Acta Psychologica 198 (2019) 102890

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Does expertise moderate the seductive allure of reductive explanations?

Emily J. Hopkins[a,b,*], Deena Skolnick Weisberg[a,c], Jordan C.V. Taylor[a]

[a] University of Pennsylvania, United States of America
[b] University of Scranton, United States of America
[c] Villanova University, United States of America

ARTICLE INFO

Keywords:
Expertise
Explanations
Seductive allure
Decision making
Science
Philosophy

ABSTRACT

Non-experts are unduly attracted to explanations of scientific phenomena that contain irrelevant reductive language (e.g., explanations of biological phenomena that mention chemistry; Hopkins, Weisberg, & Taylor, 2016). To determine if expertise would reduce this reasoning error, the current study recruited individuals with graduate-level training in six scientific fields and in philosophy ($N = 580$) and asked them to judge explanations for phenomena from those fields. Like the novices in Hopkins et al. (2016), scientists' ratings of bad explanations were influenced by reductive information when viewing phenomena from outside their field of expertise, but they were less likely to show this bias when reasoning about their own field. Higher levels of educational attainment did improve detection of bad explanations. These results indicate that advanced training in science or logic can lead to more accurate reasoning about explanations, but does not mitigate the reductive allure effect.

1. Introduction

What makes an explanation persuasive? Although intuitively one might point to the content of the explanation as playing the primary role in this decision, a large body of literature in psychology demonstrates that people's judgments of explanations can be swayed by non-explanatory factors such as an explanation's length (e.g., Kikas, 2003; Langer, Blank, & Chanowitz, 1978). This is a particularly troubling issue in science, where the true structure of the world runs often counter to people's intuitions (Shtulman, 2015), amplifying the difficulty in accurately judging explanations. Yet people are regularly presented with scientific explanations for questions as diverse as why climate change is occurring or why a particular treatment for a disease is the most effective. Their ability to evaluate the quality of scientific explanations like these thus has important consequences for their daily lives, as well as larger issues of public policy.

In the current work, we examined one possible avenue towards ameliorating people's abilities to reason appropriately about scientific explanations: expertise. On the one hand, it seems obvious that advanced training in science would help to protect against errors of reasoning about scientific topics. Domain-based expertise is known to affect a wide variety of processes in perception (see review in Stokes, 2018), as when chess experts parse and remember the layout of a chessboard differently than novices (DeGroot, 1965). Expertise with certain categories of objects (e.g., cars, birds) changes the way the brain processes those objects (Gauthier, Skudlarski, Gore, & Anderson, 2000). Expertise also, unsurprisingly, improves the ability to reason within a scientific domain. For example, physics experts are less likely than novices to fall prey to common misconceptions about physical systems (e.g., that heat is a substance that flows through a system; Slotta, Chi, & Joram, 1995; see also Masson, Potvin, Riopel, & Foisy, 2014).

On the other hand, experts are by no means immune to misconceptions and biases. For example, intelligence and cognitive ability are largely unrelated to the bias to reason from our own point of view (see Stanovich, West, & Toplak, 2013, for a review). Goldberg and Thompson-Schill (2009) demonstrated that even biology professors were slower and less accurate to respond to statements about biology that contradicted tenets of naïve biology (e.g., plants are alive) than to statements that corresponded to these tenets (e.g., animals are alive; see also Kelemen, Rottman, & Seston, 2012; Shtulman & Harrington, 2015). Experts also tend to over-estimate their knowledge in their domain of expertise (Fisher & Keil, 2015; Lawson, 2006). These findings suggest that even experts within a field may find it difficult to overcome reasoning biases. Further, all of this previous work has tested experts only within their domain of expertise. This leaves open the question of whether advanced training in a particular science might allow participants to more accurately judge explanations only within that science, or whether any effects of training might extend to evaluating explanations from other fields.

The current study investigated in detail whether and how expertise in science might affect reasoning about scientific explanations. To do

* Corresponding author at: University of Scranton, Alumni Memorial Hall Room 200, Scranton, PA 18510, United States of America.
E-mail address: emily.hopkins@scranton.edu (E.J. Hopkins).

https://doi.org/10.1016/j.actpsy.2019.102890
Received 27 April 2018; Received in revised form 13 February 2019; Accepted 9 July 2019

so, we drew on previous work on the *seductive allure effect*: the finding that including irrelevant neuroscience information in explanations of psychological phenomena makes non-expert participants judge these explanations more favorably than explanations containing no neuroscience information (Fernandez-Duque, Evans, Christian, & Hodges, 2015; Rhodes, Rodriguez, & Shah, 2014; Weisberg, Keil, Goodstein, Rawson, & Gray, 2008; Weisberg, Taylor, & Hopkins, 2015). This phenomenon represents a specific instance of a more general *reductive allure* effect: People prefer scientific explanations that explain higher-level phenomena with respect to component parts or more fundamental processes (e.g., explaining a biological phenomenon with reference to its underlying chemistry; Hopkins, Weisberg, & Taylor, 2016). This result is in line with recent work finding that individuals prefer explanations that refer to more causal mechanisms (Zemla, Sloman, Bechlivanidis, & Lagnado, 2017).

But this tendency represents an error in judgment: Explanations are not necessarily improved by the addition of reductive information. Indeed, the stimuli in Hopkins et al. (2016) were carefully constructed so that the reductive information did not add explanatory value. Participants' preference for such explanations over logically equivalent ones that do not contain reductive information is thus misplaced. Given this, it is important to explore reasons for this attraction and, especially, possible avenues for combatting it. In the current paper, we focus on the latter issue, investigating whether expertise in science could attenuate this attraction. Training in science could confer a general disposition to evaluate explanations more critically, or the specific skills needed to judge explanations from one's own field more carefully, or both.

