide substantial support for the PTS model's assumptions and predictions. However, more work is needed to distinguish it from other integrative models of response conflict.

AUTHOR NOTE

Address correspondence to T. L. Seymour, Psychology Department, University of California, Santa Cruz, CA 95064 (e-mail: nogard@ucsc.edu) or E. H. Schumacher, School of Psychology, Georgia Institute of Technology, Atlanta, GA 30332 (e-mail: eschu@gatech.edu).

REFERENCES

Atkinson, R. C., & Juola, J. F. (1974). Search and decision processes in recognition memory. In D. H. Krantz, R. C. Atkinson, & R. D.

Lucke (Eds.), *Contemporary developments in mathematical psychology: Vol. 1. Learning, memory, and thinking* (pp. 243-293). San Francisco: Freeman.

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, **108**, 624-652.

Botvinick, M. M., Cohen, J. D., & Carter, C. S. (2004). Conflict monitoring and anterior cingulate cortex: An update. *Trends in Cognitive Sciences*, **8**, 539-546.

Botvinick, M. M., Nystrom, L. E., Fissell, K., Carter, C. S., & Cohen, J. D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, **402**, 179-181.

Burle, B., Allain, S., Vidal, F., & Hasbroucq, T. (2005). Sequential compatibility effects and cognitive control: Does conflict really matter? *Journal of Experimental Psychology: Human Perception & Performance*, **31**, 831-837.

Burle, B., Possamaï, C.-A., Vidal, F., Bonnet, M., & Hasbroucq, T. (2002). Executive control in the Simon effect: An electromyographic and distributional analysis. *Psychological Research*, **66**, 324-336.

Burle, B., Roger, C., Allain, S., Vidal, F., & Hasbroucq, T. (2008). Error negativity does not reflect conflict: A reappraisal of conflict monitoring and anterior cingulate cortex activity. *Journal of Cognitive Neuroscience*, **20**, 1637-1655.

Burle, B., Vidal, F., Tandonnet, C., & Hasbroucq, T. (2004). Psychological evidence for response inhibition in choice reaction time task. *Brain & Cognition*, **56**, 141-152.

Carbonnell, L., & Falkenstein, M. (2006). Does the error negativity reflect the degree of response conflict? *Brain Research*, **1095**, 124-130.

Carter, C. S., Braver, T. S., Barch, D. M., Botvinick, M. M., Noll, D., & Cohen, J. D. (1998). Anterior cingulate cortex, error detection, and the online monitoring of performance. *Science*, **280**, 747-749.

Cohen, J. D., Dunbar, K., & McClelland, J. L. (1990). On the control of automatic processes: A parallel distributed processing account of the Stroop effect. *Psychological Review*, **97**, 332-361.

Coles, M. G. H., Gratton, G., Bashore, T. R., Eriksen, C. W., & Donchin, E. (1985). A psychophysiological investigation of the continuous flow model of human information processing. *Journal of Experimental Psychology: Human Perception & Performance*, **11**, 529-553.

Coles, M. G. H., Scheffers, M. K., & Fournier, L. (1995). Where did you go wrong? Errors, partial errors, and the nature of human information processing. *Acta Psychologica*, **90**, 129-144.

Dodson, C. S. (2007). Retrieval-based illusory recollections: Why study–test contextual changes impair source memory. *Memory & Cognition*, **35**, 1211-1221.

Dodson, C. S., Bawa, S., & Slotnick, S. D. (2007). Aging, source memory, and misrecollections. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **33**, 169-181.

Dodson, C. S., Holland, P. W., & Shimamura, A. P. (1998). On the recollection of specific- and partial-source information. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **24**, 1121-1136.

Dodson, C. S., & Johnson, M. K. (1996). Some problems with the process-dissociation approach to memory. *Journal of Experimental Psychology: General*, **125**, 181-194.

Dosher, B. A. (1984). Discriminating preexperimental (semantic) from learned (episodic) associations: A speed–accuracy study. *Cognitive Psychology*, **16**, 519-555.

Farwell, L. A., & Donchin, E. (1991). The truth will out: Interrogative polygraphy ("lie detection") with event-related brain potentials. *Psychophysiology*, **28**, 531-547.

Gehring, W. J., Goss, B., Coles, M. G. H., Meyer, D. E., & Donchin, E. (1993). A neural system for error detection and compensation. *Psychological Science*, **4**, 385-390.

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, **91**, 1-67.

Gronlund, S. D., & Ratcliff, R. (1989). Time course of item and associative information: Implications for global memory models. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **15**, 846-858.

