# Who is susceptible to conjunction fallacies in category-based induction?

**AIDAN FEENEY**
*Durham University, Stockton-on-Tees, England*

**PATRICK SHAFTO**
*Massachusetts Institute of Technology, Cambridge, Massachusetts*

AND

**DARREN DUNNING**
*Durham University, Stockton-on-Tees, England*

Recent evidence suggests that the conjunction fallacy observed in people's probabilistic reasoning is also to be found in their evaluations of inductive argument strength. We presented 130 participants with materials likely to produce a conjunction fallacy either by virtue of a shared categorical or a causal relationship between the categories in the argument. We also took a measure of participants' cognitive ability. We observed conjunction fallacies overall with both sets of materials but found an association with ability for the categorical materials only. Our results have implications for accounts of individual differences in reasoning, for the relevance theory of induction, and for the recent claim that causal knowledge is important in inductive reasoning.

Conjunction fallacies occur in reasoning when people assign a higher probability to a conjunction than to one or other of its constituents. In the most famous demonstration in the literature (Tversky & Kahneman, 1983) people read a description of Linda, a 31-year-old, smart, outspoken woman who was a philosophy major, concerned with discrimination and social justice, and a participant in antinuclear demonstrations. When asked to rank various statements about Linda in order of their probability, people rank the statement "Linda is a bank teller and is active in the feminist movement" above the statement "Linda is a bank teller," resulting in an apparent normative contradiction.

Recently, Medin, Coley, Storms, and Hayes (2003) have demonstrated a parallel conjunction fallacy in category-based induction. The effect was demonstrated for two different kinds of materials: (1) category reinforcement arguments and (2) causal arguments (see Table 1). Both sets of arguments concerned related categories. The categories in Arguments 1–3 are related by virtue of a shared property. Accordingly, Medin et al. referred to them as property-reinforcement materials. However, we prefer the term *category reinforcement materials* as in each case the shared property (e.g., living at altitude) can be thought of as a categorical relation, and in many cases what is shared corresponds more directly to a category (e.g., *cows*, *pigs*, and *sheep* are all farm animals). The categories in Arguments 4–6, on the other hand, are causally related.

Medin et al. (2003) found that, across a number of sets of arguments, the mean ratings of inductive strength (on a 9-point scale) for two-conclusion category arguments, like 1 and 4, were greater than the means for the corresponding single-conclusion category arguments, like 2 and 3 and 5 and 6. These judgments are normatively contradictory, as the probability of a conjunction cannot be greater than the probability of its constituents, and replicate the general conjunction fallacy effect. Although Medin et al. do not systematically analyze differences between judgments based on causal and categorical knowledge, differences between categorical and causal materials highlight the fact that the fallacy may be based on different kinds of knowledge about the relationships among categories. One important difference between these arguments is that whereas the relation between the premise and conclusion categories is the same in both single-category reinforcement arguments, there are different premise–conclusion relations in the single causal arguments. For example, because mice eat grain, a property might be directly transmitted from grain to mice. However, because owls eat mice—not grain—a property is likely to be indirectly transmitted from grain to owls via mice. Thus, the causal relationship in 5 is closer than that in 6, and we refer to arguments like 5 as *single close* and to arguments like 6 as *single distant*.

Although the conjunction fallacy effect is decidedly robust, applying to diverse experimental judgments based on

A. Feeney, aidan.feeney@durham.ac.uk

**Table 1**
**Arguments Used to Demonstrate Category Reinforcement**
**and Causal Conjunction Fallacies**

Category Reinforcement Arguments

| 1. Two category | Andean people have property $X12$ |
| | Therefore, Himalayan people and Alpine people have property $X12$ |
| 2. Single category | Andean people have property $X12$ |
| | Therefore, Himalayan people have property $X12$ |
| 3. Single category | Andean people have property $X12$ |
| | Therefore, Alpine people have property $X12$ |

Causal Arguments

| 4. Two category | Grain has property $M4$ |
| | Therefore, Mice and Owls have property $M4$ |
| 5. Single close | Grain has property $M4$ |
| | Therefore, Mice have property $M4$ |
| 6. Single distant | Grain has property $M4$ |
| | Therefore, Owls have property $M4$ |

Note—Adapted from Medin, Coley, Storms, and Hayes (2003).