An initial indication that expertise can moderate the seductive allure of reductive information comes from Weisberg et al. (2008, Study 3), which found that neuroscience experts did not show the seductive allure effect when neuroscience was added to psychology explanations. Similarly, although both undergraduate students and MTurk workers showed the reductive allure effect in Hopkins et al. (2016), the effect was smaller for the students. Furthermore, students gave lower ratings overall than the MTurk workers, particularly for bad explanations. Because 90% of the MTurk workers had completed at least some college, this suggests that the experience of currently being in an academic setting, rather than educational attainment per se, may minimize the effect somewhat, perhaps by priming skepticism or critical thinking. Thus, experts, who work in their field daily, may be less susceptible to the reductive allure effect than non-experts because they are used to treating new information with skepticism. Or, we may see that expertise has no effect on this bias because intelligence and education do not always protect against reasoning errors (Stanovich et al., 2013).

The current study addresses more specifically the potential role of expertise in ameliorating the reductive allure effect for explanations across the sciences. Does advanced training in science protect against the allure of reductive information? If so, does this advanced training confer general immunity to this effect or only for one's chosen discipline? Answering these questions can help to provide insight on the differences between experts and novices with respect to science and on how expertise functions in general.

In this study, we thus present the same explanation judgment task as in Hopkins et al. (2016) to experts in six scientific fields, matching the fields from which the stimulus items were drawn: social science (sociology and political science), psychology, neuroscience, biology, chemistry, and physics. We asked all participants to judge explanations in each of these fields. Half of these explanations were genuinely informative about why the target phenomenon happens (*good*), and half were logically circular or contained superfluous, non-explanatory information (*bad*). Additionally, half of the explanations were *horizontal*, meaning they were given at the same level as the phenomenon (e.g., biological explanations for biological phenomena), and half were *reductive*, meaning they referred to the immediately more fundamental discipline (e.g., chemical explanations for biological phenomena). Crucially, as in the original studies on the seductive allure effect, this reductive information did not affect the logic of the explanation itself and hence should not have added any value, as confirmed through discussion with expert consultants in each of the six target sciences (see Hopkins et al., 2016 for more details).

We expect that advanced training in a field of science will confer some protection against the reductive allure effect within that field, meaning that these participants should judge explanations from their own field in the same way regardless of whether they contain irrelevant reductive information. However, this training may be inadequate to protect participants against the effect in general, meaning that participants may still judge explanations from other disciplines as better when they contain irrelevant reductive information.

Finally, we also recruited a sample of philosophy experts. Philosophy as a discipline is focused on analyzing arguments and explanations, but does not provide discipline-based training in science. Studying the responses of these experts will thus allow us to compare the effect of expertise in logic and reasoning versus content-specific science expertise on judgments of scientific explanations.

2. Method

2.1. Participants

All participants in this study reported that they either had completed or were working towards an advanced degree in one of the following fields: physics, chemistry, biology, neuroscience, psychology, social science (sociology or political science), or philosophy. These are the fields from which the stimuli in Hopkins et al. (2016) were drawn; this allows us to examine participants' ratings of explanations from their particular domain of expertise. Participants' field memberships were determined by self-report: All were asked at the end of the survey to choose which category most closely matched the field of their highest degree (physical sciences, social sciences, engineering, humanities, health, and business) and to identify their specific field. Some were additionally asked to choose their field as part of a screening survey before engaging the main survey.[1] Thirteen participants were excluded because they did not fit into one of the seven target fields of expertise.[2]

Participants were recruited through advertisements posted on academic forums and professional society mailing lists, recruitment emails sent to university departments, word of mouth, and personal networking. Participants received a $10 Amazon gift card in exchange for completing the survey. A total of 659 participants fit the criteria for inclusion, but 79 were excluded from the sample for failing attention check questions (described in the Procedure). The final sample used for all analyses thus consisted of 580 participants (261 men, 303 women, 16 did not report gender). Participants ranged in age from 20 to 78 years (M = 30.5 years). The educational attainment of participants was as follows: 152 had completed some graduate school, 201 had master's degrees, and 227 had either a PhD, MD, or JD (see Table 1 for participant demographics by field of expertise).

2.2. Design

This study used the same stimuli and procedure as Hopkins et al. (2016). All participants completed an online survey hosted by Qualtrics. Participants first completed the Ratings Explanations task. They

[1] This was implemented partway through data collection because large numbers of spam (e.g. many responses submitted from the same IP address within a few minutes) or inappropriate responses (participants who did not meet the education criteria) were being submitted to the survey.

[2] These participants either identified a field that was not one of our six targets (e.g., law, social work), or the information they provided was inconsistent across the different questions we asked about their field of expertise (e.g., one participant said biology, engineering, and chemistry at different points in the survey).

Table 1
Participant demographics by field of expertise.

Field of expertise	*N* (M/F/unreported)	Mean age (range)	Some grad school/Master's/Doctorate	Horizontal/reductive
Physics	84 (58/25/1)	29.8 years (22–71)	21/45/34%	44/40
Chemistry	83 (34/46/3)	26.6 years (21–45)	48/33/19%	40/43
Biology	81 (30/47/4)	32.2 years (21–71)	23/32/44%	42/39
Neuroscience	85 (29/55/1)	28.3 years (21–48)	45/15/40%	45/40
Psychology	81 (15/66/0)	30.1 years (22–59)	12/36/51%	37/44
Social Science	85 (39/44/2)	33.7 years (22–78)	13/42/45%	41/44
Philosophy	81 (56/20/5)	33.2 years (21–69)	19/40/41%	40/41
Total	580 (261/306/14)	30.5 years (21–78)	26/35/39%	289/291

then completed four additional components in a random order: Science Literacy, Reflective Thinking, Logical Syllogisms, and Perceptions of Science. Finally, participants responded to a set of demographic questions.

