THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY
2012, 65 (1), 92–120

Psychology Press
Taylor & Francis Group

Stimulus/response learning in masked congruency priming of faces: Evidence for covert mental classifications?

Doris Eckstein[1] **and Richard N. Henson**[2]

[1]Department of Psychology, University of Bern, Bern, Switzerland
[2]MRC Cognition and Brain Sciences Unit, Cambridge, UK

Reaction times for categorization of a probe face according to its sex or fame were contrasted as a function of whether the category of a preceding, sandwich-masked prime face was congruent or incongruent. Prime awareness was measured by the ability to later categorize the primes, and this was close to chance and typically uncorrelated with priming. When prime faces were never presented as visible probes within a test, priming was not reliable; when prime faces were also seen as probes, priming was only reliable if visible and masked presentation of faces were interleaved (not simply if primes had been visible in a previous session). In the latter case, priming was independent of experimentally induced face–response or face–category contingencies, ruling out any simple form of stimulus–response learning. We conclude that the reliable masked congruency priming reflects bindings between stimuli and multiple, abstract classifications that can be generated both overtly and covertly.

Keywords: Priming; Learning; Subliminal perception; Response priming; Unconscious; Semantic priming.

The question of whether semantic processing of stimuli is possible without awareness of those stimuli has been the subject of much controversy. The most common paradigm used to address this question is the masked priming paradigm. In the present series of experiments, we used sandwich-masked congruency priming of faces to investigate whether the priming that we observed was indicative of unconscious extraction of semantic information about faces, or whether it reflected instead some form of learning of stimuli and/or responses as faces were repeated across trials (Abrams & Greenwald, 2000; Damian, 2001). Congruency priming refers to the case when the ability to categorize a probe stimulus is facilitated (in terms of faster response times, or more accurate categorization) by a prime stimulus that is perceptually different from the probe, but of the same response category (as distinct from repetition priming, where the prime and probe are examples of the same

---

Correspondence should be addressed to Doris Eckstein, Institut für Psychologie, Universität Bern, Muesmattstrasse 45, 3000 Bern 9, Switzerland. E-mail: doris.eckstein@psy.unibe.ch

This manuscript was funded by Swiss National Foundation Fellowship No. PA001–113106/1 granted to the first author and by the UK Medical Research Council (MC_US_A060_0046). We thank Elias Mouchlianitis and Jason Taylor for collecting some of the data, Dennis Norris and Aidan Horner for helpful discussions, and the Editor and two reviewers for their thoughtful comments.

http://dx.doi.org/10.1080/17470218.2011.590595

stimulus). The reaction time to categorize a face as famous, for example, might be faster, on average, when it is preceded by another famous face (a *congruent* trial) than when it is preceded by a non-famous face (an *incongruent* trial). It is findings like this that have been used by some to infer access to semantic information about the prime. Sandwich masked priming refers to the case when the prime is presented briefly (typically $<$50 ms) and is immediately preceded and succeeded by a pattern mask (see Figure 1). The aim of this masking procedure is to render the prime invisible to participants (as assessed by separate tests; see below), such that any priming effects indicate "subliminal" processing of the prime. Congruency priming from invisible primes therefore suggests unconscious semantic access (Marcel, 1983).

Subliminal semantic priming

Historically, the debate about unconscious semantic processing has focused on how to establish whether masked primes are truly subliminal (see, e.g., Dosher, 1998; Draine & Greenwald, 1998; Greenwald & Draine, 1998; Holender, 1986; Klauer, Eder, Greenwald, & Abrams, 2007; Klauer, Greenwald, & Draine, 1998; Kouider & Dehaene, 2007; Merikle, 1982; Merikle & Daneman, 1998; Merikle & Reingold, 1998). Nonetheless, several recent studies have found congruency priming under conditions in which no awareness of the prime was detected, using clear criteria for prime awareness—for example, when an objective measure of the ability to perform the same categorization task on the prime does not differ from chance (e.g., Greenwald, Abrams, Naccache, & Dehaene, 2003; Greenwald, Klinger, & Schuh, 1995; Kunde, Kiesel, & Hoffmann, 2003; Naccache & Dehaene, 2001a, 2001b).

The debate then moved to the cause of such subliminal priming. One explanation was that it reflected a learning effect that developed over trials, particularly when the stimuli were reused as primes and probes across trials. For example, a probe stimulus might become associated with the response produced to it—so-called "stimulus–response", or "S–R", learning—such that when it is repeated later as a masked prime, the response is automatically retrieved without awareness (Damian, 2001). If this retrieval is fast, and the response retrieved to the prime is congruent with the correct response to the probe, priming can occur. Importantly, this priming would not require unconscious semantic processing of the prime (Abrams & Greenwald, 2000). Nonetheless, later studies of masked priming controlled for such S–R learning—for example, by using "novel" primes (i.e., primes that were never seen as probes; Greenwald et al., 2003; Klauer et al., 2007; Naccache & Dehaene, 2001b; Quinn & Kinoshita, 2007), or by using tasks in which S–R learning is unlikely to contribute to priming (Kinoshita & Kaplan, 2008; Kinoshita & Norris, 2009)—and still found reliable priming. Indeed, a recent meta-analysis supported at least two contributions to masked semantic priming: S–R learning and unconscious semantic access (Van den Bussche, Van Den Noortgate, & Reynvoet, 2009). Thus the purpose of the present paper is not to deny that unconscious semantic access can occur in some situations, but rather to focus on the nature of S–R learning that appears prevalent in many situations. Indeed, the present evidence that S–R learning is more abstract and flexible than previously thought suggests that tighter controls may be necessary before claims are made from masked priming paradigms about the nature of unconscious semantic processing.

Subliminal priming of faces

Most previous studies of masked semantic priming, like those cited above, have used numbers or words. We were interested whether the same subliminal effects would generalize to faces. The ability to quickly categorize and identify faces is essential for adaptive social behaviour. The way that familiar faces pop out of a crowd would suggest that some aspects of faces can be processed without awareness. Indeed, we have previously reported evidence for subliminal *repetition* priming of faces in a fame-judgement task (Henson, Mouchlianitis, Matthews, & Kouider, 2008; Kouider, Eger, Dolan, & Henson, 2008), which cannot be attributed to S–R learning

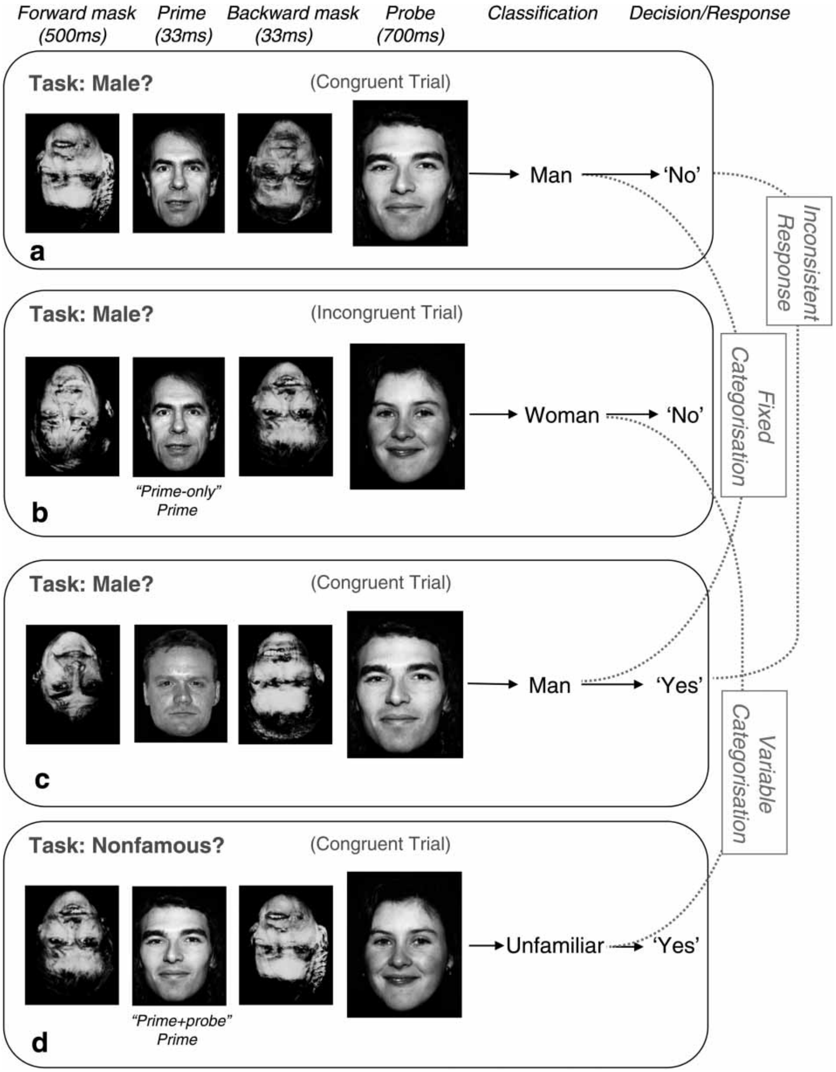

Figure 1. *Example trials in the present experiments, together with possible bindings between stimuli and various types of response.* Prime-only *primes are never shown as probes;* prime + probe *primes are shown as probes in other trials. The first, third, and fourth trials (a, c, and d) illustrate* congruent *trials, because the prime and probe were congruent with respect to the response; the second trial (b) is* incongruent. *Priming is measured as the difference in response time (RT)/accuracy between congruent and incongruent trials. In Experiments 1–3, tasks were to decide whether a face was "male or female" and "famous or nonfamous". In Experiments 4 and 5, four tasks were used ("male?", "female?", "famous?", and "nonfamous?") in which the question had to be answered with "yes" or "no". The probe face in (a) and (c) is used in two different tasks, causing the associated response to be* inconsistent *across trials (50% "no"; 50% "yes";* consistent *probes would always be associated with the same response). However, the classification entailed by those tasks is the same ("male"), and hence this probe stimulus has a* fixed categorization *(if it also appeared as a probe in a famous or nonfamous task, then it would have a* variable categorization*). Note that there were no immediate repetitions of the same prime and/or probe across and within trials in the experiments.*

because prime and probe in the unprimed condition were also associated with the same response. However, this priming could be driven largely by perceptual fluency, consistent with functional magnetic resonance imaging (fMRI) evidence that the effects were restricted to posterior parts of the ventral visual processing stream (Kouider et al., 2009). Such unconscious perceptual processing is less controversial than unconscious semantic processing (Kouider & Dehaene, 2007).

A common explanation for semantic priming is that the prime leaves residual activity within existing semantic representations, or changes the strength of connections between those representations, which then facilitates semantic processing of the probe. An example of such "abstractionist" accounts of priming (Tenpenny, 1995) in the domain of faces is the interactive activation and competition (IAC) model of face recognition developed by Burton and colleagues (Burton, Bruce, & Hancock, 1999; Burton, Bruce, & Johnston, 1990; see also Bruce & Young, 1986). In this model, a familiar prime face induces activity in a face recognition unit (FRU), which then spreads to a personal identity node (PIN) that uniquely indexes a known person, which in turn activates a number of semantic information units (SIUs) that code semantic information associated with that person. Residual activity within SIUs that are also shared with the person denoted by the probe can then facilitate identification of that face by virtue of bidirectional flow of activity back from SIUs to PINs (Burton, Bruce, & Hancock, 1999). In typical semantic priming paradigms using faces, the prime and probe pertain to famous people (known to participants), and priming is defined by the reduction in response time (RT) to make a fame decision about a probe face as a function of whether or not the person denoted by the prime is categorically (or associatively) related to the probe (e.g., a famous person with the same occupation; Wiese & Schweinberger, 2008). Such priming has been reported even when the prime was masked and for the first occurrence of a face as a prime, which precludes any contribution from S–R learning (Stone, 2008; Stone & Valentine, 2006; Wiese, Henson, & Schweinberger, 2011).

An important prediction of the IAC model is that priming will only be found for familiar faces (for which FRUs/PINs/SIUs exist), as was the case for the above masked semantic face priming studies. Indeed many studies originally failed to find priming for unfamiliar (previously unseen) faces (Ellis, Young, & Flude, 1990). More recent work has found repetition priming for unfamiliar faces (Goshen-Gottstein & Ganel, 2000), albeit smaller in size than for familiar faces, one explanation of which is the rapid formation of presemantic, perceptual representations of faces (Martin & Greer, 2011). Such an explanation could not explain semantic priming for unfamiliar faces, however (when prime and probe are typically no more perceptually similar for primed than for unprimed conditions).