different kinds of knowledge, there appear to be individual differences in susceptibility to the fallacy: People who resist it tend to report higher SAT scores. Stanovich and West (1998) have explained these individual differences in terms of a dual-process theory of reasoning (Evans, 2003; Sloman, 1996; Stanovich, 1999), which posits the existence of two reasoning systems: a fast, largely unconscious, parallel, and pragmatic System 1 working alongside a slower, memory-intensive, conscious, serial, and evolutionarily more recent System 2. Under this view, in cases where normatively correct task performance is not associated with ability, System 1 and System 2 are said to deliver the same, normatively correct, response. In cases where there is an association between performance and ability, the systems deliver different responses. System 1 is said to operate on a contextualized representation of the problem delivering the normatively incorrect response, and System 2 operates on a decontextualized representation that abstracts away from the content and context of the problem to its underlying form (see Stanovich, 1999, pp. 169–174), thus delivering the correct response. Accordingly, participants who are higher in System 2 resources (i.e., cognitive ability) are more likely to suppress the incorrect response from System 1 in favor of System 2, and give the correct response based on their decontextualized representation of the problem.

One as-yet-unexplored possibility allowed by this explanatory framework is that differences in susceptibility to reasoning fallacies may also arise due to the background knowledge that is evoked by particular content. That is, because it may be harder to construct a decontextualized representation of some problems than others, some problems may lead to higher rates of fallacious responding than others. At the moment, dual process theory offers no principled means of generating predictions about the ease with which decontextualized representations may be constructed in different knowledge domains.

In the category-based induction literature, research has suggested that causal and categorical knowledge are treated differently, and intriguingly suggests that knowledge effects may also be qualified by individual differences. For example, Rehder (2006) has investigated the robustness, in the presence of causal knowledge, of typicality and diversity, two classic effects in category-based reasoning that are indicative of categorically organized knowledge. In one experimental trial, participants learned about Kehoe ants and were asked about their willingness to project a property from specific exemplars to the entire class of Kehoe ants. The property was either blank, such as having choroidal parasites attached to its eyes, or causally related to known features of Kehoe ants (e.g., having a stinging venomous bite, which is caused by a high concentration of iron sulphate, a preexisting property of Kehoe ants).

Rehder (2006) observed significantly greater typicality and diversity effects in the blank property condition than in the causal condition: When causal knowledge was available, participants preferred to generalize based on the causal relations among features. However, an individual differences analysis revealed two patterns of responding, one group that relied primarily on causal knowledge and the other that relied primarily on categorical knowledge, though no explanation of the origin of the individual differences was offered. When considered in the context of studies with real-world concepts (e.g., Lopez, Atran, Coley, Medin, & Smith, 1997; Shafto & Coley, 2003), the results suggest that causal knowledge is preferred as a basis for inductive inference, but leave unresolved the cause of individual differences in the use of causal and categorical knowledge.

Combining the literatures on conjunction fallacies and category-based induction, we see two factors that may affect whether the conjunction fallacy is observed in category-based induction: (1) cognitive ability, and (2) relationships among categories in an argument. Consequently, our goals in this study are, first, to replicate the conjunction fallacy effects in category-based reasoning demonstrated by Medin et al. (2003) for both causal and categorical knowl-

edge. Second, we attempt to replicate findings regarding individual differences in susceptibility to the conjunction fallacy (Stanovich & West, 1998). On the basis of previous research, we expect people with higher cognitive ability to be less susceptible to the fallacy. Third, we ask whether conjunction fallacies based on causal knowledge are more frequent than those based on categorical knowledge. This last possibility comes from Rehder's (2006) finding that, for the majority of participants evaluating category-based inductive arguments, causal relations are a more compelling source of information than taxonomic relations.

## METHOD

### Participants

A total of 114 female and 16 male undergraduates at the University of Durham participated in this experiment.

### Materials and Design

Each participant completed the AH4 (Heim, 1970), a widely used group test of cognitive ability.[1] Participants also completed a reasoning questionnaire comprising 18 conjunction fallacy by category reinforcement arguments and 18 causal conjunction fallacy arguments. Some of these arguments are taken from Medin et al. (2003). However, the majority of the arguments are of our own devising. Each group of 18 arguments was made up of six sets, each containing 3 arguments, two whose conclusion concerned a single category, and a third whose conclusion concerned both of the categories described in the conclusion of the first two arguments.

An example set of category reinforcement arguments follows:

Pasta has property *X13*.
How likely is it that rice and potatoes have property *X13*?