The Rating Explanations task used a 2 (Explanation Level: horizontal, reductive) × 2 (Quality: good, bad) × 6 (Science: physics, biology, chemistry, neuroscience, psychology, social science) design. Participants were randomly assigned to either the horizontal (n = 289) or reductive (n = 291) condition (see Table 1 for a breakdown by field of expertise). Quality and science were within-subjects variables: All participants rated two explanations from each science, one good and one bad.

2.3. Materials

The Rating Explanations task used 24 different phenomena (four per science) that described concepts, principles, or research findings from each of the six sciences. Each phenomenon had four corresponding explanations: horizontal-good, horizontal-bad, reductive-good, and reductive-bad. The process of creating and piloting these stimuli is described in more detail in Hopkins et al. (2016). All phenomena and their corresponding explanations, as well as details about how these stimuli were counterbalanced, are available in the online supplemental materials and via the Open Science Framework (https://osf.io/p8vft/; see Table 2 for an example).

The horizontal-good versions of the explanations were the commonly accepted explanations given by experts in the field. The horizontal-bad explanations were missing key information necessary to explain the phenomenon. Instead, they contained either circular restatements of the original phenomenon or irrelevant information. However, the information in the bad explanations was always factually correct. Therefore, participants could not use the accuracy of the explanation as a basis for their judgments about explanation quality. Horizontal explanations contained only information from the discipline of the phenomenon (i.e., psychological phenomena were explained only in psychological terms).

To construct the reductive explanations, we arranged our target disciplines in a reductive hierarchy: social science, psychology, neuroscience, biology, chemistry, and physics. The ordering of this hierarchy was based on ways that scientific fields relate to each other; that is, one can think of the study of society-level processes as reducing to the study of individuals or the study of the brain as reducing to the biology of brain cells. It is also consistent with research on the perceived relations between academic disciplines. A recent meta-analysis of 20 studies that used bibliographic measures to map the scientific landscape found a similar ordering (Klavans & Boyack, 2009). Similarly, participants in Hopkins et al. (2016) were asked to rate the prestige, scientific rigor, and knowledge gap between experts and novices in each of these fields. The results on a composite score of these three questions largely mirrored this hierarchy, with more reductive fields being rated more highly (i.e., more prestigious, more rigorous, and with a larger knowledge gap); the one exception was neuroscience, which was rated more highly than would be expected given its place in the hierarchy.

Reductive versions of the explanations contained information from the discipline below that of the phenomenon in the reductive hierarchy (bolded text in Table 2; this emphasis was not shown to participants in the study). For example, reductive explanations for biological phenomena contained reference to chemicals or chemical reactions. Reductive explanations for physical phenomena referred to smaller particles and/or more fundamental forces. The reductive information, although factually correct, was logically irrelevant and did not affect the quality of the explanation (verified by our expert consultants, none of whom participated in this study). Although some reductive explanations contained information not provided in the horizontal ones, this additional information was never explanatory, and this was not the case for all phenomena. Controlling for whether a particular item included additional, but irrelevant, information vs. a circular restatement of the phenomenon did not change any of the results reported here.

Table 2
Sample phenomenon and explanations from biology.

Male anole lizards bob their heads up and down rhythmically as part of a mating ritual to attract females. They typically increase their rate of head-bobbing when they see a female lizard of their species. However, their rate of head-bobbing also increases when they see another male lizard of the same species, even if no female lizards are present.
Why do male lizards bob their heads when other males are nearby?

	Good	Bad
Horizontal	This happens because the male lizards are extremely territorial, and head-bobbing is a distinctive behavior typical of this particular species of lizard. During mating season when they are in competition with each other for females, males use various dominance displays to defend their territory. They perceive other males as a threat and engage in increased head-bobbing, which is a sign of aggression.	This happens because the male lizards are seeking mates, and head-bobbing is a distinctive behavior typical of this species of lizard. During mating season when they are trying to attract females, males use a variety of behaviors that are characteristic of anole lizards. They perceive the presence of other males and engage in increased head-bobbing, which is commonly seen during mating season.
Reductive	This happens because the male lizards are extremely territorial. During mating season when they are in competition with each other for females, males use various dominance displays to defend their territory. They perceive other males as a threat and engage in increased head-bobbing, which is a sign of aggression. **Aggressive behavior is known to be associated with elevated levels of testosterone and other aggression-enabling hormones**.	This happens because the male lizards are seeking mates. During mating season when they are trying to attract females, males use a variety of behaviors that are characteristic of lizards. They perceive the presence of other males and engage in increased head-bobbing, which is commonly seen during mating season. **Aggressive behavior is known to be associated with elevated levels of testosterone and other aggression-enabling hormones**.

2.4. Procedure

2.4.1. Rating explanations

All participants completed the Rating Explanations task first. There were 12 experimental trials, with an attention check trial presented after the first six (Oppenheimer, Meyvis, & Davidenko, 2009). On each trial, participants read a description of a phenomenon from one of the six sciences. After 10 s, they were able to advance to the next screen where an explanation was displayed below the phenomenon. They were told to rate the quality of the explanation on a 7-point scale from −3 (Very poor) to +3 (Very good). The attention check trial was similar in format to the experimental trials, but the explanation contained explicit instructions for participants to select 3 on the scale. Participants who did not select 3 (37 in the horizontal condition, 42 in the reductive condition) were excluded from analyses. These participants were distributed across fields of expertise (5 from physics, 12 from chemistry, 6 from biology, 14 from neuroscience, 7 from psychology, 24 from social science, and 11 from philosophy) and levels of education (11 with some graduate school, 28 with master's degrees, and 40 with doctoral degrees). Including these participants in analyses did not change any of our primary results; the few places where their inclusion made a difference are noted.