Clearly, semantic priming of the type studied by Stone (2008), for example, where prime and probe identities share an occupation, is not possible to test for unfamiliar faces, which possess no associated associative/categorical information. Such priming can be studied using congruency priming, however, since unfamiliar faces can be categorized according to dimensions like famous/nonfamous, or male/female. Only one prior study to our knowledge has investigated masked congruency priming of faces (Enns & Oriet, 2007). These authors showed evidence of priming of sex/fame categorizations, under conditions of minimal awareness of the prime. Importantly, however, this study used faces that repeated across trials—that is, primes that also appeared as probes in other trials—so the priming could reflect S–R learning. We consider such alternative "episodic" accounts of priming next.

S–R learning in priming

The importance of S–R learning in priming has been appreciated for many years (e.g., Hommel, 1998; Logan, 1990). These theories assume that a single pairing of a stimulus and a response is sufficient to form a new "episodic" representation (Tenpenny, 1995) that can be retrieved when that stimulus is repeated, thereby facilitating production of the associated response (assuming it is congruent with the response required by the task). There has been a recent resurgence of interest in such S–R learning, particularly whether it is more abstract than previously assumed (e.g., Abrams, Klinger, & Greenwald, 2002; Kiesel, Kunde, &

Hoffmann, 2007; Kiesel et al., 2006; Klauer, Musch, & Eder, 2005). Indeed, in the context of unmasked, across-trial repetition priming, Horner and Henson (2009) reported evidence for simultaneous binding of stimuli to at least three levels of response code: an action (e.g., finger press), a decision (e.g., yes/no), and a task-specific classification (e.g., famous/nonfamous). This view is compatible with the *event file* concept of Hommel and collaborators, according to which stimuli and responses are bound in a structured fashion—that is, with separate associations between relevant stimulus features, responses, and tasks (Hommel & Colzato, 2009; Waszak & Hommel, 2007). By manipulating the history of prior occurrences of a stimulus as a function of its associated responses and tasks, the present experiments go further to suggest that stimuli presented visibly as probes become associated with multiple, abstract classifications, even if those classifications are only made covertly. In the present paradigm at least, this questions the need to appeal to unconscious activation of semantic representations.

PRESENT EXPERIMENTS

We began the present series of experiments by comparing congruency priming of sex and fame decisions about faces as a function of whether the prime faces were also used as probes in other trials (*prime + probe* primes; Figure 1d), or were never shown as probes (*prime-only* primes; Figure 1b). We measured participants' ability to see the primes in a subsequent phase in which they were told about the presence of the primes and were asked to perform the same categorization task on them (rather than the probes). In Experiments 1–3, we consistently found priming from prime + probe primes, concurrent with no reliable evidence that participants could categorize the primes, but did not find evidence for priming from prime-only primes, even when they had been presented multiple times as probes for a different task in a previous session. Moreover, prime + probe priming was found for both famous and nonfamous (previously unfamiliar) faces. This pattern of findings is difficult to reconcile with abstractionist accounts of priming, like that proposed by the IAC model of face recognition, but is consistent with S–R learning theories, assuming such learning only takes place when a stimulus appears as a probe. This prompted Experiments 4 and 5, in which we explored the nature of the priming for prime + probe trials, as a function of (a) whether or not a given probe face was always associated with the same "yes/no" response across trials (*consistent* trials) or not (*inconsistent* trials), and (b) whether that face appeared as a probe in one type of sex/fame categorization (*fixed* categorization) or both types (*variable* categorization). The robust priming found in all cases led us to suggest that prime + probe priming reflects bindings between stimuli and multiple, abstract classifications, even if those classifications are only generated covertly as a consequence of a switching task context. We elaborate this and other possible explanations in the General Discussion.

EXPERIMENT 1

This experiment was designed to verify whether it is possible to observe congruency priming from subliminal faces. Given that previous research on subliminal priming with numbers and words had demonstrated the importance of whether the stimuli used as primes are also used as probes (on other trials), we compared two conditions: the prime + probe condition in which prime stimuli also appeared as probes, and the prime-only condition in which they did not (in fact, these faces were never shown unmasked prior to or during the priming phase).[1] Assuming that categorization of faces shares some characteristics with categorization of numbers and words, we predicted greater priming in the prime + probe condition than in

---

[1] While these might also be called "novel" primes (e.g., Kunde et al., 2003), we prefer to reserve that label for stimuli that are only ever presented once as a prime throughout an experiment, unlike here, where prime-only stimuli were repeated as primes on multiple trials within the experiment.

the prime-only condition, and we were interested in whether any evidence would be found for priming in the prime-only condition.

We used a sandwich masked paradigm, in which the forward and backward masks were created by superimposing a number of inverted faces (Figure 1). We have previously found such masks to be effective in minimizing awareness for an interspersed upright face (Kouider et al., 2009). Pilot studies showed that, with our stimulus set, a 33-ms duration of prime and backward mask (in conjunction with a different, 500-ms forward mask) was sufficient for participants not to notice the masked face.

Participants performed one of two speeded categorization tasks on the probe: either a male/female sex judgement or a famous/nonfamous fame judgement (task was manipulated between participants). We used two tasks to test the generalization of any priming effects (e.g., in case priming of fame judgements to nonfamous versus famous faces was biased by their different response categories, or in case priming of sex judgements was caused by perceptual similarity between same-sex prime and probes). The same set of 32 faces (with equal numbers of famous/nonfamous and male/female faces) was used in both tasks.

The experiment was divided into three phases: the main phase in which priming was measured, followed by a discrimination phase in which participants were told about the presence of the prime faces and were asked to categorize them in the same manner as they had for the probes in the main experiment. In a final debriefing phase, all the prime-only primes were shown visibly, to check that they could be categorized correctly.

Method

Participants

Twenty-six participants were paid volunteers of the Medical Research Council (MRC) volunteer panel. Data from 12 participants were used in each task (sex task: 5 men and 7 women, mean age of 32.3 years; fame task: 3 mean and 9 women, mean age = 34.5 years). One participant was replaced because of insufficient recognition of famous faces in the final prime recognition phase; another was replaced because of an outlier d' score in the prime discrimination phase (see below). When participants were replaced, conditions were assigned such that counterbalancing was maintained. All participants had normal or corrected-to-normal vision. The experiments were of the type approved by a local research ethics committee (Cambridge Psychological Research Ethics Committee Reference 2005.08).

Design

Data were analysed by analysis of variance (ANOVA) with 3 two-level factors: task (sex vs. fame; between-participants), prime type (prime + probe vs. prime-only; within-participants), and priming (congruent vs. incongruent; within-participants).

Materials and apparatus

Eight famous male, eight famous female, eight nonfamous male, and eight nonfamous female faces were selected as the most accurately categorized from a set of faces used previously (Henson et al., 2008; Kouider et al., 2009). None of the faces contained obvious sex-predictive (or fame-predictive) features, such as facial hair or excessive cosmetics. A nested, balanced, incomplete block design was used in order to ensure that faces were evenly and independently assigned to the prime + probe and prime-only conditions across participants. In a block of 32 trials, each face was shown once as a prime, and each prime + probe face was shown twice as a probe. Each of the eight conditions of interest, 2 (prime + probe/prime-only) × 2 (probe fame/sex) × 2 (prime-probe congruency), was shown four times in 32 trials, with the nontask dimension and nontask congruency (e.g., probe sex and sex congruency in the fame task) equally represented across those four presentations. Each prime was paired maximally twice with a specific probe during the whole priming phase. The average lag between the presentation of a prime + probe face as a probe and its presentation as a prime was 11.6 trials.

Sixty-four masks were created by overlaying four inverted faces (half famous, half female). In order to minimize pixel overlap with the probe face, the prime was scaled to be 80% smaller than the probe (masks received the same size reduction for masking improvement reasons). A different mask was used as the forward and backward mask on a given trial, and otherwise the masks were randomly assigned to trials, with the constraint that each mask was used once every 32 trials.

The experiment was conducted on a PC using an external button box to collect responses. A CRT monitor was used for display of stimuli using a 60-Hz refresh rate.

Procedure

The experiment consisted of three phases: a priming phase, a prime discrimination phase, and a prime recognition phase. In all phases, participants responded by pressing one of two keys using their left and right index fingers. The assignment of buttons (fingers) to responses was fixed across phases within a participant, but was counterbalanced across participants.

In the main priming phase, participants were instructed to classify the faces as male or female (the sex-task group), or as famous or nonfamous (the fame-task group), whereby emphasis was given on fast responses, without making mistakes. A total of 20 practice trials were given using a set of faces that were not used again. The priming phase consisted of six sessions of 128 trials with breaks between sessions (each face was thus shown 48 times as a probe during the experiment).

The timing for each trial timing was as follows: After a 500-ms fixation cross and a 500-ms forward mask, the prime was shown for two screen refreshes (33 ms), then replaced by the backward mask for two refreshes, followed by the probe for 700 ms, followed by a fixation cross (Figure 1). The response window duration was 2 s. Probe faces subtended an area of 8 cm (height) × 6 cm (width), covering a viewing angle of about 7.6 × 5.7 degrees.

After the priming phase, the experimenter asked the participants whether they had seen any of the primes, after which they were informed about the presence of the prime faces and performed the second prime discrimination phase. The instruction screen for this phase consisted of an explanation of the masking sequence, including an illustration of the sequence of pictures in the trial. Participants were asked to perform the same categorization task as that in the previous priming phase, but this time basing their decision on the prime face. Practice was done first using a 100-ms duration for display of the primes in order to render primes more visible. If unsure, participants were encouraged to guess. Then followed a rerandomized sequence of the first 128 trials of the priming phase, with a response window duration of 2 s.

In the final prime recognition phase, participants were asked to classify the faces as before, whereby the 16 prime-only faces were shown in a random sequence for 700 ms using a response window of 2.4 s. Participants who incorrectly categorized more than two of the prime-only primes in the final check were excluded from analysis, which resulted in removal of 1 participant from the fame-task group.

Analysis

Trials with outlier responses (RTs greater than 2.5 standard deviations of each participant's average RT) and trials where the prime duration exceeded 33 ms (due to software interruptions; 1–5 trials in every 2–3 participants) were excluded from analysis. Participants with an absolute d' greater than 1 in the prime-discrimination phase were excluded. Priming in each condition (incongruent – congruent, for both RTs and errors) was tested as being greater than zero with one-sided tests using a Type I error criterion of .05.

Results

Prime discrimination

Four participants in the fame task and 2 in the sex task reported having occasionally seen another upright face flashed before the probe face. Nonetheless, mean overall d' from the prime discrimination phase was not greater than zero, $d' = 0.10$, $SE = 0.06$, $t(23) = 1.59$, $p = .13$. When split by condition, prime discrimination was not greater than zero either for prime + probe primes, $d' = .09$, $SE = 0.08$, $t(23) = 1.15$, $p > .20$, or for prime-only primes, $d' = 0.11$, $SE = 0.08$, $t(23) = 1.31$, $p > .20$.