Pasta has property *X13*.
How likely is it that potatoes have property *X13*?

Pasta has property *X13*.
How likely is it that rice has property *X13*?

The causal conjunction fallacy materials were presented in a similar fashion. For all arguments, participants were asked to make their strength judgments on 9-point likert scales ranging from 1 (*very unlikely*) to 9 (*very likely*). Note that, unlike Medin et al., we used the same blank property in all three arguments in a set. All 36 arguments in the booklet were presented in one of four different random orders.

### Procedure

The cognitive ability and reasoning data were collected from participants on two separate occasions in lectures.

## RESULTS

Scores on the AH4 ranged from 46 to 122. The mean overall score was 94.6 ($SD = 14.3$), close to the norm for this type of sample ($M = 96.4$, $SD = 15.0$). We coded participants as high or low in ability. Participants who scored 95 or less were coded as low ($N = 64$), whereas all other participants were coded high ($N = 66$). The mean ability score for the low group was 83.1 ($SD = 11.0$), whereas the mean score for the high group was 105.7 ($SD = 5.9$).

### Conjunction Fallacy by Category Reinforcement

For the purposes of the analysis that follows, in each set we coded the single-category-conclusion arguments that had received the highest mean rating as single strong and the other as single weak. We coded a set as displaying sensitivity to the fallacy if the strength rating attributed to the two-category conclusion argument was greater than that assigned to either of the single-category-conclusion arguments. In Table 2, these data are broken down by phenomenon type and item. Overall, participants were susceptible to the fallacy on 47% of category reinforcement trials. We observed a negative association between the proportion of cases in which participants displayed the fallacy on

**Table 2**
**Percentage of Participants Displaying the Fallacy, Mean Ratings of Inductive Strength, and Standard Deviations, Broken Down by Phenomenon Type, Problem Set, and Argument**

| | | Argument Ratings | | | | | |
|---|---|---|---|---|---|---|---|
| | | Two Category | | Single Strong/ Single Close | | Single Weak/ Single Distant | |
| Phenomenon Type and Problem Set | % Susceptible* | M | SD | M | SD | M | SD |
| Category Reinforcement | | | | | | | |
| Cabbage/lettuce, spinach | 56 | 5.0 | 1.9 | 4.7 | 2.0 | 4.5 | 1.9 |
| Cows/pigs, sheep | 48 | 4.1 | 1.9 | 3.9 | 1.7 | 3.8 | 1.8 |
| Pasta/potatoes, rice | 45 | 4.7 | 2.2 | 4.7 | 2.1 | 4.4 | 2.1 |
| Polish people/Czech people, Ukrainian people | 45 | 4.4 | 2.1 | 4.3 | 2.0 | 4.1 | 1.9 |
| Andean people/Alpine people, Himalayan people | 44 | 4.3 | 2.1 | 4.3 | 2.0 | 3.9 | 1.9 |
| Brown bears/panda bears, polar bears | 42 | 4.6 | 2.1 | 4.8 | 2.0 | 4.6 | 2.0 |
| Overall | 47 | 4.5 | | 4.4 | | 4.2 | |
| Causal | | | | | | | |
| Soil/grass, cows | 62 | 4.7 | 2.2 | 5.5 | 2.1 | 3.5 | 1.9 |
| Sheep/wool, sweaters | 61 | 5.5 | 2.3 | 5.9 | 2.1 | 4.3 | 2.2 |
| Lead/pipes, plumbers | 51 | 3.3 | 2.0 | 5.4 | 2.3 | 2.5 | 1.8 |
| Cows/milk, ice cream | 49 | 5.4 | 2.1 | 6.1 | 2.1 | 4.6 | 2.0 |
| Grain/mice, owls | 44 | 3.8 | 2.1 | 4.1 | 2.2 | 3.1 | 1.9 |
| Sea water/cod, sharks** | 33 | – | | – | | – | |
| Overall | 50 | 4.5 | | 5.1 | | 3.9 | |

*% Susceptible refers to the proportion of participants who rated the two-category conclusion argument stronger than at least one of the single-category-conclusion arguments.    **We excluded this set from the analysis of strength ratings. In the order they would have appeared in the table, the mean ratings for the arguments in this set are 4.5, 4.8, and 4.5.

the category reinforcement stimuli and their scores on the AH4 [$r(130) = -.21, p < .02$]. That is, participants who did poorly on the AH4 were more likely to be susceptible to the fallacy.