After participants rated 12 explanations, the survey software randomly selected one item for which a participant had given a positive rating and one for which they had given a negative rating. For each, participants were asked to explain why they gave the rating they did and what (if any) additional information would have improved the explanation. They also answered multiple choice questions about whether reading the explanation changed their understanding of the phenomenon and whether they would like to change their initial rating of the explanation. These questions were included in the survey as part of a different project; data from these questions are available through the Open Science Framework, but will not be discussed here.

After the explanations task, participants completed four additional measures, presented in random order: Science Literacy (National Science Board, 2014), Reflective Thinking (Toplak, West, & Stanovich, 2014), Logical Reasoning (Hopkins et al., 2016), and Perceptions of Science (Fernandez-Duque et al., 2015). The complete set of questions for each of these measures are included in the online supplemental materials and via the Open Science Framework.

2.4.2. Demographics

At the end of the survey, participants answered a series of demographic questions, including gender, age, and highest degree completed. Participants were asked to pick the category that most closely matched the field of their highest degree (physical sciences, social sciences, engineering, humanities, health, and business), and to describe the exact field in a text box. They were additionally asked for their area of specialization and the number of years that they had been working in their field. They were finally asked whether they had taken any college- or graduate-level courses in anthropology, chemistry, physics, sociology, economics, neuroscience, psychology, political science, biology, or philosophy.

3. Results

The data from the Rating Explanations task were analyzed using mixed-effects regression predicting the rating given on each trial. All models included random intercepts by participant and item as well as a random effect of item on the slope for the quality variable.

We conducted three primary analyses. All three included explanation quality, explanation level and their interaction as predictors, but they differed in how the sample was divided into subgroups. The first set of analyses examined the performance of the experts in this study and compared it to the performance of the novices in Hopkins et al. (2016). In the second analysis, we subdivided trials from experts into categories based on how the expertise of the participant aligned with the field of the item they were rating. This allowed us to investigate whether expertise in a field conferred immunity to the reductive allure effect either within one's own field and/or more generally across all fields of science.

Table 3
Standardized regression coefficients for each sample.

Predictor	Novices[a]	Experts
Intercept	0.00 (0.05)	0.00 (0.05)
Quality	0.68 (0.06)***	1.05 (0.07)***
Explanation Level	0.12 (0.05)*	0.02 (0.03)
Quality × Explanation Level	−0.09 (0.06)	−0.10 (0.04)**

Note. Numbers in parentheses are the standard errors of the regression coefficients.
[a] These data were previously reported in Hopkins et al. (2016).
* $p < .05$.
** $p < .01$.
*** $p < .001$.

3.1. Comparing experts and novices

To examine whether expertise helps combat the reductive allure bias, we compared the expert participants recruited for this study (all had at least some graduate school) to the novice participants (undergraduate students and MTurk workers) from Hopkins et al. (2016). Hopkins et al. (2016) found that novices gave significantly higher ratings to good explanations ($M = 1.76$, $SD = 1.43$) than to bad explanations ($M = 0.53$, $SD = 1.93$), and they gave significantly higher ratings to explanations that contained irrelevant reductive information ($M = 1.26$, $SD = 1.71$) than to those that did not ($M = 1.04$, $SD = 1.88$). The interaction between quality and explanation level was not significant, indicating that ratings of both good and bad explanations were influenced by the allure of reductive information (Table 3).

The pattern for the expert participants in this study is somewhat different (Fig. 1). In a regression including fixed effects of quality, explanation level, and their interaction, experts also rated good explanations ($M = 1.53$, $SD = 1.56$) significantly higher than bad explanations ($M = -0.66$, $SD = 1.97$), but they did not show a significant overall difference between reductive and non-reductive explanations. However, there was a small but significant Quality × Explanation Level interaction: this indicates that, for experts, the addition of irrelevant reductive information had a larger impact on ratings of bad explanations ($M_R - M_H = 0.16$) than on ratings of good explanations ($M_R - M_H = -0.05$).

From these separate analyses, it appears as though experts show a larger effect of quality and smaller effect of reduction than novices. To determine whether expertise significantly moderated the strength of these effects, we conducted another regression analysis that included fixed effects of quality (good vs. bad), explanation level (horizontal vs. reductive), and sample (experts vs. novices) as well as all interactions between them; this analysis also controlled for education level (1 = less than high school/some high school, 2 = some college, 3 = 2-year degree/4-year degree, 4 = some graduate school, 5 = master's degree, 6 = PhD/MD/JD). This regression (Table 4) revealed a significant main effect of education: Participants with higher levels of education gave lower ratings overall.[3] However, even controlling for education level, experts' ratings indicated a larger differentiation between good and bad explanations than novices' ratings did, as indicated by the significant Quality × Sample interaction ($\beta = 0.47$, $SE = 0.03$, $p < .001$).

Although the two groups differed in their ability to detect bad

[3] The main effect of education was not significant when participants who failed the attention check were included in the analyses.