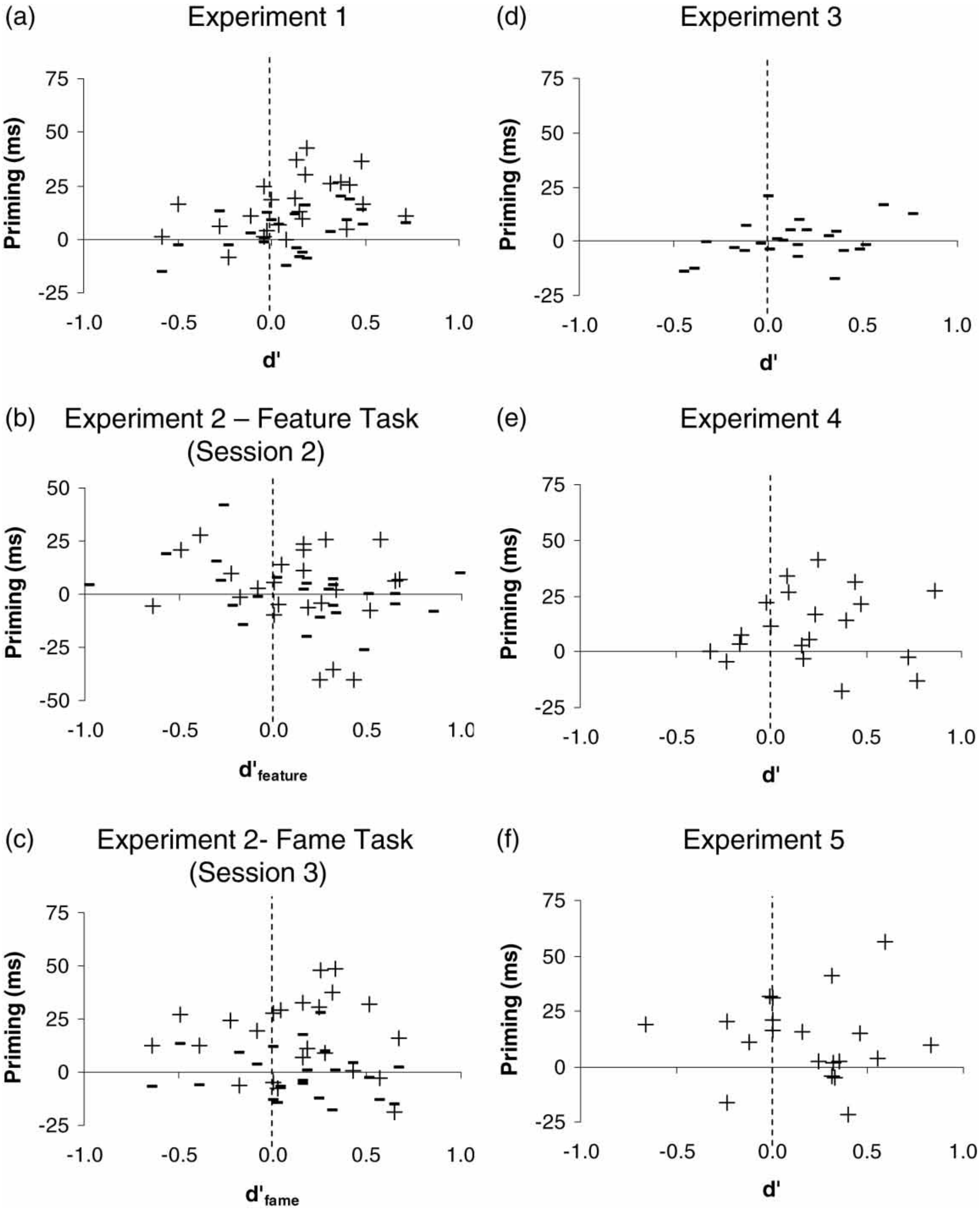

Figure 2. *Scatterplots of response time (RT) priming (incongruent–congruent) and* d′ *(from discrimination phase) across participants from Experiments 1 to 5. "+"= prime + probe trials; "–" = prime-only trials.*

The distribution of *d*′ straddled zero (Figure 2a), as expected if the true *d*′ were zero, and the variance across participants reflected random measurement error in estimation of *d*′.

Nonetheless, there was a reliable positive correlation between *d*′ and priming ($r = .51$, $p = .011$, and $r = .52$, $p = .009$, for prime-only and prime + probe primes, respectively). We therefore conducted a further regression analysis, corrected for measurement error on the predictor variable (see Klauer, Draine, & Greenwald, 1998). This analysis indicated that priming remained reliable when extrapolated to zero d' (i.e., the y-axis intercept was reliably greater than zero) for both prime-only primes, $M = 2.06$, $SE = 0.84$, $t(23) = 2.46$, $p = .01$, and for prime + probe primes, $M = 13.98$, $SE = 2.08$, $t(23) = 6.71$, $p < .001$.

Congruency priming

Error rates were less than 3% on average. RTs and error rates for each condition are given in Table 1. In the 2 × 2 × 2 omnibus, mixed-effects ANOVA on RTs, a significant effect of task indicated that overall RTs were faster in the sex task than in the fame task, $F(1, 22) = 10.07$, $MSE = 9{,}605$, $p = .004$. The interaction of task and priming was not reliable, $F(1, 22) = 0.36$, $MSE = 85.41$, $p > .20$, though the interaction of task, priming, and prime type approached significance, $F(1, 22) = 3.51$, $MSE = 27.39$, $p = .074$. The main effect of priming was significant, $F(1, 22) = 23.26$, $p < .001$, but, crucially, there was also a significant interaction of priming and prime type, $F(1, 22) = 26.79$, $MSE = 85.41$, $p < .001$. As can be seen in Table 1, priming by prime + probe primes was significant in both fame and sex tasks, but did not reach significance for prime-only primes in either task.

The same ANOVA on arcsine-transformed accuracies revealed a significant main effect of priming, $F(1, 22) = 4.54$, $MSE = 1.27$, $p = .045$, and a significant interaction of priming and prime type, $F(1, 22) = 4.56$, $MSE = 1.00$, $p = .044$. Responses in congruent prime + probe trials were more accurate than responses in incongruent prime + probe trials (which remained reliable within the fame task alone, see Table 1), while no significant difference was seen with prime-only primes. This suggests that the RT priming for prime + probe trials above was not a consequence of a speed–accuracy trade-off. No other effect approached significance.

Finally, we explored the RT priming effects as a function of the fame and the sex of the probe faces. One reason for this analysis is that, if congruency priming reflected residual activity in semantic representations, as assumed, for example, by the IAC model of Burton et al. (1999), then it would only be expected for famous faces, for which such semantic representations exist. A 2 × 2 × 2 × 2 ANOVA on the fame task, with factors of priming, prime type, probe fame, and probe sex, showed a reliable interaction between priming and probe fame, $F(1, 11) = 5.74$, $MSE = 155$, $p = .035$ (as well as a reliable main effect of priming, which was qualified by a reliable interaction between priming and prime type, as expected from the omnibus ANOVA above), but no other effects reached significance, $F\text{s} < 1.88$, $p\text{s} > .20$. Contrary to what one might expect from the IAC model, however, this interaction between probe fame and priming actually reflected greater congruency priming for nonfamous than for famous faces. Indeed, prime + probe priming was significant for nonfamous probes, both female and male ($M = 20.2$ ms, $SE = 4.6$ ms, and $M = 25.5$ ms, $SE = 5.5$ ms, respectively, one-tailed $p\text{s} < .001$). Though smaller in size, prime + probe priming was also significant for both female and male famous probes ($M = 12.4$ ms, $SE = 4.5$ ms, and $M = 14.4$ ms, $SE = 5.7$ ms, respectively, $p\text{s} < .05$). Priming was not significant for any of the four prime-only conditions, except for nonfamous, female probes ($M = 11.0$, $SE = 4.3$, $p < .05$), though the latter would not survive correction for multiple comparisons and was not replicated in Experiment 2.

The corresponding ANOVA on the sex task showed an analogous interaction between priming and probe sex, $F(1, 11) = 18.95$, $MSE = 85$, $p < .001$ (as well as a reliable main effect of priming, which was qualified by a reliable interaction between priming and prime type, as expected from the omnibus ANOVA above). This interaction reflected greater priming for male than for female faces. Prime + probe priming was significant for male probes, both famous and nonfamous ($M = 17.3$ ms, $SE = 5.7$ ms, and $M = 19.3$ ms, $SE = 5.6$ ms, respectively, $p\text{s} < .01$). Prime + probe priming approached significance for female famous probes ($M = 7.5$ ms, $SE = 4.4$ ms, $p = .056$), but not for female nonfamous probes ($M = 1.5$ ms, $SE = 4.3$ ms). Priming was not significant for female prime-only conditions, but was

Table 1. *Results of Experiments 1–5*

						Response times (ms)					*Error rates (%)*			
Experiment	*N*	*Session*	*Task*	*Probe categorization*	*Response*	*Prime type*	*Incongruent*	*Congruent*	*Priming*	*t*	*Incongruent*	*Congruent*	*Priming*	*t*
1	12 (fame)		Fame	Fixed	Probe consistent	Prime + probe	585 (45)	568 (49)	18** (3)	5.59	3.4 (0.4)	2.2 (0.4)	1.2** (0.4)	3.38
			Fame	Fixed	Prime inconsistent	Prime-only	575 (48)	573 (49)	3 (2)	1.75	2.6 (0.4)	2.8 (0.4)	–0.2 (0.3)	–0.68
	12 (sex)		Sex	Fixed	Probe consistent	Prime + probe	519 (47)	507 (54)	11** (4)	3.11	3.3 (0.5)	2.5 (0.4)	0.8 (0.6)	1.37
			Sex	Fixed	Prime inconsistent	Prime-only	514 (48)	509 (52)	4 (3)	1.31	2.8 (0.4)	2.5 (0.4)	0.4 (0.4)	0.89
2	24	1	Fame	Fixed	Probe consistent	Prime + probe	577 (40)	567 (43)	10** (3)	3.22	2.4 (2.9)	1.6 (1.3)	0.8 (0.7)	0.63
			Fame	Fixed	Prime inconsistent	Prime-only	568 (38)	564 (40)	4 (3)	1.26	2.1 (2.5)	1.4 (1.3)	0.8 (0.4)	1.32
		2[a]	Feature	Fixed	Probe consistent	Prime + probe	635 (76)	633 (82)	2 (4)	0.48	6.6 (3.6)	7.0 (7.0)	–0.3 (1.4)	0.04
			Feature	Fixed	Prime inconsistent	Prime-only	628 (81)	627 (78)	1 (3)	0.22	6.8 (5.9)	6.7 (4.5)	0.1 (1.0)	0.29
		3[b]	Fame	Fixed	Probe consistent	Prime + probe	557 (44)	541 (48)	16** (4)	4.48	3.1 (3.6)	1.2 (1.5)	1.9* (0.8)	2.23
			Fame	Fixed	Prime inconsistent	Prime-only	546 (45)	548 (52)	–2 (3)	–0.74	1.8 (1.7)	2.1 (1.8)	–0.3 (0.5)	–0.56
3	12 (fame)		Fame	Fixed	Prime consistent	Prime-only	643 (51)	638 (50)	5 (5)	1.1	9.4 (8.2)	10.1 (9.3)	–0.6 (0.5)	1.16
			Fame	Fixed	Prime inconsistent	Prime-only	644 (47)	643 (56)	1 (5)	0.2	9.2 (7.4)	9.6 (8.1)	–0.5 (0.7)	0.72
	12 (sex)		Sex	Fixed	Prime consistent	Prime-only	622 (73)	626 (73)	–4 (6)	–0.2	1.5 (1.1)	1.6 (1.1)	–0.1 (0.4)	0.26
			Sex	Fixed	Prime inconsistent	Prime-only	622 (69)	624 (77)	–2 (3)	–0.3	2.0 (1.3)	2.3 (1.6)	–0.2 (0.3)	0.67
4	20		Sex/fame	Variable	Probe consistent	Prime + probe	632 (75)	624 (74)	8* (4)	1.99	7.2 (5.4)	8.5 (7.7)	–1.4 (1.2)	–1.05
			Sex/fame	Variable	Probe inconsistent	Prime + probe	633 (77)	619 (70)	13* (4)	3.06	6.5 (5.5)	5.5 (4.0)	0.9 (1.4)	0.66
5	20		Sex/fame	Variable	Probe inconsistent	Prime + probe	644 (63)	633 (62)	11* (5)	2.42	8.4 (5.5)	7.3 (4.8)	1.1 (1.1)	1.00
			Sex/fame	Fixed	Probe inconsistent	Prime + probe	643 (69)	628 (58)	14* (5)	2.65	8.1 (4.6)	9.5 (5.0)	–1.4 (0.7)	–1.93

Note: Standard deviations of mean response times and error rates and standard errors of mean priming are given in parentheses.
[a] Prime-only primes were probes in Session 1. [b] Prime-only primes were probes in Session 2.
*$p < .05$, one-tailed. **$p < .01$, one-tailed.

for male famous and nonfamous prime-only conditions ($M = 9.3$ ms, $SE = 4.9$ ms, and $M = 8.0$, $SE = 4.3$ ms, respectively, $ps < .05$), though again these would not survive correction for multiple comparisons and were not replicated in Experiment 2. While the significance levels become questionable for such post hoc fractionation into more specific conditions, given the reduced amount of data and increased number of statistical comparisons, the main message of these analyses is that the congruency priming effects did not require previous familiarity with the face stimuli, in that they occurred equally often for nonfamous as for famous faces.