Hintzman, D. L., Caulton, D. A., & Levitin, D. J. (1998). Retrieval dynamics in recognition and list discrimination: Further evidence of separate processes of familiarity and recall. *Memory & Cognition*, **26**, 449-462.

Hintzman, D. L., & Curran, T. (1994). Retrieval dynamics of recognition and frequency judgments: Evidence for separate processes of familiarity and recall. *Journal of Memory & Language*, **33**, 1-18.

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory & Language*, **30**, 513-541.

Jacoby, L. L., Kelley, C., Brown, J., & Jasechko, J. (1989). Becoming famous overnight: Limits on the ability to avoid unconscious influences of the past. *Journal of Personality & Social Psychology*, **56**, 326-338.

Jacoby, L. L., Toth, J. P., & Yonelinas, A. P. (1993). Separating conscious and unconscious influences of memory: Measuring recollection. *Journal of Experimental Psychology: General*, **122**, 139-154.

Jacoby, L. L., Woloshyn, V., & Kelley, C. (1989). Becoming famous without being recognized: Unconscious influences of memory produced by dividing attention. *Journal of Experimental Psychology: General*, **118**, 115-125.

Kieras, D. E., Meyer, D. E., Mueller, S. T., & Seymour, T. L. (1999). Insights into working memory from the perspective of the EPIC architecture for modeling skilled perceptual–motor and cognitive human performance. In A. Miyake & P. Shah (Eds.) *Models of working memory: Mechanisms of active maintenance and executive control* (pp. 183-223). New York: Cambridge University Press.

Logan, G. D. (1994). On the ability to inhibit thought and action: A users' guide to the stop signal paradigm. In D. Dagenbach & T. H. Carr (Eds.) *Inhibitory processes in attention, memory, and language* (pp. 189-239). San Diego: Academic Press.

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, **109**, 163-203.

McElree, B., Dolan, P. O., & Jacoby, L. L. (1999). Isolating the contributions of familiarity and source information to item recognition: A time course analysis. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **25**, 563-582.

Meyer, D. E., & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: I. Basic mechanisms. *Psychological Review*, **104**, 3-65.

Monsell, S. (1996). Control of mental processes. In V. Bruce (Ed.), *Unsolved mysteries of the mind: Tutorial essays in cognition* (pp. 93-148). Hove, U.K.: Erlbaum, Taylor & Francis.

Ratcliff, R., & McKoon, G. (1989). Similarity information versus relational information: Differences in the time course of retrieval. *Cognitive Psychology*, **21**, 139-155.

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime user's guide (Version 1.1)*. Pittsburgh: Psychology Software Tools.

Schumacher, E. H., Seymour, T. L., & Schwarb, H. (2009). *Brain activation evidence for response conflict in the exclude recognition task*. Manuscript submitted for publication.

Seymour, T. L. (2001). A EPIC model of the "guilty knowledge effect": Strategic and automatic processes in recognition. *Dissertation Abstracts International: Section B. The Sciences & Engineering*, **61**, 5591.

Seymour, T. L. (2009). *Efficiency of RT versus skin conductance in the detection of concealed knowledge*. Manuscript under review.

Seymour, T. L., & Kerlin, J. R. (2008). Successful detection of verbal and visual concealed knowledge using an RT-based paradigm. *Applied Cognitive Psychology*, **22**, 475-490.

Seymour, T. L., Seifert, C. M., Shafto, M. G., & Mosmann, A. L. (2000). Using response time measures to assess "guilty knowledge." *Journal of Applied Psychology*, **85**, 30-37.

Simons, J. S., Dodson, C. S., Bell, D., & Schacter, D. L. (2004). Specific- and partial-source memory: Effects of aging. *Psychology & Aging*, **19**, 689-694.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, **18**, 643-662.

Van Boxtel, G. J., Geraats, L. H., Van den Berg-Lenssen, M. M., & Brunia, C. H. (1993). Detection of EMG onset in ERP research. *Psychophysiology*, **30**, 405-412.

Yonelinas, A. P. (2002). The nature of recollection and familiarity: A review of 30 years of research. *Journal of Memory & Language*, **46**, 441-517.

Yonelinas, A. P., & Jacoby, L. L. (1994). Dissociations of processes in recognition memory: Effects of interference and of response speed. *Canadian Journal of Experimental Psychology*, **48**, 516-534.

(Manuscript received December 12, 2007;
revision accepted for publication July 12, 2008.)