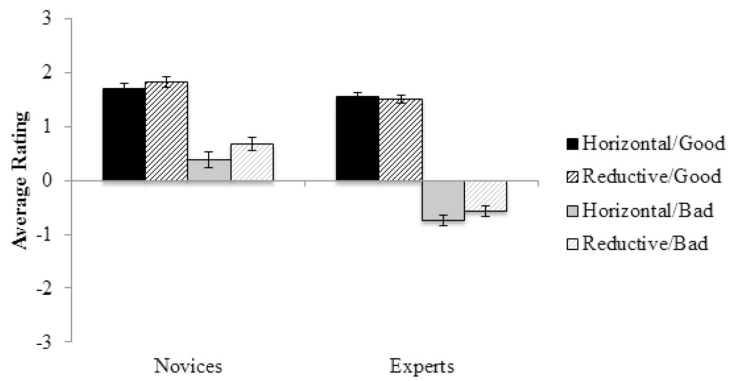

Fig. 1. Rating of explanations for both samples by quality and explanation level. Error bars are 95% confidence intervals.

Table 4
Explanation level × Quality × Sample Regression.

Predictor	β	*SE*	*t*
Intercept	0.20	0.06	3.46***
Education	−0.03	0.01	−2.30*
Quality	0.85	0.02	47.74***
Explanation Level	0.06	0.03	2.13*
Sample	−0.27	0.05	−5.71***
Quality × Explanation Level	−0.10	0.04	−2.74**
Quality × Sample	0.45	0.04	12.67***
Explanation Level × Sample	−0.07	0.06	−1.19
Quality × Explanation Level × Sample	−0.01	0.07	−0.20

* $p < .05$.
** $p < .01$.
*** $p < .001$.

explanations, there was no indication that they differed in their susceptibility to the reductive allure effect: The Explanation Level × Sample ($\beta = -0.08$, $SE = 0.05$, $p = .120$) and Quality × Explanation Level × Sample ($\beta = -0.02$, $SE = 0.07$, $p = .815$) interactions were not significant. Thus, although the groups showed different patterns of the impact of reductive information when analyzed separately, the differences between novices and experts were small and not statistically significant. The Quality × Explanation level ($\beta = -0.10$, $SE = 0.03$, $p = .004$) interaction was significant with this combined sample, further supporting the fact that experts and novices alike were influenced by reductive information in bad explanations.

3.1.1. *Analysis of auxiliary measures*

The prior analyses revealed differences between the novice and expert groups in the ability to differentiate good from bad explanations. Next, we examined the extent to which individual differences in education, logic, reflective thinking, and science literacy could explain these group differences. A difference score was computed for each participant representing the participant's average rating of the good explanations they viewed minus the average rating of the bad explanations they viewed. Thus, larger difference scores indicate that a participant was better able to differentiate good from bad explanations. When education, logic, reflective thinking, and science literacy were entered as predictors into a regression, only reflective thinking scores significantly predicted difference scores (Table 5): Participants with higher scores on the reflective thinking task showed larger differences between their ratings of good and bad explanations.[4]

[4] When participants who failed the attention check were included, science literacy scores also significantly predicted difference scores. This is likely because participants who failed the attention check scored significantly lower on the science literacy measure than participants who did not: $t(640) = 6.79$, $p < .001$.

Table 5
Effect of auxiliary measures on experts' difference scores on the explanations task.

Predictor	β	*SE*	*t*
Intercept	−1.35	0.73	−1.87+
Education	0.05	0.05	1.03
Reflective thinking	0.19	0.04	4.42***
Logic	0.01	0.04	0.32
Science literacy	0.03	0.05	0.66

\+ $p < .10$.
*** $p < .001$.

To further test our hypothesis that training in philosophy leads to improved detection of bad explanations, we examined whether science experts who had completed coursework in philosophy would outperform those who had not. Of the science expert participants, 56% reported taking at least one philosophy course as an undergraduate (evenly distributed across fields of expertise). Those who had taken undergraduate philosophy courses had significantly larger difference scores than those who had not: $t(480) = 2.55$, $p < 0.05$, $d = 0.23$.

Next, we examined whether any of these auxiliary measures moderated differences in ratings between groups. Sample was treated as a factor with four levels: MTurk workers, undergraduate students, science experts, and philosophy experts. The sample variable was backwards-difference coded such that each level was compared to the one prior to it. When difference scores were predicted from sample alone (Fig. 2, Table 6), there were significant differences between undergraduates and MTurk workers and between science experts and undergraduates. The difference between philosophy and science experts was marginally significant. However, when education was added to the model, the difference between undergraduates and science experts became non-significant. The same analysis was conducted for the other auxiliary measures (logical reasoning, reflective thinking, scientific literacy), but none had any impact on the effect of sample on difference scores.

3.2. *Analysis by type of expertise*

The prior analyses collapsed across phenomena and explanations from all sciences. To examine whether field-specific expertise affected expert participants' ratings, we next categorized each trial in terms of the relation between the field of the phenomenon and the participant's field of expertise (Table 7). Only the expert participants were included in this analysis. A trial was coded as ‘horizontal expertise’ ($n = 970$ trials) when a participant viewed a phenomenon that was from their field of expertise (e.g., a biology expert rating an explanation for a biological phenomenon). A trial was coded as ‘reductive expertise’ ($n = 810$ trials) when a participant viewed a phenomenon for which their field was the immediately reductive level (e.g., a biology expert rating an explanation for a neuroscience phenomenon), since the

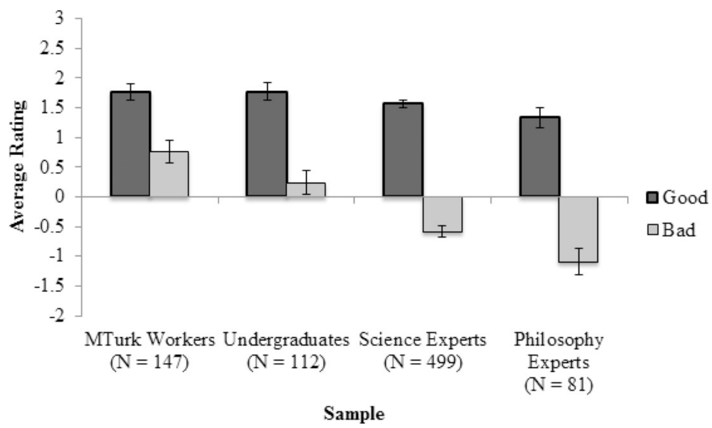

Fig. 2. Average ratings on the explanation task by sample and explanation quality. Error bars represent 95% confidence intervals.