Discussion

Experiment 1 showed congruency priming for both sex- and fame-categorization tasks (in both RT and accuracy), under conditions in which participants' ability to perform the same task on the prime (when subsequently alerted to its presence) was not distinguishable from chance. This suggests some form of subliminal processing. However, priming was significantly greater for primes that appeared as probes on other trials than for primes never seen as probes. One explanation for this is that the subliminal processing reflected the implicit retrieval of a response that has become associated with a face during prior trials in which that face appeared as a probe—that is, a form of stimulus–response (S–R) learning (Abrams & Greenwald, 2000; Damian, 2001).[2]

Furthermore, masked congruency priming was found for nonfamous as well as famous faces, even when the task was unrelated to fame (i.e., for the sex task). Prima facie, this would seem difficult to explain by any abstractionist theory, such as the IAC model of face recognition (Burton et al., 1999), that appeals to priming-related changes within preexisting perceptual and/or semantic representations. Nonetheless, one could argue that the multiple unmasked presentations across trials of the nonfamous faces in the prime + probe condition (but not when masked in the prime-only condition) are sufficient to establish new representations of those faces (e.g., Martin & Greer, 2011). Indeed, the cause of prime + probe priming for both nonfamous and famous faces could reflect some form of perceptual learning—for example, improved bindings between the features of each face. According to such an account, participants would only need to see a face a few times visibly for it to function as an effective masked prime, with no need to classify it according to any of the task-relevant dimensions (unlike an S–R learning account). The repeated conscious or prolonged nature of processing of a face when it appears as a probe may refine its perceptual representation, such that when it appears later as a masked prime, its perceptual processing occurs faster, leaving relatively more time for postperceptual processing of the (semantic) information necessary for the current task, which might in turn facilitate semantic processing of the following probe (when congruent), hence priming. While a perceptual learning account would seem incompatible with a previous failure to find masked congruency priming from simple preexposure to stimuli (Experiment 3 of Damian, 2001), Damian's experiment used familiar words, which are arguably less perceptually complex than the faces used here. Therefore we deemed the perceptual learning hypothesis to be worth testing for our specific stimuli and paradigm.

EXPERIMENT 2

An obvious way to test this perceptual learning account would simply be to preexpose one half of

---

[2] To test the hypothesis that S–R learning was incremental, we performed a post hoc analysis where priming in the prime + probe condition was split as a function of the number of prior presentations of the face as a probe, binned every 4 presentations (i.e., 0–3 prior presentations as a probe, 4–7, etc., leading to 6 levels of exposure). The interaction between congruency and bin did not reach significance, $F(5, 18) = 1.04$, $MSE = 423$, $p > .20$. However, the congruency effect was not significantly different from zero after the first 0–3 presentations, $M = 5.7$ ms, $SE = 7.0$ ms, $t(23) = 0.81$, $p > .20$, whereas it reached significance in all later bins ($M > 11$ ms, $SE < 6.3$ ms), $t(23) > 1.75$, $p < .05$, except Bin 5 ($M = 8.7$, $SE = 7.0$ ms), $t(23) = 1.24$, $p = .11$. Thus, while noisy (owing to small numbers of trials), these binned data suggest fairly rapid learning of an S–R association, in which a face requires only a few trials of being categorized as a (visible) probe before masked priming by that face (as an invisible prime) occurs.

the faces without any explicit task (to prevent any S–R learning of overt responses). However, the lack of an explicit task may not encourage sufficient perceptual processing of the faces. In Experiment 2, we therefore exposed faces as probes in a first session employing one task, before using them as masked primes in a second session that employed a different task that was unrelated to the first task. In fact, one set of stimuli was used for prime + probe trials in the first task, but for prime-only trials in the second task (while a second set of stimuli, which was used for prime-only trials in the first task, was used for prime + probe trials in the second task)—as shown in Figure 3.

For the first session, we used the same fame task as that in Experiment 1. For the second session, however, we switched to a new task, which was to decide whether the features of the face were “wide apart” or “compressed” (see Method; henceforth, the “feature” task). We reasoned that this task requires careful, configural perceptual processing of the faces. The reason for choosing such an ad hoc classification was because, had we used a more natural categorization (e.g., the sex task in Experiment 1), participants may have spontaneously, covertly categorized probe stimuli (e.g., as “male” or “female”) during the first session, even though they were only required to classify overtly according to “famous”/“nonfamous”. In other words, participants may automatically classify visible faces according to certain dimensions, such as their sex, for evolutionary reasons. We reasoned that it was unlikely that participants would covertly classify faces as wide apart/compressed in the first session, given that this classification was only explained to them at the start of the second session. This would mean that any priming for prime-only primes in the second session could not reflect retrieval of spontaneously generated responses from their visible appearance in the first session (S–R learning) and therefore provide clear support for the perceptual learning account.

Finally, we added a third session in which participants reverted back to the initial fame task,

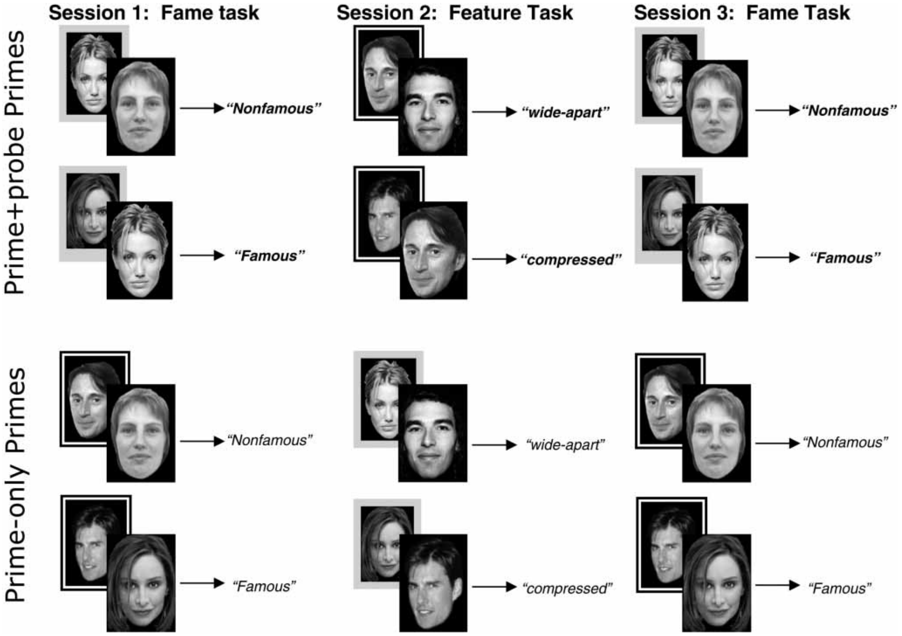

Figure 3. *Illustration of the design used in Experiment 2. Faces used as prime-only primes in Session 2 had been used as prime + probe primes in Session 1 (indicated by light grey border), while faces used as prime + probe faces in Session 2 had been used as prime-only primes in Session 1 (indicated by dark grey border). An analogous swap occurred between Sessions 2 and 3.*

and in which prime + probe trials in the second session (with the feature task) were used as prime-only trials (and in which prime-only trials in the second session were used as prime + probe trials). One reason for this final session was to provide a further test of the perceptual learning account, which should again predict priming for prime-only trials, but now in the fame task. Another reason was to compare the size of any prime-only priming effects across first and last sessions, when using exactly the same stimuli and task, differing only in whether those stimuli had been seen as probes during the intervening second session.

Method

Participants

Twenty-seven participants were recruited using same methods as those in Experiment 1. The data of 3 participants were replaced, 1 participant for not recognizing 5 out of 8 famous faces and 2 participants for using the same response in more than 80% of the trials in the feature task. This left 12 men and 12 women (mean age 28.1 years).

Design

A within-participant design was implemented with the factors prime type (prime + probe vs. prime only), session (1–3), and priming (congruent vs. incongruent). Given the a priori predictions for the different tasks in Sessions 1–3, the data were analysed by 2 × 2 prime type by priming ANOVAs on each session separately.

Materials and apparatus

The four most accurately categorized faces of each of the famous male, famous female, nonfamous male, and nonfamous female probe categories were selected from Experiment 1. Five perceptual binary classification tasks were piloted, of which the wide-apart/compressed task produced performance closest to an equal split of our stimuli according to the two alternative responses. Assignment of faces to priming conditions and apparatus was the same as that in Experiment 1.

Procedure and analysis

The procedure and analysis were identical to those in Experiment 1, with the following exceptions: (a) In the main priming phase, participants first did 256 trials using the fame task (Session 1), then 256 trials using the feature task (Session 2), and then 256 trials using the fame task again (Session 3); (b) in the feature task, participants were instructed to decide “as quickly as possible whether eyes, nose and mouth are rather compressed or rather wide-apart” (compared to the average); (c) 20 practice trials were given at the start of the first fame and feature sessions, but not at the start of the final fame session; (d) faces used as prime-only faces in the fame task were used as prime + probe faces in the feature task, and vice versa, with each prime + probe face appearing 32 times as a (visible) probe within a session; and (e) in the prime discrimination phase, 32 trials of the fame task and 32 trials of the feature task were run, with the instruction that participants should now classify the primes. Because (unlike the fame and sex tasks) the feature task is somewhat subjective, “errors” were defined as responses that differed from the modal response given for each stimulus by each participant.

Results

Prime discrimination

Five participants reported having occasionally seen another upright face flashing up before the probe face. Average d' was not significantly greater than zero, $d'_{\text{fame}} = 0.12$, $SE = 0.10$, $t(23) = 1.21$, $p > .20$, and $d'_{\text{feature}} = 0.15$, $SE = 0.14$, $t(23) = 1.11$, $p > .20$, and there was no obvious relationship with the amount of priming in either task (Figures 2b and 2c). Correlation of priming with d' was not significantly greater than zero (rs ranging from −.33 to .15, $p > .20$ in each condition/task).

Congruency priming

Error rates for the fame task were less than 3% on average, while “error” rates for the more subjective feature task (see Method) were less than 7% on average (Table 1).

The ANOVA on the initial (fame) session did not show an interaction between priming and prime type that reached significance, $F(1, 23) = 2.17$, $MSE = 110.92$, $p = .15$, though, as in Experiment 1, priming was significant for prime + probe primes but not for prime-only primes (see Table 1). Both the main effect of priming, $F(1, 23) = 10.48$, $MSE = 106.53$, $p = .004$, and the main effect of prime type, $F(1, 23) = 7.22$, $MSE = 101.21$, $p = .013$, were significant.

The ANOVA on the second (feature) session did not show an interaction between priming and prime type that reached significance, $F(1, 23) = 0.08$, $MSE = 130.43$, $p > .20$. Unlike the fame task in the first session, priming was not significant for either prime + probe primes or prime-only primes (see Table 1). The main effect of prime type, however, was again significant, $F(1, 23) = 10.02$, $MSE = 94.15$, $p = .004$.[3]

The ANOVA on the final (fame) session did show a significant interaction between priming and prime type, $F(1, 23) = 21.26$, $MSE = 93.59$, $p < .001$, as well as a significant main effect of priming, $F(1, 23) = 8.48$, $MSE = 146.31$, $p = .008$, and priming was once again significant for prime + probe primes but not for prime-only primes (Table 1). Importantly, prime-only priming was numerically smaller, rather than bigger, in the final session than in the initial session, with a direct comparison of prime-only priming across sessions showing no evidence of an effect of intervening exposure to the prime stimuli in the second session, $M = -5.57$, $SE = 3.69$, $t(23) = -1.51$, $p > .20$. By contrast, prime + probe priming showed a trend for greater priming across first and third sessions, $M = 6.30$, $SE = 4.35$, $t(23) = 1.45$, $p = .08$ (one-tailed).

The same analysis on arcsine-transformed accuracies on each session separately did not reveal significant priming in the initial, $F(1, 23) = 1.99$, $MSE = 110$, $p = .17$, second, $F(1, 23) = 0.07$, $MSE = 110$, $p > .20$, or final session, $F(1, 23) = 2.18$, $MSE = 140$, $p = .15$. The interaction between priming and prime type approached significance in the final session, $F(1, 23) = 3.77$, $MSE = 190$, $p = .064$, with significant priming for prime + probe primes, $t(23) = 2.23$, $SE = 4$, $p = .018$, reflecting more errors for incongruent than for incongruent prime + probe trials (and demonstrating that the analogous RT priming did not reflect a speed–accuracy trade-off). No other priming effects on accuracy reached significance, ts < 1.5, $p > .05$.

Finally, as in Experiment 1, we explored the RT priming effects as a function of the fame and the sex of the probe faces. For the fame task, and averaging across initial and final sessions to maximize power, the $2 \times 2 \times 2 \times 2$ ANOVA, with factors of priming, prime type, probe fame, and probe sex, showed only a reliable interaction between priming and probe fame, $F(1, 23) = 12.90$, $MSE = 237$, $p < .005$ (as well as a reliable main effect of priming, which was qualified by a reliable interaction between priming and prime type, as expected from the omnibus ANOVAs above), but no other effects reached significance, Fs < 2.06, ps $> .17$. As in Experiment 1, this interaction reflected greater priming for nonfamous than famous probes. Prime + probe priming was significant for nonfamous probes, both female and male ($M = 23.5$ ms, $SE = 3.9$ ms, and $M = 19.3$ ms, $SE = 5.6$ ms, respectively, ps $< .005$) and approached significance for female famous probes ($M = 7.4$ ms, $SE = 4.6$ ms, $p = .06$), but not for male famous probes ($M = 4.3$ ms, $SE = 3.9$ ms, $p = .14$). Priming was not significant for any of the four prime-only conditions ($M < 2.8$ ms, $SE > 3.9$ ms). An analogous ANOVA performed on the feature task in the second session showed no effects of probe fame, only a borderline interaction between priming and probe sex, which suggested more positive priming for female than male faces. However, no effects of priming reached significance for any of the eight prime + probe and prime-only conditions ($M < 8.0$ ms, $SE > 8.7$ ms, $p > .18$).