Mean ratings of argument strength, broken down by item and argument, are shown in Table 2. The results of a 2 (ability) $\times$ 3 (argument: two category, single strong, and single weak) $\times$ 6 (set) ANOVA contained a significant main effect of argument [$F(2,256) = 12.81, MS_e = 1.53, p < .001$]. This main effect was qualified by a significant interaction between argument and ability [$F(2,256) = 5.11, MS_e = 1.53, p < .01$]. The means involved in this interaction are displayed in Figure 1.

Planned comparisons revealed that the low group attributed significantly greater strength to two-category arguments than to either single weak ($p < .001$) or single strong ($p < .05$) arguments. For the high group, strength ratings for the two-category arguments did not differ significantly from strength ratings for either weak or strong single-category arguments.

The results of the ANOVA contained a significant main effect of set [$F(5,640) = 9.40, MS_e = 4.42, p < .001$], but the interaction between set and argument was not significant [$F(10,1280) = 1.14, MS_e = 1.19, p = .33$]. Examination of the means in Table 2 reveals that, whereas the mean for the single weak argument was never greater than the mean for the two category argument, the size of the difference between these means varied considerably between sets. Post hoc tests revealed that the difference was significant for four out of six sets. For the Poles set, the difference was marginal ($p < .06$), but for the bears set, the two-category conclusion ratings—although slightly greater than those for one of the other arguments—were not significantly so.

## Causal Conjunction Fallacy

The item involving sea water, cod, and sharks produced relatively low levels of susceptibility to the fallacy (see Table 2). One reason for this discrepancy may be that there are independent routes for causal transmission from seawater to cod and from seawater to sharks. Thus, cod and sharks may be an equal causal distance from seawater. Accordingly, we did not include this item in any of the analysis that follows.

Excluding this item, participants were susceptible to the causal conjunction fallacy on 53% of trials. However, more prevalent than susceptibility to the fallacy was use of the causal distance between the premise category and the conclusion category as a cue to argument strength. This was particularly true for the single close arguments. On 81% of trials, participants rated these arguments stronger than the two-category argument. By contrast, participants rated the single distant argument stronger than the two-category argument on only 50% of trials.

In contrast with our findings with the category reinforcement materials, there was little evidence that higher ability participants are more able to resist the causal conjunction fallacy. The negative association between susceptibility to the fallacy and ability was nonsignificant [$r(130) = -.09, p > .3$]. There was some evidence of individual differences in the tendency to use causal knowledge. The tendency to rate the single close argument stronger than the two-category argument was associated with ability [$r(130) = .22, p < .02$]. High-ability participants rated the Close argument stronger than the two-category argument on 85% of occasions, whereas for low-ability participants, this figure was 77.5% [$t(128) = 2.30, p < .03$].

A 2 (ability) $\times$ 3 (argument) $\times$ 5 (set) ANOVA on mean ratings of argument strength for the causal conjunction fallacy materials revealed significant main effects of all three variables. The significant main effect of ability [$F(1,128) = 4.59, MS_e = 23.44, p < .04$] was due to the low group's giving significantly higher ratings of inductive strength ($M = 4.7$) than did high participants ($M = 4.3$). However, neither the interaction between ability and argument [$F(2,256) < 1$] (see Figure 1), nor the three-way interaction [$F(8,1024) = 1.30, MS_e = 1.86, p = .24$] was significant.

