



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/120794/>

Version: Accepted Version

Article:

Hanczakowski, M., Zawadzka, K. and Higham, P.A. (2014) The dud-alternative effect in memory for associations: Putting confidence into local context. *Psychonomic Bulletin & Review*, 21 (2). pp. 543-548. ISSN: 1069-9384

<https://doi.org/10.3758/s13423-013-0497-x>

Reuse

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

The Dud-Alternative Effect in Memory for Associations: Putting Confidence into Local
Context

Maciej Hanczakowski¹, Katarzyna Zawadzka², and Philip A. Higham²

¹Cardiff University, UK

²University of Southampton, UK

Author Notes

Maciej Hanczakowski, School of Psychology, Cardiff University; Katarzyna Zawadzka, Psychology, University of Southampton; Philip A. Higham, Psychology, University of Southampton.

The authors would like to thank Dylan Jones for helpful comments while developing this project.

Correspondence concerning this article should be directed to Maciej Hanczakowski, School of Psychology, Cardiff University 64 Park Place, Cardiff, CF10 3AT, United Kingdom, email: HanczakowskiM@cardiff.ac.uk, ph: +44 (0)299 2087 5030.

Article published in *Psychonomic Bulletin & Review*.

The final publication is available at Springer via <http://dx.doi.org/10.3758/s13423-013-0497-x>

Abstract

When participants are asked to provide confidence judgments for each provided alternative in a multiple-choice memory task, such judgments are inflated if assessed alternatives are accompanied by an implausible (dud) alternative. This finding, termed the dud-alternative effect, has been recently documented in a memory setting with a line-up procedure (Charman, Wells, & Joy, 2011). In the present study we develop a novel paradigm to investigate the dud-alternative effect in memory. The paradigm utilizes a multiple-choice associative recognition task in which dud alternatives can be rejected on the basis of their unfamiliarity. In two experiments we demonstrate a reliable dud-alternative effect with our novel procedure. The results demonstrate that the dud-alternative effect in episodic memory is not limited to tasks based on perceptual factors but is a general phenomenon concerning confidence judgments.

Keywords: Duds, Confidence, Recognition, Metacognition

The Dud-Alternative Effect in Memory for Associations: Putting Confidence into Local Context.

When faced with a task of remembering, people often consider evidence supporting multiple alternatives. How sure is a person that she parked her car today at the entrance and not closer to the parking gate? How sure is she that she saw the person number five robbing the old lady and not the person number four? It seems logical that when assessing confidence for multiple alternatives, confidence is reduced if there is strong evidence supporting each alternative (Clark, 1997). But what happens with confidence when there is very little evidence in support of one (or more) of the alternatives? The most intuitive prediction would be that nothing much happens – the implausible alternative is simply eliminated and confidence depends on evidence supporting the remaining alternatives. However, some recent evidence indicates that this intuition is incorrect and such alternatives do have an important effect on confidence, an effect that is examined in the present study.

Windschitl and Chambers (2004) introduced the *dud-alternative effect* whereby the inclusion of implausible (dud) alternatives in a multiple-choice question inflates the judged likelihood of other, plausible alternatives relative to a case in which duds are absent. For example, in the dud-absent condition of their Experiment 1, participants were asked which of the following alternatives was selected in a recent survey of American 7-10 year-olds to be their favorite food for dinner: Pizza or Hamburger. In the dud-present condition, the question included two additional dud alternatives: Eggplant Parmesan and Grilled Fish. Windschitl and Chambers found that likelihood judgments for the plausible alternatives (Pizza and Hamburger) were higher in the dud condition compared to the no-dud condition. The effect generalized across a number of different tasks and conditions. For example, the judged likelihood of winning a raffle was increased if some players (duds) were added that had very few tickets. Also, the effect occurred regardless of whether the task directed participants'

attention to a particular alternative (e.g., “How likely do you think it is that *pizza* was the most frequently selected option?”) or if ratings were required for all alternatives. The finding is non-normative because adding more alternatives should either reduce (if they are plausible) or leave unchanged (if they are completely implausible) the judged likelihood of the original, plausible alternatives, but they should never increase it.

Windschitl and Chambers (2004) accounted for their finding with a *contrast hypothesis*. According to it, participants judge the likelihood of a given alternative by comparing it (explicitly or implicitly) against the likelihood of the other alternatives. By this hypothesis, the inclusion of duds increases the number of comparisons that strongly favor the focal alternative currently being judged, which, in turn, increases its judged likelihood.