---

[3] In order to test for spurious associations between the fame decision and the wide-apart decision, a wideness score was computed for each face—that is, the average proportion of "wide-apart" decisions for a probe face across participants. Scores ranged from .09 to .83, whereby 4 famous and 4 nonfamous faces had scores lower than or equal to .5, and 4 famous and 4 nonfamous faces had scores higher than .5. This suggests that there was no correlation across items between the two decisions.

Discussion

In this experiment, we tested a perceptual learning hypothesis that might account for the priming from prime + probe primes, but not prime-only primes, in Experiment 1. According to this hypothesis, multiple visible perceptions of an attended stimulus are sufficient for that stimulus to be processed more rapidly when reused as a masked prime (e.g., via establishment of a new perceptual representation; Martin & Greer, 2011), in turn allowing more time for semantic analysis and hence congruency priming. We tested this hypothesis by measuring priming by prime-only primes in one task before and after visible presentation of the prime faces as probes in a different task. There was no support for the perceptual learning account, however: Priming was not reliable for (a) prime-only primes in the feature task of the second session, which had appeared 32 times as probes in the fame task of the first session, nor for (b) prime-only primes in the fame task of the third session, which had appeared 32 times as probes in the feature task of the second session. At the same time, priming for prime + probe trials within the fame task was reliable in both the first and the third session, replicating Experiment 1. The prime faces in these prime + probe trials had been seen visibly as probes no more often[4] than those used for the prime-only trials in the second and third sessions, suggesting that they cause priming not through perceptual learning, but through association with some form of task-relevant response (i.e., S–R learning).

It should be noted that priming was not reliable for prime + probe trials within the feature task of the second session. Such priming might have been expected according to an S–R learning account, in the same way that prime + probe priming was found in the fame task. One reason may relate to task differences: The ad hoc classification of faces according to the spatial distribution of their features may involve a range of strategies/processes, some of which are less prone to priming. Indeed, masked congruency priming might be restricted to "natural" categorizations (like familiarity and sex) that participants perform in everyday life. Alternatively, priming in the feature task may have emerged as reliable with more trials (e.g., there were only one half as many trials per session in this experiment as in Experiment 1). Nonetheless, the fact that priming for prime + probe primes was reliable in the final fame task, together with the concurrent absence of reliable priming for prime-only primes within that same session (in conjunction with the reliable interaction between the size of priming across prime + probe and prime-only conditions within that session), suggests at a minimum that perceptual learning is not a sufficient explanation for the present prime + probe priming.

In the next experiment, we therefore returned to an S–R learning account of the pattern of priming in Experiments 1–2. Indeed, further consideration of such an account suggested another possible reason for the absence of priming for prime-only trials: a difference in the consistency of the response associated with a given face when that face was used in the prime-only relative to prime + probe conditions.

EXPERIMENT 3

While the data from Experiments 1–2 are consistent with a response becoming associated with a visible probe face, it is also logically possible that a response can become associated with the "invisible" prime face that occurred on the same trial. For example, the prime, probe, and response might become bound into a single event-record (Hommel, 1998). In this case, however, Experiments 1 and 2 were not a fair test of any S–R learning between the prime and the response.

---

[4] In fact, such prime + probe faces were seen less often on average within a session than prime-only faces were across session, though it should be noted that prime + probe faces were nonetheless seen more recently on average than prime-only faces. This issue is considered in the General Discussion.

This is because each prime face was paired equally often, on average, with each of the two possible responses (e.g., 50% female/50% male). This would prevent a consistent mapping between a response and a prime face in the prime-only condition from being learned. This is unlike prime faces that also occurred as probes (in the prime + probe condition), which would have had a more consistent mapping between a face (whether occurring as prime or probe) and a response. This is because the response is determined by the probe, so that when collapsing across all presentations as a prime or a probe, faces in the prime + probe condition were associated with one response on 75% of trials (and the other response on 25%; see Figure 4). This response would be the correct response when that face appeared as a prime in a congruent trial and so could facilitate RTs for congruent trials. In other words, differences in the consistency of mapping between a response and a face, whether that face appears as a prime or probe, could explain the greater priming found for the prime + probe condition than for the prime-only condition. In yet other words, if responses are learned to both prime and probe faces within a trial, there would be a bias towards one response (the congruent response) in the prime + probe condition, but not in the prime-only condition.

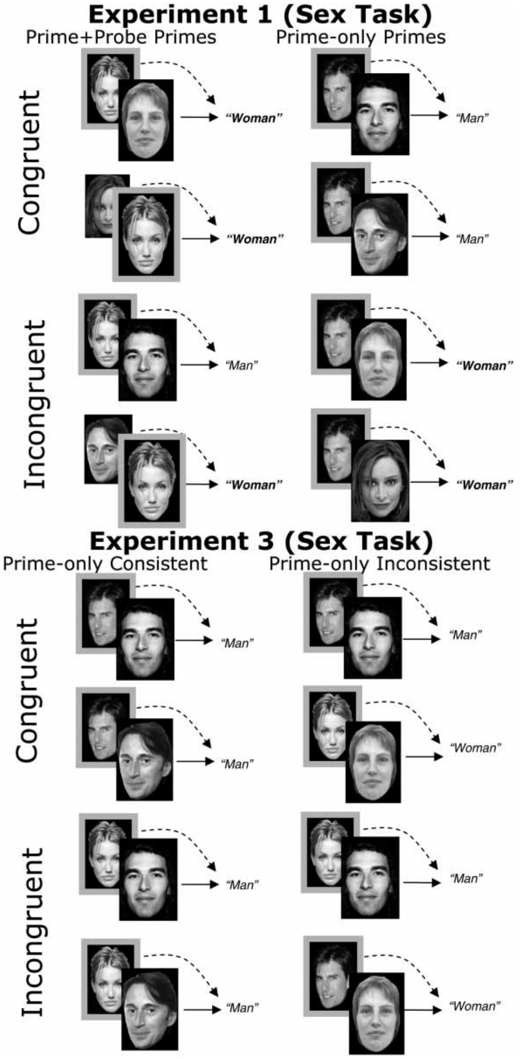

Figure 4. *Illustration of possible bias in prime–response bindings in Experiments 1 (and 2) and its manipulation in Experiment 3 (critical faces indicated by light-grey border).*

This possibility was explored in Experiment 3, which examined solely prime-only primes, but in which primes were either consistently associated with one response (consistent primes), or associated with both responses equally (inconsistent primes). One set of 64 faces was used as probes only, and another set of 16 faces was used as primes only. Prime–response mapping was either *consistent* or *inconsistent*. Prime faces with a consistent mapping always occurred within trials with the same response (as determined by the probe). Note that this consistent response could be either congruent or incongruent with the prime category (i.e., the prime category could be either congruent or incongruent with the probe category), meaning that any S–R learning to the prime should maximize priming (i.e., the difference between congruent and incongruent trials). Prime faces with inconsistent mapping had a 50% chance of being associated with a congruent or incongruent response.

Method

Participants

The same recruitment procedure was used as that in Experiment 1. There were 12 participants in the sex task (2 men and 10 women, mean age 28.7 years) and 12 in the fame task (3 men and 9 women, mean age 29.4 years).

Design

A 2 × 2 × 2 design was implemented with the factor task (fame vs. sex) varied between participants and the factors priming (congruent vs. incongruent) and prime–response consistency (consistent vs. inconsistent) varied within participants.

Materials and procedure

The set of prime faces was identical to that in Experiment 2. A total of 64 additional nonambiguous faces were selected from the same previously standardized face set (Henson et al., 2008; Kouider, Eger, Dolan, & Henson, 2009) to form the probe set. Each prime was paired evenly with the different types of probes within the limits of consistency. In the fame task, for instance, a consistent incongruent famous female prime face was evenly paired with nonfamous female and nonfamous male probe faces, and a inconsistent famous female prime face was evenly paired with all types of faces. Each of the 64 probe faces was shown once in 64 trials, each of the 16 prime faces was shown twice as a prime in 32 trials, and each of the 32 Consistency × Fame × Sex × Fame Congruency × Sex Congruency conditions was shown once in 32 trials, in a randomized order. Each prime was paired maximally once with a specific probe during the priming test. A total of 16 practice trials were given using a set of faces not used in the test. The priming test consisted of four sessions of 128 trials with breaks between sessions. Remaining materials and procedure were the same as those in Experiment 1.

Results

Prime discrimination

None of the participants reported having seen a prime during the main priming phase. Average d' in the prime discrimination phase was not significantly greater than zero, $d' = 0.09$, $SE = 0.06$, $t(23) = 1.44$, $p = .17$. The distribution of d' was near-Gaussian around zero (Figure 2d). Correlations of consistent and inconsistent priming with d' did not reach significance ($r = .26$ and .13, respectively, $p > .20$).

Congruency priming

Error rates were less than 5% on average (Table 1). No main effect or interaction reached significance in either ANOVA on RTs or arcsine-transformed accuracies, all $F(1, 23) < 1$. RT and accuracy priming were not reliable for either consistent or inconsistent prime-only primes in either task (see Table 1).

In case S–R learning took time to develop, as suggested in Experiment 1, the same ANOVA was repeated but excluding data from the first of the four sessions (see Footnote 2). Neither the main effect of priming nor any interaction with priming was significant, $Fs(1, 23) < 1.3$, $ps > .20$.

A power analysis performed for this experiment, using parameters estimated from the prime + probe conditions (collapsed across tasks) in Experiment 1 (which actually had less consistent responses for primes than the consistent prime-only trials in the present experiment), indicated that a sample size of 20 participants sufficed to detect a medium priming effect with a power of 0.9. Therefore, the absence of significant priming in this experiment is unlikely to be due to lack of power.

Discussion

Experiment 3 failed to find any evidence that responses can become associated with masked primes, in that no reliable priming was found for consistent (or inconsistent) prime-only primes. What becomes clear from Experiments 1–3 is that (a) we were unable to obtain reproducible evidence for masked priming by faces that were never presented visibly for categorization (i.e., as probes) during the experiment, and that (b) S–R binding is stronger, or is more likely, between probe faces and the response than between prime faces and the response, either because the probes are clearly visible (unlike the primes), or because it is the probe that drives the response. In Experiments 4–5, therefore, we concentrated on S–R learning to probe stimuli (i.e., prime + probe condition), as a function of response consistency and categorization consistency.

PREVIEW OF EXPERIMENTS 4 AND 5

Experiments 4 and 5 were variants of one design, in which a set of 16 faces was used as probes and as primes. The focus was on any S–R learning occurring to probes (as indexed by priming when they were reused as primes in other trials). The experiments consisted of a sequence of short blocks of 16 trials during which participants performed one task. Four different tasks alternated across blocks, which were to decide whether a probe face was "famous?", "nonfamous?", "male?", or "female?". Note that this entailed a different key mapping from that of the previous experiments, in that keys were assigned to "yes"/"no" responses (e.g., key "1": "yes", key "2": "no"), rather than to category labels, and the tasks differed in the polarity of the classification—that is, "Is the face female? (yes/no)".

In Experiment 4, we compared priming from consistent versus inconsistent probe–response mappings. In the *consistent probe response* condition, a given face only appeared in tasks requiring the same response (e.g., a famous male face that occurred as a probe in the "nonfamous?" and "female?" tasks would have a consistent response, i.e., "no" in this case); in the *inconsistent probe response* condition, a given face would equally often prompt "yes" and "no" responses (e.g., a famous male face that occurred as a probe in the "nonfamous?" and "male?" tasks; see Figure 1 for an example).

Because the results from Experiment 4 suggested that any contribution of S–R bindings at the level of a "yes" versus "no" decision or specific finger press was small, Experiment 5 examined S–R learning at the more abstract level of the classification (i.e., of a face as either "male/female" or "famous/nonfamous"). This was done by comparing fixed versus variable categorization of probes (with inconsistent response mappings).[5] Faces in the *fixed probe categorization* condition only appeared as probes in the "famous?" and "nonfamous?" tasks, or in the "male?" and "female?" tasks. Faces in the *variable probe categorization* condition could appear, for example, in the "nonfamous?" and the "female?" task (Experiment 4 used variable probe categorizations, therefore).