Table 6
Standardized regression coefficients for model predicting difference scores.

Predictor	Model 1	Model 2
Education		0.11 (0.04)**
Undergraduates vs. MTurk workers	0.41 (0.12)***	0.47 (0.12)***
Science experts vs. undergraduates	0.48 (0.10)***	0.15 (0.15)
Philosophy experts vs. science experts	0.21 (0.11)+	0.20 (0.11)+

Note. Numbers in parentheses are standard errors of the regression coefficients.
+ $p < .10$.
** $p < .01$.
*** $p < .001$.

Table 7
Coding trials by type of expertise.

Participant field	Phenomenon					
	Phys	Chem	Bio	Neuro	Psych	Soc
Physicists	H	R	G	G	G	G
Chemists	G	H	R	G	G	G
Biologists	G	G	H	R	G	G
Neuroscientists	G	G	G	H	R	G
Psychologists	G	G	G	G	H	R
Social Scientists	G	G	G	G	G	H
Philosophers	L	L	L	L	L	L

Note. H = horizontal expertise, R = reductive expertise, G = general science expertise, L = logic expertise.

reductive information was always drawn from the immediately more reductive scientific field. All other trials from science experts were coded as general science expertise ($n = 4040$ trials), where the phenomenon and its corresponding explanations came from outside the participant's field (e.g., a biology expert rating a social science phenomenon). All trials from philosophy experts were coded as 'logic expertise' ($n = 972$ trials). Comparing 'logic expertise' and 'general science expertise' trials enables us to further examine the difference between training in logic and argumentation vs. any training in science. Examining the performance on horizontal and reductive expertise trials enables us to determine the effect of specific expertise in the field(s) from which the explanations were drawn.

The regression model predicting the rating given on a trial from the fixed effects of explanation level, quality, and expertise type is shown in Table 8. Expertise was simple-effects coded, such that the coefficient for each type of expertise represents the effect of that level compared to the reference level (general science expertise trials).

As in the prior analyses, there was a significant main effect of Quality and a significant Quality × Explanation Level interaction: All participants gave higher ratings to good explanations than to bad

Table 8
Explanation level × Quality × Expertise Regression.

Predictor	β	SE	t
Intercept	−0.05	0.05	−0.90
Quality	1.08	0.07	16.11***
Explanation Level	0.01	0.03	0.33
Expertise (logic)	−0.21	0.04	−5.13***
Expertise (horizontal)	−0.15	0.03	−5.79***
Expertise (reductive)	−0.06	0.03	−1.94+
Quality × Explanation Level	−0.12	0.04	−2.70**
Quality × Expertise (l)	0.16	0.05	2.99**
Quality × Expertise (h)	0.08	0.05	1.45
Quality × Expertise (r)	0.05	0.06	0.81
Explanation Level × Expertise (l)	−0.08	0.08	−0.99
Explanation Level × Expertise (h)	−0.09	0.05	−1.66+
Explanation Level × Expertise (r)	0.03	0.06	0.55
Quality × Explanation Level × Expertise (l)	0.08	0.11	0.72
Quality × Explanation Level × Expertise (h)	−0.04	0.11	−0.42
Quality × Explanation Level × Expertise (r)	−0.18	0.12	−1.54

+ $p < .10$.
** $p < .01$.
*** $p < .001$.

explanations, and the effect of quality was stronger in the horizontal compared to the reductive condition (Fig. 3). Participants gave lower ratings overall on any trials within their areas of expertise: Ratings were significantly lower on logic and horizontal expertise trials than on general science expertise trials, and ratings on reductive expertise trials were marginally lower than on general science expertise trials. Furthermore, the difference between good and bad explanations was significantly larger for philosophy experts (logic expertise trials) compared to general science expertise trials. This supports the idea that expertise either in a particular topic area or in logic leads to greater skepticism; training in logic further leads to a stronger differentiation between good and bad explanations.

Participants were somewhat less vulnerable to the reductive allure effect when they were rating explanations from within their own field compared to explanations from outside their field, as indicated by the marginally significant Explanation Level × Expertise interaction for horizontal expertise trials: On horizontal expertise trials, participants actually gave lower ratings to reductive ($M = 0.17$, $SD = 2.17$) than horizontal explanations ($M = 0.31$, $SD = 2.16$). In contrast, participants gave higher ratings to reductive than horizontal explanations on general science expertise ($M_R = 0.61$; $M_H = 0.51$) trials.