There was a significant main effect of argument [$F(2,256) = 105.32, MS_e = 2.31, p < .001$], which was qualified by a significant interaction between argument and set [$F(8,1024) = 152.69, MS_e = 1.86, p < .001$]. The means involved in this interaction are presented in Table 2. Post hoc tests revealed that, for all sets, ratings for the two-conclusion argument were significantly greater than the
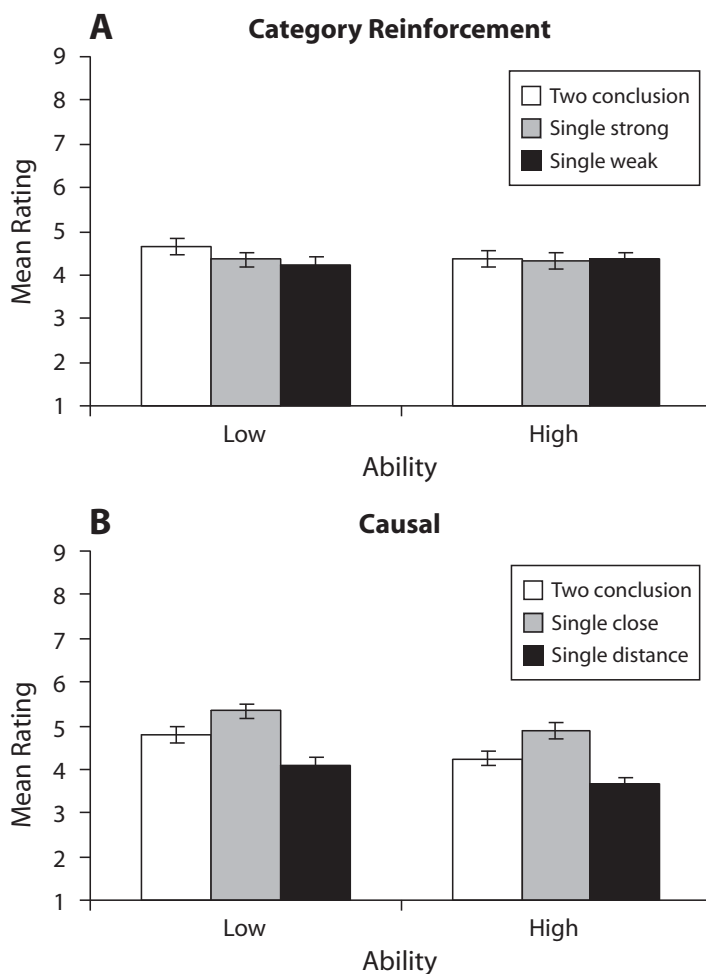


Figure 1. Mean argument strength ratings broken down by ability and argument for (A) category reinforcement materials and (B) causal materials.

ratings for the single distant argument, and significantly lower than those for the single close arguments. The interaction is due to variation in the size of these differences.

## Comparing Conjunction Fallacies

A paired samples $t$ test revealed that susceptibility to the conjunction fallacy was greater with causal materials than with category reinforcement materials [$t(129) = 2.99$, $p < .005$]. We also carried out a 2 (fallacy) $\times$ 2 (ability) $\times$ 2 (strength) mixed design ANOVA on the mean difference between ratings for each of the single-category arguments and ratings for the two-conclusion argument. For the purposes of this analysis, we collapsed across sets. In Figure 2, we present the means involved in the nonsignificant [$F(1,128) < 1$], three-way interaction from that analysis. Single sample $t$ tests on the means involved revealed that, with the exception of the difference scores produced by high-ability participants for category reinforcement materials, all scores differed significantly from 0. An examination of Figure 2 reveals that, although there was, overall, no conjunction fallacy for the causal close arguments, the biggest conjunction fallacy was found for the causal distant arguments. Notably, high- and low-ability participants appear to treat the causal arguments in a similar fashion. However, for the category reinforcement materials, only low-ability participants are susceptible to the fallacy.
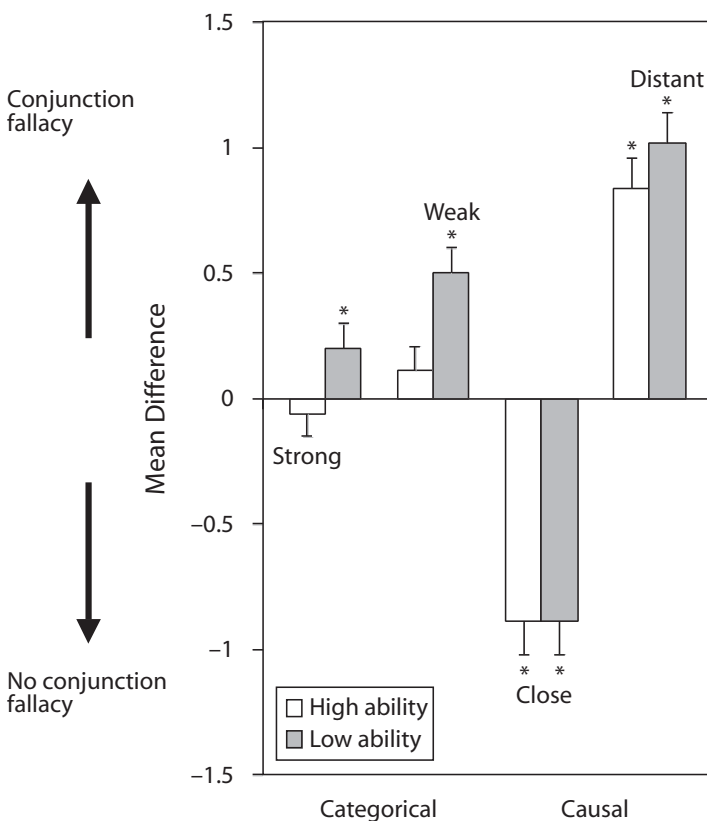