EXPERIMENT 4

In this experiment, the effect of practising a specific response to a probe face (e.g., a "yes" decision/right finger press) was investigated, by comparing priming from probes with consistent versus inconsistent responses across trials (analogous to the consistent vs. inconsistent responses to primes in Experiment 3). Every participant performed alternating blocks of sex and fame categorizations, whereby a given face occurred as a prime in all task blocks and appeared as a probe in two of the four tasks used, such that the response to a specific face alternated between two tasks (inconsistent response) or remained the same throughout the experiment (consistent response). Different participants were assigned to different face–condition mappings, thus ensuring that the same faces rotated between conditions across participants. We predicted greater priming from faces that had appeared as consistent versus inconsistent probes.

Method

Participants

Thirty-five participants were recruited using the same methods as those in Experiments 1–3. The data of 7 participants were replaced for low performance in the "nonfamous" task (more than 40% incorrect response trials). This higher number of replacements than in previous experiments reflected the increased difficulty of the experiment, which was demanding due to the frequent switching between different tasks. Seven participants were replaced for low recognition of at least two faces, and 1 participant was replaced

---

[5] We use "fixed" versus "variable", rather than "consistent" versus "inconsistent", in that "yes" and "no" responses are mutually exclusive (hence consistent vs. inconsistent responses), whereas "male" and "famous" are not (hence fixed vs. variable classifications).

because of a d' greater than 1. This left 8 men and 12 women (mean age 31.8 years).

Design

A 2 × 2, within-participant design was implemented with the factors probe–response consistency (consistent vs. inconsistent) and priming (congruent vs. incongruent).

Materials and apparatus

Stimuli were identical to the prime set used in Experiment 3, and the same apparatus was used. The counterbalancing was such that each face was shown as a probe in two of the four tasks throughout the experiment, and the 16 faces were evenly distributed across the tasks such that each of the 4 faces of the Fame × Sex categories were assigned to one of the Consistency × Task conditions. The assignment of faces to conditions was counterbalanced across participants using an incomplete nested block design. Probe faces were shown twice in a block, and all prime faces were shown once in each block. Each prime was paired maximally four times with a specific probe during the priming test. The conditions were completely balanced across the three last sessions (288 trials) of the test. The average lag between presentation of a face as a probe and subsequent presentation of the same face as a prime was at 16 trials.

Procedure

The procedure was identical to that of Experiments 1 and 2 apart from the following changes. Trials in the main priming phase were divided into blocks of 16, during which the task was constant. At the start of such a block, a display appeared, with, for instance "NEXT TASK: Is it a female face?", which remained until the participant pressed a key. The key word for the category, for instance "female?", was also displayed at the top of the screen throughout the block as a reminder. Eight practice trials were given with each task, using a set of faces not used again. The priming phase consisted of four sessions of 96 trials with breaks between sessions. The prime discrimination phase was identical to that in Experiment 1, except that only 64 trials were shown, 16 trials with each task.

Analysis

Data trimming was identical to that in Experiments 1–3. Priming analysis was restricted to Sessions 2–4 (i.e., excluding Session 1) because (a) participants required some practice at the task-switching procedure, (b) Sessions 2–4 were fully counterbalanced with respect to the use of each face across experimental conditions, and (c) this entailed three prior visible presentations of every face as a probe in every task, which should establish any S–R learning (see Footnote 2).

Results

Prime discrimination

Five participants reported having occasionally seen another upright face flashing up before the probe face. Mean prime discrimination d' was significantly greater than zero, $d' = 0.22$, $SE = 0.07$, $t(19) = 2.93$, $p = .009$, though there was no obvious relationship with priming (Figure 2e): Correlation between priming and d' was not significant (rs of .04 and .10, $p > .20$ in both conditions).

Congruency priming

Error rates were less than 7% on average (Table 1). Collapsing across priming condition, mean RTs to probes associated with consistent responses were reliably faster than mean RTs to probes associated with inconsistent responses, $M = 620$ ms ($SD = 79$ ms), $M = 629$ ms ($SD = 77$ ms), respectively, $t(19) = 2.09$, $SE = 3.72$, $p < .05$. This confirmed that a consistent response across tasks generally facilitated responses to probes, as expected.

The main 2 × 2 ANOVA on RTs showed a reliable main effect of priming, $F(1, 19) = 13.32$, $MSE = 182.23$, $p = .002$, and the interaction of priming with probe response consistency did not reach significance, $F(1, 19) = 0.73$, $MSE = 193.50$, $p > .20$. Priming was significant in both the consistent and inconsistent conditions (Table 1).

An additional ANOVA on arcsine-transformed accuracies revealed a trend for participants to make fewer errors in the inconsistent than in the consistent condition, $F(1, 19) = 3.21$, $MSE = 110$, $p = .09$. The reason for this unexpected trend is unclear, so it was not explored further. Neither the main effect of priming nor the interaction of priming with consistency was reliable, $Fs(1, 19) < 1.2$, $p > .10$.

Discussion

Assuming that (a) visible stimuli become associated directly with the response that they cue (i.e., when presented as probes), and (b) this response can be unconsciously triggered when those stimuli are shown as masked primes in later trials, we predicted greater priming from primes that had appeared as consistent rather than inconsistent probes. There was support for the first of these assumptions—that probe stimuli retain some association with previous responses—because overall RTs were faster for consistent than for inconsistent probes (regardless of priming). However, there was no support for the second of these assumptions, in that priming was not significantly greater for primes that had appeared as consistent than for those that had appeared as inconsistent probes (indeed, priming was numerically greater when primes had appeared as inconsistent probes). This suggests that priming was not driven primarily by automatic retrieval of a yes/no decision or a specific motor action.

One possibility is that the priming was driven by stimuli becoming bound to multiple, more abstract response representations (Abrams et al., 2002; Horner & Henson, 2009; Koch & Allport, 2006; Waszak & Hommel, 2007; Waszak, Hommel, & Allport, 2003). For instance, a face might become associated with both a "famous" label from its appearance in a "famous?" task, and a "woman" label from its appearance in a "female?" task: what Horner and Henson (2009) called "classifications". When that face then appears as a prime in a given task, automatic retrieval of both of these categorizations (or even only the task-relevant one) could then cause priming. Note that retrieval of a "famous" classification from a prime appearing in a "nonfamous?" task would still allow a decision to be made without semantic analysis of that face, in that the "famous" label can quickly be mapped to a "no" decision (see General Discussion). This possibility that stimuli were associated with abstract classifications was explored in Experiment 5, where probe faces were classified according to either one category (e.g., fame) or both (fame and sex).

Finally, we note that the mean d' was greater than zero in this experiment, unlike in Experiments 1–3. Thus one should keep in mind that prime presentations may not have been subliminal in the present experiment (at least for some participants and/or trials, for whatever reason). Nonetheless, there was no obvious relationship between d' and the amount of priming shown (Figure 2e), suggesting that prime visibility was not important for the current priming effects.

EXPERIMENT 5

Experiment 4 demonstrated priming even from prime faces that were not associated with a consistent yes/no response. In Experiment 4, faces could appear as probes in one of the two fame tasks (i.e., "famous?" or "nonfamous?") and in one of the two sex tasks (i.e., "male?" or "female?"). We termed this situation *variable probe categorization*. In order to address the hypothesis that priming is due to learning of an abstract classification rather than yes/no decision, Experiment 5 compared the variable probe categorization condition with a *fixed probe categorization* condition, where a probe face was shown in one type of categorization only. Hence, probe faces in the fixed probe categorization condition appeared either in the "famous?" and "nonfamous?" tasks, or in the "male?" and "female?" tasks (while appearing as a prime in all tasks). In other words, faces in the fixed probe categorization condition were effectively "prime-only" primes with respect to the nonprobed categorization (even though not prime-only primes with respect to the whole session). Note that in all conditions, probe responses were necessarily inconsistent. While we expected to replicate priming in the variable probe categorization conditions of Experiment 4, we expected to find smaller priming in the fixed probe categorization condition of Experiment 5. This is because in this condition, probe faces were categorized according to one category only, which should reduce overall priming when measured across both categorizations. As a concrete example, we expected that priming from a nonfamous male face would be smaller when it had appeared as a probe only in the "famous?" and "nonfamous?" (fixed classification) tasks than when it appeared in the "male?" and "nonfamous?" (variable classification) tasks, because both categorizations were learned in the latter case but not in the former case (see Figure 1). In other words, priming was not expected for those trials in the fixed probe categorization condition where the task performed (e.g., "male?") entailed a categorization different from when the prime stimulus had appeared as a probe (e.g., "famous?" and "nonfamous?" tasks).

Method

Participants

Twenty-five participants were recruited using the same methods as those in Experiments 1–3. The data of 1 participant were replaced for low performance in the "nonfamous" task, 3 for low recognition of at least two faces, and 1 because of an absolute d' greater than 1 in the prime discrimination phase. This left 8 men and 12 women (mean age 32.8 years).

Design

A 2×2, within-participant design was implemented with the factors probe categorization (variable vs. fixed) and priming (congruent vs. incongruent).

Materials and apparatus

Materials and apparatus were the same as those in Experiment 4. The only difference in design was that some of the faces (the fixed task faces) occurred as probes in either both fame tasks or both sex tasks, but never in a fame task and a sex task. Because the sequence of tasks alternated between sex and fame task blocks, fixed task probe faces were presented in every second block (e.g., in all male and female tasks), whereas variable task probe faces were presented either one or three blocks apart (e.g., in all male and famous tasks). The average lag between presentation of a face as a probe and subsequent presentation of the same face as a prime was 12 and 16 trials for the fixed-categorization and variable-categorization conditions, respectively.

Procedure and analysis

The procedure and analysis were identical to those in Experiment 4.

Results

Prime discrimination

Seven participants reported having occasionally seen another upright face flashing up before the probe face. Average d' was significantly greater than zero, $d' = 0.18$, $SE = 0.08$, $t(19) = 2.34$, $p = .03$, but there was no obvious relationship with the amount of priming (Figure 2f). Correlation of priming with d' was not significantly greater than zero (rs at $-.41$ and $.19$, $p > .05$ in both conditions; see General Discussion for further mention).

Congruency priming

Error rates were less than 9% on average (Table 1). The ANOVA on RTs showed a reliable main effect of priming, $F(1, 19) = 11.67$, $MSE = 256.93$, $p = .003$. The nonsignificant two-way interaction between probe categorization and priming, $F(1, 19) = 0.24$, $MSE = 216.92$, $p > .20$, did not confirm any difference between priming for fixed and for variable probe categorization primes. Priming was reliable for both variable and fixed probe categorization primes (see Table 1).

As a further check, priming for the fixed probe condition was split according to trials in which the prime occurred in the *same task* as that in which it had appeared as a probe (e.g., for faces that occurred as probes in the "famous?/nonfamous?" tasks and were now appearing as primes in the same tasks), and trials in which the prime occurred in a *different task* from that in which it appeared as a probe (e.g., for faces that occurred as probes in the "famous?/nonfamous?" tasks and were now appearing as primes in "male?/female?" tasks; see also Figure 1d). Priming was reliable for both types of trial: same task, $M = 13.7$ ms, $SD = 27.1$ ms, $t(19) = 2.26$, $p = .049$, and different task, $M = 14.2$ ms, $SD = 36.6$ ms, $t(19) = 1.74$, $p = .018$.

An additional ANOVA on arcsine-transformed accuracies revealed that neither the main effects, all $F(1, 19) < 1.8$, $p > .20$, nor the interaction of priming and probe categorization were reliable, $F(1, 19) = 2.15$, $MSE = 90$, $p = .16$.

Discussion

The data for the variable task categorization condition replicated the reliable priming found for inconsistent responses in Experiment 4, reinforcing the minimal role of S–R learning at the level of the "yes/no" decision or lower (e.g., finger press). However, our prediction that the data for the fixed task categorization condition would show reduced priming was not supported. This prediction was based on the assumption that, if priming is caused by retrieval of classifications previously associated with a face, then such priming should not occur when that face appeared as a prime in a task that entailed a different classification. This should reduce overall priming in the fixed task categorization, and yet this priming was actually numerically larger than that in the variable probe categorization condition. Indeed, priming was reliable even for that subset of trials where a prime face was categorized according to a category different from that for which it had been categorized as a probe. We now consider how these results, together with those from Experiments 1–4, might be explained.