4. Discussion

Previous research has discovered a *reductive allure effect*, whereby scientific explanations are rated as better when they contain irrelevant

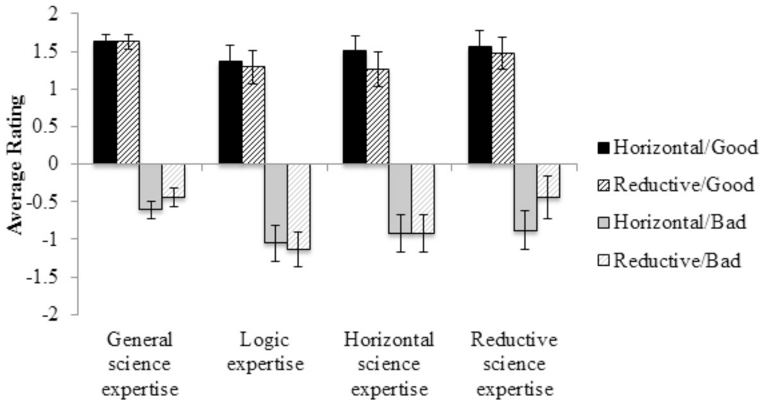

Fig. 3. Average ratings of explanations by explanation level, quality, and type of expertise. Error bars are 95% confidence intervals.

information from the neighboring reductive field (Hopkins et al., 2016). The main goal of this study was to explore the effect of expertise in order to determine whether advanced training, either in a specific field of science or in philosophy, could help to combat this reasoning error.

We first examined whether training in any field of science or in philosophy could help one to avoid this bias for reductive information in general. Although separate analyses suggested that experts may be slightly less susceptible to reductive information than novices, the difference between samples was not significant when they were analyzed together, suggesting that expertise does not inoculate against the reductive allure bias.

The groups did differ in their ratings of explanation quality: Although all groups were able to tell good from bad explanations, the difference between ratings of good and bad explanations was larger for undergraduates than for MTurk workers, and for science experts than for undergraduates; philosophy experts had marginally larger difference scores than science experts. Controlling for educational attainment reduced the difference between science experts and undergraduates, suggesting that the difference between these two groups is largely quantitative – more education leads to better discrimination of good and bad explanations.

However, controlling for education did not affect the difference between undergraduates and MTurk workers or the difference between philosophy and science experts, suggesting the differences between these groups may be more qualitative in nature. As discussed in the introduction, undergraduates may differ from MTurk workers because the fact of being currently immersed in an educational setting may increase skepticism or prime critical thinking. This is further supported by the fact that science experts gave marginally lower ratings to explanations within their area of expertise; being regularly immersed in a subject area may lead to a more critical eye when evaluating information from that subject area (consistent with Weisberg et al., 2008, Study 3).

Training in philosophy that focuses on logic and argumentation may lead to a more critical eye across a variety of explanations than training in science that focuses on a particular content area. Philosophers made significantly larger distinctions between good and bad explanations than scientists who were rating explanations for which they had no specific expertise. This conclusion was further supported by the fact that the science experts who had taken courses in philosophy had better discrimination between good and bad explanations than those who had not. However, the difference between philosophers and science experts overall was only marginally significant, and we cannot know the extent to which our science experts received training in logic and argumentation in the course of their science education. It is possible that the difference between these groups is small because the amount of domain-general training received by science experts is more similar to philosophical training than we thought. Future research should examine more closely the prevalence of this type of training and the effect it has on judgments of explanations.

Although expertise in general did not appear to inoculate against the allure of reductive explanations, we found that expert participants were somewhat less susceptible to the allure of reductive information when they were rating phenomena from within their own field. Their expertise may have allowed them to be more critical about explanations that contained information they were highly familiar with and to avoid being influenced by the logically irrelevant reductive information. This effect, although only marginally significant here, is consistent with Weisberg et al. (2008, Study 3), in which neuroscience experts rated explanations with neuroscience information significantly less highly than novices did.

It is possible that individuals with different types of expertise have different ideas about what it means for an explanation to be reductive. Reduction, as we have defined it here, involves providing an explanation for a phenomenon from one scientific discipline using the language and concepts of a more fundamental science (Kemeny & Oppenheim, 1956; Oppenheim & Putnam, 1958). Numerous theories exist regarding exactly when and how events from one field can be successfully reduced to components from another (Nagel, 1961), and even whether reduction is metaphysically possible for any or all sciences (for a review, see van Riel, 2014). Given this disagreement, it is possible that our participants also had different ideas about when and how reduction can be successful. It may also be the case that different disciplines encourage specific ideas about the use and value of reduction.

These data show that advanced training in either science or philosophy strengthens the ability to detect bad explanations. More content knowledge, experience evaluating arguments and proof, and greater understanding of the scientific method likely all contribute to this ability. Although advanced training in science did not confer any general immunity against the allure of irrelevant reductive information, there was a hint that some types of training may help avoid this bias: Participants were marginally less swayed by reductive information when they were evaluating explanations from their own field. Thus, it is possible that domain-specific expertise can help people detect the irrelevant information and avoid being persuaded by it, although more work is needed to test this possibility.

These results demonstrate that general training in logic and analysis, as well as specific training in a scientific discipline, can affect what kinds of explanations one considers to be satisfying. The current study thus provides an important case study into the complex effects of attaining expertise in a field: Doing so may protect against some types of erroneous judgments, but not all. To examine these issues further, future work should address exactly which types of training could be effective at changing individuals' explanatory values, as well as how much training is necessary. It may be that acquiring additional content knowledge within a field improves one's ability to filter out irrelevant information or that developing deeper knowledge of the mechanisms behind certain phenomena makes one more critical of spurious explanations. These aspects could then be translated into teaching strategies at the undergraduate and graduate level, allowing students to avoid falling prey to the reductive allure effect.

Declaration of Competing Interest

None.

Acknowledgements

This research was funded by the John Templeton Foundation through the *Varieties of Understanding* project. We would like to thank Don Berry, Martha Farah, Simon Garcia, Cindy Keeler, Phil Nelson, Sam Reznik, and Paul Sneigowski for their assistance with stimulus creation.

References

DeGroot, A. D. (1965). *Thought and choice in chess.* New York: Basic Books.

Fernandez-Duque, D., Evans, J., Christian, C., & Hodges, S. D. (2015). Superfluous neuroscience information makes explanations of psychological phenomena more appealing. *Journal of Cognitive Neuroscience, 27*(5), 926–944. https://doi.org/10.1162/jocn_a_00750.