Although not the focus of Windschitl and Chambers (2004) research, the authors nevertheless predicted that a similar effect would occur with episodic memory judgments. Recently, Charman, Wells, and Joy (2011) confirmed this prediction using a line-up identification procedure. They found that including duds in the line-ups (i.e., foils that were highly perceptually dissimilar to the actual perpetrator) increased confidence that the remaining, non-dud foils matched the perpetrator. Thus, consistent with Windschitl and Chambers’ predictions, the study by Charman et al. indicated that dud alternatives can increase confidence for non-dud alternatives in episodic memory tasks as well as other likelihood judgment tasks.

However, Charman et al. (2011) suggested that the dud-alternative effect in line-ups might stem from a different mechanism than the analogous effect with non-episodic likelihood judgments. Specifically, they argued against the contrast hypothesis and in favor of a *perceptual similarity* account. This account postulates that inclusion of duds in line-ups causes the non-dud alternatives to be perceived as more perceptually similar to the memory representation of the actual perpetrator, which, in turn, elevates confidence.¹

Although Charman et al. (2011) did not delineate the exact psychological mechanism of how duds affect perceived similarity between non-dud line-up member and the memory representation of a perpetrator, work by Tversky (1977) may be helpful in this regard. Tversky argued that changing the local context in which similarity judgments are made by including a novel alternative could broaden the set of features that are considered diagnostic of category membership, a phenomenon referred to as the *extension effect*. Applying this logic to Charman's et al. recognition results, if all non-dud members of a line up share a particular feature (e.g., blonde hair) with the memory representation of the perpetrator, that feature is nondiagnostic for recognition: All non-dud members are perceived to be equally similar to the memory representation with respect to hair color and so the feature cannot be used to differentiate amongst the candidates and may not even be considered at all. However, when the line-up also includes a dark-haired suspect (dud), hair color acquires diagnostic value, boosting the perceived similarity between non-duds and the memory representation. That is, in this new context, non-duds share the feature "blonde hair" with the perpetrator whereas the dud does not, and so the presence of the dud facilitates the perceived similarity (and judged likelihood) between non-dud line-up members and the memory representation.

However, despite this potential role of feature diagnosticity on perceived similarity, in our view, the perceptual similarity hypothesis proposed by Charman et al. (2011) is overly narrow. When several alternatives of what happened in the past are considered, it is rare that the question boils down to which alternative is most perceptually similar to what is in memory. Consequently, the main empirical objective of the present study is to demonstrate that the dud-alternative effect occurs in episodic memory tasks for which perceptual similarity plays a minimal role.

For this demonstration, we developed a novel paradigm. Our procedure is based on associative recognition, a memory task requiring retrieval of inter-item associations (e.g.,

Kelley & Wixted, 2001; Verde, 2004). Participants in our paradigm decided which of the provided alternatives was paired with a given cue in the study phase. A condition in which two non-dud alternatives were provided, one being a target paired with this cue and the other being a target paired with other cues, was pitted against a condition in which the third, dud alternative was also included. The dud alternative was novel, not paired previously with any cue, and thus easily rejected because of low familiarity. Given that the paradigm is based on memory for perceptually impoverished but semantically rich materials, namely pairs of words, we assume that in our task, perceptual similarity of presented alternatives to the correct answer plays a negligible role. Thus, if the dud-alternative effect in memory is restricted to situations in which duds affect perceptual similarity of non-dud alternatives to targets, then we would predict no dud-alternative effect in our paradigm. If, however, the dud-alternative effect extends beyond such limited circumstances and obtains more broadly in confidence judgments provided in an episodic memory setting, we can expect this effect to occur in our paradigm.

Experiment 1

Method

Participants. Forty undergraduates from Cardiff University participated for course credit or small monetary compensation (age: $M = 19.13$; 4 males).

Materials and Design. The experiment consisted of two study-test blocks, with one block corresponding to the dud-present condition and the other to the dud-absent condition. The order of the two block types was counterbalanced across participants.