GENERAL DISCUSSION

Congruency priming by masked primes has frequently been used to make inferences about unconscious semantic processing. In Experiments 1–3, we confirmed the robustness of masked congruency priming of fame and sex categorizations of faces that were used as primes and probes across trials (prime + probe primes), which contrasted with no convincing evidence for priming by faces never used as probes within the same session (prime-only primes). There was no evidence that the amount of this prime + probe priming was related to the conscious ability to categorize the masked prime, which was generally close to chance, suggesting that the priming was subliminal. This prime + probe priming was found for both famous faces (confirmed as known by participants) and nonfamous faces (unseen prior to the experiment and confirmed as unknown by participants), suggesting that it did not depend on preexisting semantic representations, as assumed by abstractionist models like the IAC model of face recognition (Burton et al., 1999). Furthermore, repeated, attended processing of a face as a probe in a previous session (in Experiment 2) was not sufficient to cause priming when that face was used as a prime-only prime in a different task, suggesting that masked congruency priming was not caused by the gradual formation of new perceptual representations either (Martin & Greer, 2011).

These findings suggested to us that the masked congruency priming entailed some form of S–R learning across trials in which a face appears as a probe, such as the formation of episodic associations or bindings between that face and some type of response code (Abrams & Greenwald, 2000; Damian, 2001). Any such S–R learning would seem surprisingly flexible, however, in that priming was found even when stimuli appeared as probes that were equally often associated with a yes and a no decision (the inconsistent response condition of Experiment 4), and even when a stimulus appeared as probe in one categorization but as a prime in a different categorization (when alternating between categorizations in the fixed categorization condition of Experiment 5). Below we propose a generalized S–R account of the present data, in which stimuli become simultaneously bound to multiple "classifications" (Horner & Henson, 2009), even if those classifications are generated covertly by virtue of a switching task context. We then compare this proposal with other existing theories and findings.

A generalized form of stimulus–response learning?

The fact that reliable priming was found even when probe stimuli were not consistently paired with a specific yes/no decision (or specific finger press) in Experiment 4 suggests that such learning occurs at an abstract level, for example by stimuli becoming bound to "classifications". This type of abstract response code was invoked by Horner and Henson (2009) to explain why positive repetition priming was found even when the polarity of the task was reversed between initial and repeated presentation of an (unmasked) stimulus (and even under conditions when alternative causes, such as perceptual or conceptual facilitation, were unlikely). So, for example, the association of a stimulus with a "bigger" classification when it was initially presented (primed) in a "bigger than a shoebox?" task was postulated as causing faster responses (relative to unprimed stimuli) when that stimulus was repeated in a "smaller than a shoebox?" task, via rapid retrieval of the "bigger" classification. This retrieved classification then allows a "no" decision to be made without requiring any repeated conceptual processing of the stimulus. Similarly, in the present masked paradigm, the rapid retrieval of a "woman" classification that is cued by a masked face prime, owing to its prior appearance as a probe in a "female?" task, could initiate generation of a "no" decision in a "male?" task, even before such generation is initiated by the probe. This could speed up decisions to a congruent probe (e.g., accelerating accumulation of evidence for one decision) and/or slow down decisions to an incongruent probe (e.g., decelerating evidence accumulation owing to interference between two incompatible decisions), leading to RT priming. This would explain why priming was unaffected by the consistency of the yes/no decision associated with probes in Experiment 4.[6]

Furthermore, the finding that priming was reliable even when a face appeared as a probe in both fame and sex categorizations (the variable categorization conditions of Experiments 4–5) suggests that a stimulus can simultaneously become bound to multiple different classifications (e.g., "woman" and "familiar"). This would then cause priming when that stimulus occurs as a probe in any of the present "famous?", "nonfamous?", "male?", or "female?" tasks: Even if both classifications were cued by the masked prime, selection of the one relevant to the current task could explain the congruency priming along the lines above.

The most puzzling finding for an S–R learning account was that priming was still reliable for those subset of trials in the fixed categorization of Experiment 5, in which a stimulus occurred as a prime in one task block (e.g., "female?") for which the categorization was different from that in another task block in which it appeared as a probe (e.g., "famous?"). This finding also contrasts with the results of Experiment 2, in which no evidence was found for priming by primes that were previously used in another task (i.e., no significant prime-only priming, not even for the fame task in Session 3). One possible explanation is that only when participants have practised several tasks (even if on different stimuli) do they begin to covertly categorize all stimuli according to those tasks. Indeed, the design in Experiments 4 and 5, whereby participants switched frequently between task blocks, might have fostered such covert classification, analogous to the frequent switching between tasks that has been shown to activate both the relevant and the irrelevant response rule for masked primes (Kiesel et al., 2007). In this way, a probe stimulus may become bound to both "familiar" and "woman" classifications, even if some of those classifications are only generated covertly. Such covert "stimulus-classification" bindings would facilitate performance when a stimulus appears as a prime in either task.

Some support for this possibility again comes from the across-trial repetition priming domain (Dennis, Carder, & Perfect, 2008). These authors found that RTs/errors to *unprimed* stimuli were

---

[6] In other experiments, Horner and Henson (2009) also found evidence for binding of stimuli with both decisions (yes/no) and actions (e.g., left vs. right finger press)—that is, evidence for simultaneous S–R bindings at multiple levels of response code. It is not clear why there was no evidence for such decision/action bindings in addition to classification bindings in the present paradigm, but we suspect that the frequent task switching (and hence decision/action switching) in Experiments 4–5, which was not the case in the single repetition priming of Horner and Henson, caused participants to devalue, or reduce attention to, any retrieval of responses at the level of decisions or actions.

affected by the presence or absence of primed stimuli that had been previously categorized in a different task. This is consistent with our proposal that participants have a tendency to "run off" the procedures required for all tasks they have practised previously (even on new, i.e., unprimed, stimuli in the Dennis et al. study). Even though the currently relevant task (e.g., our sex task) must dominate to produce the actual response (in order to maintain correct performance), the "partial activation" of the other task (e.g., our fame task) may be sufficient to allow binding of the stimulus to the implicitly generated classification (e.g., "familiar"), in addition to the explicitly generated classification (e.g., "woman").

Relation to event files

Our generalized S–R learning proposal resembles some of the ideas behind the structured "event file" concept of Hommel (1998): Not only might there be binding of multiple relevant and irrelevant stimulus features on one side of an event file, but also parallel binding to multiple classifications on the other side. The paradigms used to investigate event files have focused on the interference effects arising from irrelevant features/responses. So, for example, bindings between stimuli and task codes can cause stimulus-driven interference effects in task-switching paradigms (e.g., Waszak & Hommel, 2007). The longer RTs to incongruent trials in the prime + probe than prime-only conditions in both tasks of Experiments 1–2 might be consistent with some form of interference from stimuli that had become bound to such task codes (i.e., prime + probe but not prime-only stimuli). However, it is unclear why retrieval of a generic task label from a masked prime would slow down RTs more for incongruent than for congruent trials (unless that label were a specific classification, as in our generalized S–R learning hypothesis above). Task-level interference effects would also seem to predict less priming in our variable categorization than in our fixed categorization conditions, no support for which was found in Experiment 5. Nonetheless, the numerical pattern of slower RTs (and increased errors) for incongruent prime + probe trials than incongruent prime-only trials (but little difference between congruent prime + probe trials and congruent prime-only trials) does suggest that retrieval of classifications might interfere with incongruent decisions to a greater extent than it facilitates congruent decisions. Future testing of this possibility could include the addition of unrelated, or no-prime, trials.

Relation to action triggers

Our generalized S–R learning proposal can also be discussed in relation to the "action-trigger hypothesis" of Kunde et al. (2003). This hypothesis was initially used to explain masked congruency priming in a number comparison task by numbers that never appeared as probes, but that lay within the range of numbers relevant to the task. According to this hypothesis, some stimuli can trigger actions automatically, if they fulfil the release preconditions for those actions. For instance, if participants expect to see numbers in the range from 1 to 9 in a "smaller/larger than 5" categorization task, the stimulus preconditions for giving the "smaller than 5" response are defined as "seeing one of the numbers 1, 2, 3, or 4". Indeed, participants appeared to use such a definition in Kunde et al.'s study, as there was significant priming by novel primes that came from the expected range. Likewise, once they were familiar with the 16 faces in Experiments 4–5, our participants might have been able to set action trigger preconditions that they applied to any of the faces in any of the tasks. For example, participants may have extracted certain facial features that correlate with certain categorizations (e.g., broad jaw for the male faces; bright teeth for famous faces) and used these as preconditions (or "feature–response" bindings).

One difference between our generalized S–R hypothesis and the action-trigger hypothesis concerns the predictions for novel (or never categorized) stimuli. The action-trigger account was proposed specifically to explain priming from novel primes (provided those primes were within the set of stimuli on which participants might expect to perform a given task), whereas our "generalized S–R hypothesis" has no mechanism for binding classifications to previously unpresented stimuli. The data from the (restricted) domain of numbers clearly support the "action-trigger" account. Does the same apply to the (unrestricted) domain of faces? The prime-only conditions of Experiments 1–3 are relevant in this context. If stimuli must be perceived clearly in order to become bound to a classification, then the "generalized S–R hypothesis" predicts no priming in this prime-only condition. However, assuming that participants can extract certain facial features that can be used for the relevant categorization (action triggers), it is possible that action triggers for those features could be established that are sufficient to cause priming even from stimuli never seen as probes. We found no evidence to support such prime-only priming, which is why we favour our generalized S–R learning account of the present data.

Perceptual learning

In Experiment 2, we tested an alternative, "perceptual learning" explanation of our pattern of prime + probe priming, but not prime-only priming. According to this explanation, participants do not learn an S–R binding, but simply learn about a stimulus itself ("S learning").[7] In other words, conscious awareness of (and possibly attention to) a stimulus may be sufficient to learn something about that stimulus, such that when it recurs as a masked prime, that stimulus is perceived more efficiently, facilitating task-related processing of the prime, which helps the processing of subsequent congruent probes and/or hinders the processing of subsequent incongruent probes, causing priming. This perceptual learning would be minimal for stimuli only ever used as masked primes in the present experiments (given the evidence for minimal awareness of such primes), but would occur for stimuli repeatedly used as probes as well as primes, regardless of the specific responses associated with them when they appeared as probes. Such stimulus learning might be more important for complex stimuli such as faces (particularly unfamiliar faces) than the words or numbers often used to study masked semantic priming.

Experiment 2 found no support for this perceptual learning account, because there was no evidence for priming from stimuli that only occurred as primes in one session ("prime-only" primes within that session) even though they had occurred multiple times as probes in a previous session (but using a different categorization). Nonetheless it remains possible that perceptual learning is only short-lived, such that any perceptual learning of the probe stimuli in the first or second sessions of Experiment 2 had "worn off" by the time that those stimuli appeared as prime-only primes within the second and third sessions, respectively. Thus perceptual learning might have caused prime + probe priming in Experiments 4–5, where the lag between stimuli appearing as probes and as primes was relatively short, but not in Experiment 2, where that lag was longer. This could be tested in further experiments. Nonetheless, we think this perceptual learning account is generally unlikely, given other previous demonstrations that prior exposure to stimuli is not sufficient for them to cause priming when later used as masked primes (Damian, 2001).

Does the present priming reflect subliminal processing?

We believe the present priming effects are likely to be subliminal because participants' ability to categorize the primes was consistently low across all five experiments. While it is always difficult to prove that each participant's true d' was zero (e.g., Rouder, Morey, Speckman, & Pratte, 2007), the d' values in the present experiments were comparable, if not lower, than in previous claims of subliminal priming. In Experiments 1–3, d' was not significantly different from zero, and in Experiments 4–5, average d' was lower than 0.25, which is still very low discrimination. It is also important to note that the objective

---

[7] It is of course also possible that participants learned something about responses ("R learning")—for example, forming stronger associations between a decision (e.g., "yes") and an action (e.g., left index finger press). However, such learning would facilitate RTs to both congruent and incongruent trials, so not produce priming (given that congruency depends on the relationship between prime and probe stimulus, not the response per se).

d' measure used in this study is a conservative assessment of conscious perception of primes, for the following reasons: (a) Prime perception was measured in a separate task, where participants were given time to concentrate on the prime and ignore the probe, in contrast to the priming test, where participants' attention was on the probe instead—so it is reasonable to assume that the true d' was lower than the d' estimated in the discrimination task; (b) the variance of d' was low, because participants with high d' were excluded from data analysis; and (c) the assumption that any d' measure is an exclusive measure of conscious perception is theoretically difficult to sustain—more probably, d' is the result of both conscious and unconscious processes of perception (cf. Joordens & Merikle, 1993; Reingold & Merikle, 1988).

Furthermore, even if primes on some trials could be perceived well enough to be categorized, the scatter plots in Figure 2 show no indication of a relationship across participants between their d' score and the amount of priming that they exhibited. Based on these considerations, we think that conscious perception of primes was minimal and did not contribute to the patterns of priming found.

Do the present data question subliminal semantic processing?