**Figure 2. Mean differences between the strength ratings for the single- and two-conclusion arguments, collapsed across sets. Positive scores indicate a conjunction fallacy, whereas negative scores indicate that the single-conclusion argument was rated stronger than the two-conclusion argument. Asterisked bars differ significantly from 0 (no difference).**

## DISCUSSION

We have replicated Medin et al.'s (2003) finding of a conjunction fallacy in category-based induction. Our results suggest that people are somewhat more susceptible to the causal version of the fallacy than to the category-reinforcement version. However, in interpreting this result, it should be borne in mind that we used a relatively small number of items, the causal and category-reinforcement materials were different, and so may have differed in difficulty, and we excluded one of the causal items from further analysis. Further research is required before we can make strong claims about differences in susceptibility to different versions of the fallacy. Nonetheless, we have found strong evidence that participants used causal knowledge when evaluating causal arguments. We have also observed that the causal conjunction fallacy differs from the fallacy that is obtained by category reinforcement. There was a significant relationship between ability and susceptibility to the conjunction fallacy by category reinforcement; high-ability participants were less likely to be susceptible. In contrast, there was no significant relationship between ability and susceptibility to the causal conjunction fallacy. However, the latter finding is qualified by one significant result suggesting that high-ability participants were more likely to use causal knowledge to overcome part of the fallacy and thereby rate the single close arguments stronger than the two-conclusion arguments.

How could a dual-process model account for the fact that ability affects reasoning about categorical, but not causal, relations? We consider three possibilities. (1) Reasoning about causal relations may be easier than reasoning about categorical relations, allowing both System 1 and System 2 to produce the normative answer for causal materials, but not for the category reinforcement materials. However, this does not account for the fact that there were even fewer normatively correct responses to causal materials than to category reinforcement materials. (2) Causal arguments elicited a stronger response from System 1, making the incorrect response more difficult for System 2 to inhibit (see also Handley, Capon, Beveridge, Dennis, & Evans, 2004; Moutier & Houdé, 2003). On this account, only very high-ability participants would be more likely to inhibit the incorrect response produced by System 1 and respond based on the output of System 2. This would account for the slightly greater rate of nonnormative responding and for the weaker association with ability in causal cases. However, this account would appear to predict more nonnormative responding to the causal materials than we actually observed. (3) Reasoning based on causal knowledge is more difficult because of the need to integrate knowledge about the causal relations among categories with a representation of the probabilistic form of the problem. This would predict that the causal fallacy would be more difficult than the category reinforcement version, consistent with our results, and would also account for the association between ability and the tendency to use certain aspects of causal structure. However, this account appears to be at odds with previous dual process accounts, which claim that prior knowledge is not incor-

porated in decontextualized problem representations in System 2.

Our results also have implications for the relevance framework in category-based induction (Medin et al., 2003). In the relevance framework, both conjunction fallacies are explained by claiming that the addition of a second conclusion category makes the common category or causal chain more available. Our findings suggest that, perhaps, their account holds for the category reinforcement materials only. When reasoning about these materials, we think it likely that only the high-ability participants were able to resist the information about a common category that was more available when there were three categories in the argument. For the causal materials, however, we observed that the strongest ratings of argument strength were for the single close arguments. This suggests that the causal relationship was highly available when evaluating the single close arguments, even though there was just one category in the conclusion of those arguments.

In contrast with the typical conjunction fallacy, where one constituent category (feminist) is rated more probable than the conjunction, whereas the other is rated less probable (bank teller), for the category reinforcement materials, we observed that low-ability participants rated the conjunction more probable than either of its constituents. This emphasizes the difference between the fallacy due to category reinforcement and the fallacy due to representativeness. In the category reinforcement version, information about the common category is most available when there are two categories in the conclusion. Certainly, these results suggest that the category reinforcement effect does not occur because the premise category is more representative of one conclusion category than the other. In this sense, Medin et al.'s (2003) category reinforcement effect is different from other conjunction fallacies that have been reported.