To generate the materials for the two blocks, 64 4-8 letter nouns with average ratings on various dimensions (e.g., lexical frequency, concreteness, imageability) were chosen from the MRC database. Thirty-two words were used as study materials and the remaining 32 words were used only as duds in a final test of the dud-present condition., The 32 study

words were randomly divided into two lists of 16 words, with separate lists to be used in each block (Figure 1). For each list of 16 words, eight words were randomly chosen to act as cues and eight to act as targets. Each of four cues (1-4) within each list was randomly paired with four different targets (A-D) to make 16 pairs (4 X 4), whereas each of the remaining four cues (5-8) was paired with each of the remaining four targets (E-H) to make another 16 pairs. Each resultant list of 32 pairs was presented three times within the study phase of each block for a total of 96 study trials per block. List repetition did not occur until the previous list presentation was complete and a different random order of pairs was used within each list presentation. Thus, in the study phase of each block, every cue and target word was seen for a total of 12 times each (four times within each list X three presentations of the list). Assignment of lists to blocks was counterbalanced across participants. (see Verde, 2004, for a similar arrangement of materials, referred to as interference sets). The aim of using interference sets was to allow for multiple presentations of individual cues and targets (to increase their familiarity) without producing ceiling effects in performance.

At each test trial, a single cue was presented with two or three alternatives corresponding to the dud-absent and dud-present conditions (blocks), respectively. In the dud-absent condition, the alternatives included the target for the presented cue plus one of the four targets not paired with the cue (non-dud foil). The same two alternatives were presented in the dud-present condition along with a new word not shown at study (dud foil). The order of the two or three alternatives on the screen was random. For both test conditions/blocks, every cue was presented four times corresponding to the four pairs for which the cue was presented at study. Similarly, every target was presented eight times in each block, four times as a correct alternative and four times as a non-dud foil. Every dud in the dud-present block was presented once only.

Procedure. Participants were first instructed that they should memorize the pairs and that all pairs will be repeated several times. At study each pair was presented for two seconds with a 0.5 second interval. During study, there was no indication that the study list was complete and was about to be repeated either the first or second time.

At test, participants were informed that they would be presented with individual cues from the study phase together with some alternatives out of which one and only one was a target previously paired with a given cue. Participants were asked to rate each alternative for their confidence that it was paired with a given cue on a scale from 1 (very low confidence) to 6 (very high confidence). Responding in the test was self-paced. After the test for the first block was finished, the second study phase ensued and the test for this list followed.

Results and Discussion

Table 1 presents mean confidence ratings for this experiment. Duds received very low confidence ratings ($M = 1.15$; $SD = 0.27$), indicating that participants were highly confident that duds were not paired with cues. Following Windschitl and Chambers (2004), we did not actually require participants to choose a target from amongst the alternatives. Instead, participants only rated confidence in each alternative. However, by using confidence ratings, it is possible to determine target selection rate by examining the proportion of trials on which a target alternative was given the highest confidence rating from among presented alternatives. When a target and one of the remaining alternatives received the same highest confidence judgment, the trial was excluded from the analysis.² A t-test comparing the proportion of included trials with targets receiving the highest confidence ratings failed to reveal a significant difference between the dud-present ($M = .67$; $SD = .15$) and dud-absent ($M = .69$; $SD = .17$) conditions, $t < 1$. This null result suggests that including duds does not affect participants' ability to identify targets.

The confidence ratings given to non-dud alternatives were initially analyzed with an ANOVA that included the order of conditions as a factor. However, this factor failed to produce any significant effects, so we collapsed the data from the two counterbalancing conditions. A resulting 2 (alternative status) x 2 (test condition/block) repeated-measures ANOVA yielded a significant main effect of alternative status, $F(1, 39) = 68.87$, $MSE = 0.89$, $p < .001$, $\eta_p^2 = .64$, which indicates that participants gave higher judgments for targets than for non-dud foils. This main effect reflects a certain level of memory for associations. More interestingly, the ANOVA also yielded a significant main effect of test condition, $F(1, 39) = 6.68$, $MSE = 0.29$, $p = .014$, $\eta_p^2 = .15$, with higher confidence ratings when duds were included in the test than if they were not. The interaction was not significant, $F(1, 39) = 1.63$, $MSE = 0.35$, $p > .20$, indicating that duds affected confidence for both targets and non-dud foils. These results indicate that the dud-alternative effect is present in our paradigm, extending this effect to a memory task in which the role of perceptual features of the assessed alternatives is negligible.²

The present results document the dud-alternative effect in associative recognition. Inclusion of dud alternatives in test trials led to higher confidence ratings for both targets and non-dud foils. Thus, at test, providing local context in the form of a highly implausible alternative elevates confidence for the remaining alternatives, even if the memory task is clearly not dependent on assessing perceptual similarity of alternatives to information stored in memory. This result demonstrates that the dud-alternative effect is a general phenomenon of episodic memory when people assess confidence for multiple alternatives concerning past events. It is not limited to perceptually-driven tasks like the line-up procedure.