While much of the evidence for unconscious semantic processing has come from masked congruency priming experiments like the present one, it is also important to note that we are not claiming that unconscious semantic processing (here of faces) does not occur. Indeed, as noted in the introduction, several recent studies have reported masked semantic priming under conditions in which S–R learning is unlikely. In other words, despite our lack of reliable priming with prime-only primes, we are not claiming that S–R learning is a sufficient explanation of all examples of masked semantic priming, or even just a sufficient explanation of masked congruency priming of faces. Finkbeiner and Palermo (2009), for example, recently reported priming of sex judgements by masked primes that were also never presented as probes (i.e., prime-only primes), suggesting that some subliminal categorization of faces is possible. We do not yet know the conditions under which the size of such subliminal congruency priming becomes detectable (with statistical power comparable to that of the present experiments), but do note that Finkbeiner and Palermo used only 2 faces as primes and only 10 faces as probes (compared to the minimum of 8 primes and 16 to 64 probes in the present experiments); it is possible that such a small prime set increased the likelihood of establishing feature–response bindings, analogous to the action-trigger account described earlier. Nonetheless, the main conclusion from the present experiments relates to the robust subliminal congruency priming that we repeatedly found when primes *did* appear as probes, even when those occurrences were not confined to a specific response nor specific categorization.

Finally, it is important to distinguish the present masked congruency priming from other types of masked semantic priming in which the prime and probe have a preexisting semantic relationship. The prime and probe faces used here had no preexisting relationship (they were randomly assigned to congruent or incongruent trials), but when a prime and probe are already related, either categorically or associatively, priming on a fame task can be reliably found, as described in the introduction (e.g., Stone & Valentine, 2006). Indeed, Wiese et al. (2011) found reliable masked priming in a fame task for prime-only primes that were associated with the same classification response in both the primed and unprimed case (e.g., the face of Angelina Jolie preceded by the masked name of Brad Pitt—the primed case—relative to the face of Angelina Jolie preceded by the masked name of Boris Yeltsin—the unprimed case). This type of priming would seem difficult to explain in terms of any type of S–R learning. More generally, the strength of subliminal semantic priming may relate to the structure of categories, such as whether members have a family resemblance on the dimension required by the categorization task (Quinn & Kinoshita, 2007).

CONCLUSIONS

Using masked presentations of faces in sex or fame categorization tasks, we found clear evidence that congruency priming of reaction times was due to some form of stimulus–response learning, in that it was only found for prime faces that also appeared as probes in other, intermixed trials. Such priming was found even if those prime and probe trials entailed a different categorization. The pattern of results was not compatible with any of the main theories of face or response priming. We suggest that stimuli become bound to multiple classifications, even if those classifications are generated covertly by virtue of a repeatedly switching task context. While future experiments are needed to test further this "generalized S–R learning" hypothesis and relate it to other theories of event records and action triggers, the potential for such abstract and flexible learning is important because it further cautions against the assumption that masked congruency priming implies subliminal semantic processing.

Original manuscript received 4 October 2010
Accepted revision received 1 April 2011
First published online 19 July 2011

REFERENCES

Abrams, R. L., & Greenwald, A. G. (2000). Parts outweigh the whole (word) in unconsious analysis of meaning. *Psychological Science*, *11*(2), 118–124.

Abrams, R. L., Klinger, M. R., & Greenwald, A. G. (2002). Subliminal words activate semantic categories (not automated motor responses). *Psychonomic Bulletin & Review*, *9*(1), 100–106.

Bruce, V., & Young, A. W. (1986). Understanding face recognition. *British Journal of Psychology*, *77,* 305–327.

Burton, A. M., Bruce, V., & Hancock, P. J. B. (1999). From pixels to people: A model of familiar face recognition. *Cognitive Science*, *23,* 1–31.

Burton, A. M., Bruce, V., & Johnston, R. A. (1990). Understanding face recognition with an interactive activation model. *British Journal of Psychology*, *81,* 361–380.

Burton, Bruce, & Johnston, 1999

Damian, M. F. (2001). Congruity effects evoked by subliminally presented primes: Automaticity rather than semantic processing. *Journal of Experimental Psychology: Human Perception and Performance*, *27* (1), 154–165.

Dennis, I., Carder, H. P., & Perfect, T. J. (2008, April). *A response congruence effect in cross-task repetition priming*. Paper presented at the talk presented to the Experimental Psychology Society, Cambridge, UK: University of Cambridge.

Dosher, B. A. (1998). The response window regression method—Some problematic assumptions: Comment on Draine & Greenwald (1998). *Journal of Experimental Psychology: General*, *127*(3), 311–317.

Draine, S. C., & Greenwald, A. G. (1998). Replicable unconscious semantic priming. *Journal of Experimental Psychology: General*, *127*(3), 286–303.

Ellis, A. W., Young, A. W., & Flude, B. M. (1990). Repetition priming and face processing—Priming occurs within the system that responds to the identity of a face. *Quarterly Journal of Experimental Psychology, Section A: Human Experimental Psychology*, *42*(3), 495–512.

Enns, J. T., & Oriet, C. (2007). Visual similarity in masking and priming: The critical role of task relevance. *Advances in Cognitive Psychology*, *3*(1–2), 211–226.

Finkbeiner, M., & Palermo, R. (2009). The role of spatial attention in nonconscious processing: A comparison of face and non-face stimuli. *Psychological Science*, *20,* 42–51.

Goshen-Gottstein, Y., & Ganel, T. (2000). Repetition priming for familiar and unfamiliar faces in a sex-judgment task: Evidence for a common route for the processing of sex and identity. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *26,* 1198–1214.

Greenwald, A. G., Abrams, R. L., Naccache, L., & Dehaene, S. (2003). Long-term semantic memory versus contextual memory in unconscious number processing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *29,* 235–247.

Greenwald, A. G., & Draine, S. C. (1998). Distinguishing unconscious from conscious cognition—reasonable assumptions and replicable findings: Reply to Merikle and Reingold (1998) and Dosher (1998). *Journal of Experimental Psychology: General*, *127*(3), 320–324.

Greenwald, A. G., Klinger, M. R., & Schuh, E. S. (1995). Activation by marginally perceptible ("subliminal") stimuli: Dissociation of unconscious from conscious cognition. *Journal of Experimental Psychology: General*, *124*(1), 22–42.

Henson, R. N., Mouchlianitis, E., Matthews, W. J., & Kouider, S. (2008). Electrophysiological correlates of masked face priming. *NeuroImage*, *40,* 884–895.

Holender, D. (1986). Semantic activation without conscious identification in dichotic listening, parafoveal vision, and visual masking: A survey and appraisal. *Behavioral and Brain Sciences*, *9,* 1–66.

Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus–response episodes. *Visual Cognition*, *5*(1/2), 183–216.

Hommel, B., & Colzato, L. S. (2009). When an object is more than a binding of its features: Evidence for two mechanisms of visual feature integration. *Visual Cognition*, *17*, 120–140.

Horner, A. J., & Henson, R. N. (2009). Bindings between stimuli and multiple response codes dominate long-lag repetition priming in speeded classification tasks. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *35*(3), 757–779.

Joordens, S., & Merikle, P. M. (1993). Independence or redundancy? Two models of conscious and unconscious influences. *Journal of Experimental Psychology: General*, *122,* 462–467.

Kiesel, A., Kunde, W., & Hoffmann, J. (2007). Unconscious priming according to multiple S–R rules. *Cognition*, *104,* 89–105.

Kiesel, A., Wagener, A., Kunde, W., Hoffmann, J., Fallgatter, A. J., & Stöcker, C. (2006). Unconscious manipulation of free choice in humans. *Consciousness and Cognition*, *15,* 397–408.

Kinoshita, S., & Kaplan, L. (2008). Priming of abstract letter identities in the letter match task. *Quarterly Journal of Experimental Psychology*, *61,* 1873–1885.

Kinoshita, S., & Norris, D. (2009). Transposed-letter priming of pre-lexical orthographic representations. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *35,* 1–18.

Klauer, K. C., Draine, S. C., & Greenwald, A. G. (1998). An unbiased errors-in-variables approach to detecting unconscious cognition. *British Journal of Mathematical and Statistical Psychology*, *51,* 253–267.

Klauer, K. C., Eder, A. B., Greenwald, A. G., & Abrams, R. L. (2007). Priming of semantic classifications by novel subliminal prime words. *Consciousness and Cognition*, *16,* 63–83.

Klauer, K. C., Greenwald, A. G., & Draine, S. C. (1998). Correcting for measurement error in detecting unconscious cognition: Comment on Draine & Greenwald (1998). *Journal of Experimental Psychology: General*, *127*(3), 318–319.

Klauer, K. C., Musch, J., & Eder, A. B. (2005). Priming of semantic classifications: Late and response related, or earlier and more central? *Psychonomic Bulletin & Review*, *12*(5), 897–903.

Koch, I., & Allport, A. (2006). Cue-based preparation and stimulus-based priming of tasks in task switching. *Memory & Cognition*, *34*(2), 433–444.

Kouider, S., & Dehaene, S. (2007). Levels of processing during non-conscious perception: A critical review of visual masking. *Philosophical Transactions of the Royal Society*, *362,* 857–875.

Kouider, Eger, Dolan, & Henson, 2008

Kouider, S., Eger, E., Dolan, R., & Henson, R. N. (2009). Activity in face-responsive brain regions is modulated by invisible, attended faces: Evidence from masked priming. *Cerebral Cortex*, *19,* 13–23.

Kunde, W., Kiesel, A., & Hoffmann, J. (2003). Conscious control over the content of unconscious cognition. *Cognition*, *88*(2), 223–242.

Logan, G. D. (1990). Repetition priming and automaticity: Common underlying mechanisms? *Cognitive Psychology*, *22,* 1–35.

Marcel, A. L. (1983). Conscious and unconscious perception: Experiments on visual masking and word recognition. *Cognitive Psychology*, *15,* 197–237.

Martin, D., & Greer, J. (2011). Getting to know you: From view-dependent to view-invariant repetition priming for unfamiliar faces. *The Quarterly Journal of Experimental Psychology*, *64,* 217–223.

Merikle, P. M. (1982). Unconscious perception revisited. *Perception and Psychophysics*, *31*(3), 298–301.

Merikle, P. M., & Daneman, M. (1998). Psychological investigations of unconsious perception. *Journal of Consciousness Studies*, *5*(1), 18.

Merikle, P. M., & Reingold, E. M. (1998). On demonstrating unconsious perception: Comment on Draine & Greenwald (1998). *Journal of Experimental Psychology: General*, *127*(3), 304–310.

Naccache, L., & Dehaene, S. (2001a). The priming method: Imaging unconscious repetition priming reveals an abstract representation of number in the parietal lobes. *Cerebral Cortex*, *11,* 966–974.

Naccache, L., & Dehaene, S. (2001b). Unconscious semantic priming extends to novel unseen stimuli. *Cognition*, *80,* 223–237.

Quinn, W. M., & Kinoshita, S. (2007). Congruence effect in semantic categorization with masked primes with narrow and broad categories. *Journal of Memory and Language*, *58,* 286–306.

Reingold, E. M., & Merikle, P. M. (1988). Using direct and indirect measures to study perception without awareness. *Perception and Psychophysics*, *44,* 563–575.

Rouder, J. N., Morey, R. D., Speckman, P. L., & Pratte, M. S. (2007). Detecting chance: A solution to the null sensitivity problem in subliminal priming. *Psychonomic Bulletin & Review*, *14*(4), 597–605.

Stone, A. (2008). Categorical priming of famous person recognition: A hitherto overlooked methodological factor can resolve a long-standing debate. *Cognition*, *108*(3), 874–880.

Stone, A., & Valentine, T. (2006). The categorical structure of knowledge for famous people (and a novel application of centre-surround theory). *Cognition*, *104*(3), 535–564.

Tenpenny, P. L. (1995). Abstractionist versus episodic theories of repetition priming and word identification. *Psychonomic Bulletin & Review*, *2,* 339–363.

Van den Bussche, E., Van Den Noortgate, W., & Reynvoet, B. (2009). Mechanisms of masked priming: A meta-analysis. *Psychological Bulletin*, *135,* 452–477.

Waszak, F., & Hommel, B. (2007). The costs and benefits of cross-task priming. *Memory & Cognition*, *35*(5), 1175–1186.

Waszak, F., Hommel, B., & Allport, A. (2003). Task switching and long-term priming: Role of episodic stimulus-task bindings in task-shift costs. *Cognitive Psychology*, *46,* 361–433.

Wiese, H., Henson, R. N., & Schweinberger, S. R. (2011). *Determinants of non-strategic associative and categorical priming in person recognition*. Manuscript submitted for publication.

Wiese, H., & Schweinberger, S. R. (2008). Event-related potentials indicate different processes to mediate categorical and associative priming in person recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *34*(5), 1246–1263.