Fisher, M., & Keil, F. C. (2015). The curse of expertise: When more knowledge leads to miscalibrated explanatory insight. *Cognitive Science, 40*, 1251–1269. https://doi.org/10.1111/cogs.12280.

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3*(2), 191–197. https://doi.org/10.1038/72140.

Goldberg, R. F., & Thompson-Schill, S. L. (2009). Developmental “roots” in mature biological knowledge. *Psychological Science, 20*(4), 480–487. https://doi.org/10.1111/j.1467-9280.2009.02320.x.

Hopkins, E. J., Weisberg, D. S., & Taylor, J. C. V. (2016). The seductive allure is a reductive allure: People prefer scientific explanations that contain logically irrelevant reductive information. *Cognition, 155*, 67–76. https://doi.org/10.1016/j.cognition.2016.06.011.

Kelemen, D., Rottman, J., & Seston, R. (2012). Professional physical scientists display tenacious teleological tendencies: Purpose-based reasoning as a cognitive default. *Journal of Experimental Psychology: General, 142*(4), 1074–1083. https://doi.org/10.1037/a0030399.

Kemeny, J. G., & Oppenheim, P. (1956). On reduction. *Philosophical Studies, 7*(1–2), 6–19.

Kikas, E. (2003). University students' conceptions of different physical phenomena. *Journal of Adult Development, 10*(3), 139–150. https://doi.org/10.1023/A:1023410212892.

Klavans, R., & Boyack, K. W. (2009). Toward a consensus map of science. *Journal of the American Society for Information Science and Technology, 60*(3), 455–476. https://doi.org/10.1002/asi.

Langer, E. J., Blank, A., & Chanowitz, B. (1978). The mindlessness of ostensibly thoughtful action: The role of “placebic” information in interpersonal interaction. *Journal of Personality and Social Psychology, 36*(6), 635–642. https://doi.org/10.1037/0022-3514.36.6.635.

Lawson, R. (2006). The science of cycology: Failures to understand how everyday objects work. *Memory & Cognition, 34*(8), 1667–1675. https://doi.org/10.3758/BF03195929.

Masson, S., Potvin, P., Riopel, M., & Foisy, L. M. B. (2014). Differences in brain activation between novices and experts in science during a task involving a common misconception in electricity. *Mind, Brain, and Education, 8*(1), 44–55. https://doi.org/10.1111/mbe.12043.

Nagel, E. (1961). *The structure of science: Problems in the logic of explanation.* New York: Harcourt, Brace & World, Inc.

National Science Board (2014). *Science and engineering indicators.* Arlington, VA: National Science Foundation.

Oppenheim, P., & Putnam, H. (1958). Unity of science as a working hypothesis. In H. Feigl, M. Scriven, & G. Maxwell (Vol. Eds.), *Minnesota studies in the philosophy of science. Vol. 2. Minnesota studies in the philosophy of science* (pp. 3–36). Minneapolis, MN: University of Minnesota Press.

Oppenheimer, D. M., Meyvis, T., & Davidenko, N. (2009). Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology, 45*(4), 867–872. https://doi.org/10.1016/j.jesp.2009.03.009.

Rhodes, R. E., Rodriguez, F., & Shah, P. (2014). Explaining the alluring influence of neuroscience information on scientific reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(5), 1432–1440. https://doi.org/10.1037/a0036844.

Shtulman, A. (2015). How lay cognition constrains scientific cognition. *Philosophy Compass*, 1–14. https://doi.org/10.1111/phc3.12260.

Shtulman, A., & Harrington, K. (2015). Tensions between science and intuition across the lifespan. *Topics in Cognitive Science, 8*, 118–137. https://doi.org/10.1111/tops.12174.

Slotta, J., Chi, M., & Joram, E. (1995). Assessing students' misclassifications of physics concepts: An ontological basis for conceptual change. *Cognition and Instruction, 13*(3), 373–400. https://doi.org/10.1207/s1532690xci1303_2.

Stanovich, K. E., West, R. F., & Toplak, M. E. (2013). Myside bias, rational thinking, and intelligence. *Current Directions in Psychological Science, 22*(4), 259–264. https://doi.org/10.1177/0963721413480174.

Stokes, D. R. (2018). On perceptual expertise. Retrieved from https://philpapers.org/archive/STOOPE.pdf.

Toplak, M. E., West, R. F., & Stanovich, K. E. (2014). Assessing miserly information processing: An expansion of the Cognitive Reflection Test. *Thinking & Reasoning, 20*(2), 147–168. https://doi.org/10.1080/13546783.2013.844729.

van Riel, R. (2014). *Conceptions of reduction in the philosophy of science. The concept of reduction. Philosophical studies series*, *Vol. 121*. Cham: Springer.

Weisberg, D. S., Keil, F. C., Goodstein, J., Rawson, E., & Gray, J. R. (2008). The seductive allure of neuroscience explanations. *Journal of Cognitive Neuroscience, 20*(3), 470–477. https://doi.org/10.1162/jocn.2008.20040.

Weisberg, D. S., Taylor, J. C. V., & Hopkins, E. J. (2015). Deconstructing the seductive allure of neuroscience explanations. *Judgment and Decision making, 10*(5), 429–441. https://doi.org/10.1162/jocn.2008.20040.

Zemla, J. C., Sloman, S. A., Bechlivanidis, C., & Lagnado, D. A. (2017). Evaluating everyday explanations. *Psychonomic Bulletin & Review, 24*(5), 1488–1500. Advance online publication https://doi.org/10.3758/s13423-017-1258-z.