However, Experiment 1 suffers from one important caveat. Although we have argued that duds constitute a local context for the remaining, non-dud alternatives on a given test

trial, one could argue that our implementation of duds in Experiment 1 has more to do with global than local context. Because the presence of duds was manipulated between blocks, all test trials in the dud-present block included duds. It is thus possible that duds do not exert their influence on confidence at the level of individual trials but rather at the level of the whole test. To tease apart these two possibilities, Experiment 2 manipulated the presence of duds within a single test. If duds serve as local context and elevate confidence only in trials in which they are included, we should replicate the dud-alternative effect in Experiment 2. If, in contrast, duds exert their influence at the general level of the whole test, we should obtain no dud-alternative effect in Experiment 2.

Experiment 2

Method

Participants. Twenty-four undergraduates from Cardiff University participated for course credit (age: $M = 19.17$; all females).

Materials, Procedure, and Design. One of the lists of cue-target pairs developed for the previous experiment was used together with half of the words that served as duds in the final test. The study phase was the same as either study phase in Experiment 1. In the test phase, half of the trials were presented with only two alternatives (target and non-dud foil) whereas the remaining half included an additional dud. Presentation order of the two trial types was random. The assignment of items to the two test conditions was counterbalanced across participants.

Results and Discussion

Table 1 presents mean confidence for this experiment. Confidence ratings for duds were again, as intended, very low ($M = 1.12$; $SD = 0.21$). The first analysis again looked at

the proportion of trials in which targets received the highest confidence rating.³ A t-test showed no significant difference between the dud-present ($M = .73$; $SD = .16$) and dud-absent ($M = .70$; $SD = .16$) conditions, replicating Experiment 1 and showing that inclusion of duds does not affect participants' ability to identify targets. A 2 (alternative status) x 2 (test condition) repeated measures ANOVA performed on the confidence ratings for the non-dud alternatives yielded a significant main effect of status, $F(1, 23) = 50.50$, $MSE = 1.01$, $p < .001$, $\eta_p^2 = .69$, which again indicates that participants gave higher judgments for targets than for non-dud foils. More importantly, the main effect of the test condition was also significant, $F(1, 23) = 9.51$, $MSE = 0.08$, $p = .005$, $\eta_p^2 = .29$. These results are consistent with the results of Experiment 1 and indicate that the presence of duds inflated confidence judgments. The interaction was not significant, $F < 1$.

Thus, Experiment 2 again documented that duds inflate confidence for non-dud alternatives, this time with the presence of duds manipulated within a single test. This result suggests that duds serve as local context affecting confidence at the level of individual test trials. The theoretical underpinnings of these results are discussed next.

General Discussion

In the present study we developed a novel paradigm for investigating the dud-alternative effect in memory. Using a variant of the associative recognition task, we created conditions under which an unfamiliar and therefore easily rejected dud alternative was included in the test, affecting confidence for the remaining, non-dud alternatives. In two experiments, using a design in which the presence of duds was manipulated between blocks (Experiment 1), and a design in which the presence of duds was manipulated within a single test (Experiment 2), we documented that inclusion of duds inflates confidence for non-dud

alternatives, whether they are targets or non-dud foils. These results speak to the robustness of the dud-alternative effect for confidence judgments concerning past events.

Regarding the mechanism of the dud-alternative effect, it is first important to note that this effect occurs robustly across different tasks. To date, it has been documented in the non-episodic likelihood judgments tasks (Windschitl & Chambers, 2004), as well as episodic memory tasks such as line-up identification (Charman et al., 2011), and associative recognition (the present study). The mechanism of this effect should thus be general enough to account for its various manifestations. The present study is clearly inconsistent with the perceptual-similarity hypothesis developed by Charman et al. for their line-up study. Our results and those of Windschitl and Chambers indicate that the dud-alternative effect in episodic memory requires a broader theoretical framework.