In conclusion, we have replicated both of Medin et al.'s (2003) conjunction fallacies and shown that susceptibility to the category reinforcement fallacy—but not the causal fallacy—is associated with ability. The dual process approach assumes that differences in ability are causally related to susceptibility to these fallacies; but because our design was correlational, it is possible that other factors related to ability cause the observed differences. Future work using experimental designs should investigate this possibility. Nevertheless, our results support recent findings suggesting that causal knowledge is particularly important for inductive reasoning (see Rehder, 2006; Rehder & Hastie, 2001; Shafto & Coley, 2003; Shafto, Kemp, Baraff, Coley, & Tenenbaum, 2005) and suggest that causal and categorical knowledge are treated differently in reasoning. More broadly, these results suggest that any complete theory of reasoning should address both how individual differences in cognitive ability and different kinds of prior knowledge affect reasoning.

**AUTHOR NOTE**

Correspondence relating to this article should be addressed to A. Feeney, Department of Psychology, Durham University Queen's Campus, Thornaby, Stockton-on-Tees TS17 6BH, England (e-mail: aidan.feeney@durham.ac.uk).

**REFERENCES**

Alexopoulos, D. S. (1997). Reliability and validity of Heim's AH4 in Greece. *Personality & Individual Differences*, **22**, 429-432.

Evans, J. St. B. T. (2003). In two minds: Dual process accounts of reasoning. *Trends in Cognitive Sciences*, **7**, 454-459.

Handley, S. J., Capon, A., Beveridge, M., Dennis, I., & Evans, J. St. B. T. (2004). Working memory and inhibitory control in the development of children's reasoning. *Thinking & Reasoning*, **10**, 175-196.

Heim, A. W. (1970). *AH4 group test of intelligence* [Manual]. London: National Foundation for Educational Research.

Lopez, A., Atran, S., Coley, J. D., Medin, D., & Smith, E. E. (1997). The tree of life: Universal and cultural features of folkbiological taxonomies and inductions. *Cognitive Psychology*, **32**, 251-295.

Medin, D., Coley, J. D., Storms, G., & Hayes, B. (2003). A relevance theory of induction. *Psychonomic Bulletin & Review*, **10**, 517-532.

Moutier, S., & Houdé, O. (2003). Judgement under uncertainty and conjunction fallacy inhibition training. *Thinking & Reasoning*, **9**, 185-201.

Rehder, B. (2006). When similarity and causality compete in category-based property induction. *Memory & Cognition*, **34**, 3-16.

Rehder, B., & Hastie, R. (2001). Causal knowledge and categories: The effects of causal beliefs on categorization, induction, and similarity. *Journal of Experimental Psychology: General*, **130**, 323-360.

Shafto, P., & Coley, J. D. (2003). Development of categorization and reasoning in the natural world: Novices to experts, naive similarity to ecological knowledge. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, **29**, 641-649.

Shafto, P., Kemp, C., Baraff, E., Coley, J. D., & Tenenbaum, J. B. (2005). *Context sensitive induction*. Paper presented at the 27th Annual Conference of the Cognitive Science Society, Stresa, Italy.

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin*, **119**, 3-22.

Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning.* Mahwah, NJ: Erlbaum.

Stanovich, K. E., & West, R. F. (1998). Individual differences in framing and conjunction effects. *Thinking & Reasoning*, **4**, 289-317

Tversky, A., & Kahneman, D. (1983). Extensional vs. intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, **90**, 293-3l5.

**NOTE**

1. There are two 65-item sections in the AH4, each administered in separate 10-min sessions. The first part contains verbal items concerning direction, verbal opposites, numerical series, verbal analogies, simple arithmetic computations, and synonyms. The second part of the test contains diagrammatic items requiring judgments about analogies, sames, subtractions, series, and superimpositions. The test is scored by summing the raw scores from each of its component parts. Heim (1970) reported correlations between Parts 1 and 2 ranging from .60 to .81. Test–retest consistency and internal reliability for the AH4 are high (Alexopoulos, 1997; Heim, 1970). Correlations of .60 (Alexopoulos, 1997) and .69 (Heim, 1970) have been found between the AH4 and Raven's matrices.