In our view, the results of the investigations of the dud-alternative effect conducted to date remain consistent with the contrast hypothesis developed by Windschitl and Chambers (2004). Duds create a local context to which evidence supporting the assessed alternative is compared. Because duds differ greatly from the focal alternatives on the assessed dimension, when comparison against a dud is made, the perceived value on the assessed dimension for the focal alternative is greater, leading to higher confidence judgments. Charman et al. (2011) claimed that the contrast hypothesis was inconsistent with their results from the line-up procedure. They argued against this hypothesis, because duds in their line-ups increased confidence not only for the endorsed non-dud foil (they used only target-absent line-ups) but also for the other, unendorsed non-dud foil. Moreover, duds affected confidence for participants who were precluded from choosing in a line-up. Charman et al. reasoned that the contrast hypothesis predicts the dud-alternative effect only for the explicitly chosen alternative, which is compared against other alternatives.

However, in our view, Charman et al. (2011) might have unnecessarily restricted the meaning of ‘focal alternative’ in their analysis of the contrast hypothesis. Our reading of Windschitl and Chamber’s (2004) use of the term suggests that a ‘focal alternative’ is any alternative for which confidence is currently being assessed. Consequently, confidence for an assessed alternative could be affected by the presence of duds via the contrast mechanism even if it is not ultimately endorsed. We consider the contrast hypothesis a possible overarching account of the dud-alternative effect that is able to provide a framework for understanding various manifestations of this effect.

A somewhat more detailed implementation of the contrast hypothesis may be offered by the aforementioned analysis of Tversky (1977). Tversky proposed that a local context may determine which features of the stimuli are considered diagnostic for a similarity judgment. In the case of perceptual similarity, providing a dud alternative may highlight and thus render diagnostic perceptual features of assessed line-up members that would otherwise be ignored. But this logic may be extended to other cases of similarity as well. In the case of the associative recognition task, the fact that two alternatives were both studied may not be deemed diagnostic when all words provided in the test trial were studied. However, when duds are included and non-duds alternatives are contrasted with duds, the fact that non-dud alternatives were studied becomes salient and thus diagnostic to the decision of which alternative was associated with the cue, inflating confidence judgments. The idea of diagnosticity of features developed by Tversky can be thus seen as a detailed specification of the contrast hypothesis proposed by Windschitl and Chambers (2004).

We need to acknowledge, however, that an alternative account of the dud-alternative effect is also possible. Both Windschitl and Chambers (2004) and Charman et al. (2011) discussed a hypothesis that could be referred to as a *recalibration hypothesis*. According to this hypothesis, duds do not truly affect underlying subjective confidence for the assessed

alternatives but instead change the scale on which confidence judgments are made. When duds are absent, confidence assigned to non-dud foils may create an anchor at the low end of the scale (e.g., Tversky & Kahneman, 1974). However, when a dud alternative is included, the complete lack of confidence associated with this alternative may cause the confidence scale to be recalibrated. Recalibration sets a new, lower anchor for the confidence scale, pushing the assigned confidence values for the remaining alternatives upward despite no actual change in the underlying subjective confidence. Under this scenario, the dud-alternative effect is more about how people define the low end of a response scale given a certain set of alternatives than it is about true changes in subjective confidence.

Our results have some consequences for the recalibration hypothesis. Specifically, the fact that we observed the dud-alternative effect in Experiment 2, when the presence of duds was manipulated within a single test, suggests that if duds induce recalibration, then recalibration occurs not at the level of an entire test but at the level of individual trials. In other words, if the recalibration hypothesis is correct, our results would suggest that the way subjective confidence is translated into values on confidence scale is highly malleable and may be adjusted on the trial-by-trial basis. At present, the status of this novel, continuous recalibration hypothesis is unclear and further tests with dependent measures other than confidence judgments may be necessary to investigate it. We believe that our novel paradigm for investigating the dud-alternative effect can be used in further studies on this issue.

References

- Charman, S. D., Wells, G. L., & Joy, S. W. (2011). The dud effect: Adding highly dissimilar fillers increases confidence in lineup identifications. *Law & Human Behavior, 35*, 479-500. doi: <http://dx.doi.org/10.1007/s10979-010-9261-1>
- Clark, S. E. (1997). A familiarity-based account of confidence-accuracy inversions in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 23*, 232-238. doi: <http://dx.doi.org/10.1037/0278-7393.23.1.232>
- Tversky, A. & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science, 185*, 1124–1130. doi: <http://dx.doi.org/10.1126/science.185.4157.1124>
- Kelley, R., & Wixted, J.T. (2001). On the nature of associative information in recognition memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 27*, 701-722. doi: <http://dx.doi.org/10.1037//0278-7393.27.3.701>
- Tulving, E. (1981). Similarity relations in recognition. *Journal of Verbal Learning and Verbal Behavior, 20*, 479-496. doi: [http://dx.doi.org/10.1016/S0022-5371\(81\)90129-8](http://dx.doi.org/10.1016/S0022-5371(81)90129-8)
- Tversky, A. (1977). Features of similarity. *Psychological Review, 84*, 327-352. doi: <http://dx.doi.org/10.1037//0033-295X.84.4.327>
- Verde, M. F. (2004). Associative interference in recognition memory: A dual-process account. *Memory & Cognition, 32*, 1273-1283. doi: <http://dx.doi.org/10.3758/BF03206318>
- Windschitl, P. D., & Chambers, J. R. (2004). The dud-alternative effect in likelihood judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 198-215. DOI: <http://dx.doi.org/10.1037/0278-7393.30.1.198>

Footnotes

1. Tulving (1981) distinguished between perceptual and ephoric similarity in memory tasks, where perceptual similarity refers to similarity between items physically presented in a test and ephoric similarity refers to similarity between items presented in a test and information stored in memory. In these terms, the hypothesis considered by Charman et al. (2011) states that perceptual dissimilarity between the assessed alternative and a dud leads to increased ephoric similarity between the assessed alternative and stored memory for perpetrator. Here, however, we will refer to perceptual similarity as referring to perceived similarity between assessed alternative and memory for perpetrator.
2. There were 5.15 trials (16% of all trials) excluded from the dud-present list and 4.15 trials (13% of all trials) excluded from the dud-absent list. Importantly, the number of excluded trials did not differ between the lists, $t(39) = 1.33$, $SE = .75$, $p = .19$.
3. We once again excluded from the analysis trials for which a target and one of the remaining alternatives received the same highest confidence judgment. There were 2.33 trials (15% of all trials) excluded from the dud-present condition and 1.92 trials (12% of all trials) excluded from the dud-absent condition. The number of excluded trials did not differ between conditions, $t < 1$.

Table 1. *Mean Confidence Judgments In Experiments 1 and 2 as a Function of Test Condition (Dud-Present vs. Dud Absent) and Alternative Status (Target, Non-Dud Foil, Dud Alternative). Standard Deviations Are Given in Parentheses.*

Item	Experimental Condition	
	Dud-present	Dud-absent
Experiment 1		
Target	4.50 (0.58)	4.40 (0.66)
Non-dud foil	3.38 (0.94)	3.04 (0.79)
Dud alternative	1.15 (0.27)	-
Experiment 2		
Target	4.62 (0.61)	4.46 (0.66)
Non-dud foil	3.17 (1.00)	2.98 (0.85)
Dud alternative	1.12 (0.21)	-

<u>Study List (32 unique pairs)</u>			<u>Test List</u>			
3X		Cues	Targets		Cues	Targets
	1	1	A	17	5	E
	2	1	B	18	5	F
	3	1	C	19	5	G
	4	1	D	20	5	H
	5	2	A	21	6	E
	6	2	B	22	6	F
	7	2	C	23	6	G
	8	2	D	24	6	H
	9	3	A	25	7	E
	10	3	B	26	7	F
	11	3	C	27	7	G
	12	3	D	28	7	H
	13	4	A	29	8	E
	14	4	B	30	8	F
	15	4	C	31	8	G
16	4	D	32	8	H	

<p><u>Dud-Absent Trial</u></p> <p style="text-align: center;">A (target)</p> <p>1 - ?</p> <p style="text-align: center;">E (non-dud foil)</p>
<p><u>Dud-Present Trial</u></p> <p style="text-align: center;">A (target)</p> <p>1 - ?</p> <p style="text-align: center;">E (non-dud foil)</p> <p style="text-align: center;">X (dud foil)</p>

Figure 1. Schematic depiction of the design used in Experiment 1. In the study list, numbers designate eight different word cues and letters designate eight different word targets. For each of the two study-test blocks, 32 unique cue-target pairs were generated by pairing each cue with four targets as shown above. In each block, the 32 pairs were presented in a random order for study, and then the same pairs were presented again a second time in a different random order, and then a third time in third random order. This meant that each cue and target was viewed 12 times in total by the end of the 96 study trials for a given block, making them very familiar. At test, participants were required to rate confidence in each alternative on each trial. The alternatives were a target and a non-dud (familiar) foil in the dud-absent condition/block. These same two words plus a new dud word not presented at study were the alternatives in the dud-present condition/block.