Anomalous Evidence, Confidence Change, and Theory Change

Joshua A. Hemmerich, a Kellie Van Voorhis, a Jennifer Wiley b

a Department of Medicine, The University of Chicago
b Department of Psychology, The University of Illinois at Chicago

Received 13 December 2012; received in revised form 18 May 2015; accepted 24 May 2015

Abstract

A novel experimental paradigm that measured theory change and confidence in participants’ theories was used in three experiments to test the effects of anomalous evidence. Experiment 1 varied the amount of anomalous evidence to see if “dose size” made incremental changes in confidence toward theory change. Experiment 2 varied whether anomalous evidence was convergent (of multiple types) or replicating (similar finding repeated). Experiment 3 varied whether participants were provided with an alternative theory that explained the anomalous evidence. All experiments showed that participants’ confidence changes were commensurate with the amount of anomalous evidence presented, and that larger decreases in confidence predicted theory changes. Convergent evidence and the presentation of an alternative theory led to larger confidence change. Convergent evidence also caused more theory changes. Even when people do not change theories, factors pertinent to the evidence and alternative theories decrease their confidence in their current theory and move them incrementally closer to theory change.

Keywords: Theory change; Confidence change; Conceptual change; Anomalous evidence; Theory confidence

1. Introduction

Human beings, young and old, expert and novice, use theories to make sense of the wide array of empirical evidence that they encounter on a day-to-day basis. Often times, however, we are confronted by evidence that does not fit our current theories. When prior knowledge is at odds with newly encountered information, conceptual change is required in order to best fit and integrate the new knowledge, as opposed to modes of learning that involve adding more knowledge to existing frameworks. How anomalous evidence can

Correspondence should be sent to Jennifer Wiley, Department of Psychology, University of Illinois at Chicago, 1007 W Harrison St MC 285, Chicago, IL 60647. E-mail: jwiley@uic.edu
cause people to change their explanatory mental models of phenomena has long been a central question in cognitive science and in research on conceptual change (Chinn & Brewer, 1992), and numerous possible pathways to conceptual change that involve different cognitive mechanisms have been proposed (Carey, 1988; Chi, 2013; Chinn & Samarpungavan, 2009; Duit, Treagust, & Widodo, 2013; Gentner et al., 1997; Thagard, 1992; Vosniadou, 2013).

The specific component of conceptual change that is explored in the current studies is *theory change*. These studies investigate when encounters with the new data will make people stop believing in their accepted theory and start believing in a different one that can better explain the newly encountered anomalous evidence (Chinn & Brewer, 1992). *Theory change* is defined in our experimental paradigm as occurring when one abandons a previously accepted theory and replaces it with another to explain the same causal phenomenon. It is not the accumulation of factual information but the adaptation of beliefs toward a mental model that fits empirical observations (diSessa & Sherin, 1998).

Here, the definition of an *accepted theory* is a person’s current mental model or causal explanation for a phenomenon. As a mental model, it can be “run” to make predictions about outcomes and changes in one variable based on others. Among college-aged adults, such a theory can be thought of as a cohesive mental model and not a collection of fragmented ideas (Gopnik & Wellman, 1994; Vosniadou, 2002). Being consistent with empirical evidence is an epistemic defining characteristic of a good theory (Pluta, Chinn, & Duncan, 2011). *Anomalous evidence* are data which would not be predicted by, and are inconsistent with, a person’s mental model (Chinn & Brewer, 1993a). People can encounter anomalous evidence as part of educational instruction or formal interventions, but also through everyday encounters with new information. When people encounter empirical evidence that contradicts what their accepted theory about a causal relationship would predict, this discrepancy must be resolved. For change to occur people must recognize that the new evidence conflicts with their current understanding. In such instances, it is possible that people will make a lateral move to another more coherent theory, but this does not always happen (Chi, 2013; Guzzetti, Snyder, Glass, & Gamas, 1993; Thagard, 1992). Sometimes people instead maintain belief in their theory in the face of evidence that contradicts its predictions (Chinn & Brewer, 1998).

Previous research has demonstrated that theory change is only one of several possible outcomes in the process of anomalous evidence resolution (Chinn & Brewer, 1992, 1993a,b, 2000). A documented taxonomy of strategies to deal with anomalous evidence (Chinn & Brewer, 1998) includes (a) ignoring, (b) rejecting, (c) professing uncertainty about its validity, (d) exclusion from the domain of one’s theory, (e) holding in abeyance, (f) reinterpretation, (g) accepting it and making peripheral theory changes, and (h) accepting it and completely changing to another theory. Only two of these strategies allow for anomalous evidence to cause any sort of overt change to an individual’s accepted theory (Chinn & Brewer, 1998). Many other studies also suggest that anomalous evidence sometimes has no effect on adults, children, or experts (Chan, Burtis, & Bereiter, 1997; Chi, Feltovich, & Glaser, 1981; van Dijk & Zeelenberg, 2003; Kuhn,
At other times people might make errors interpreting the correlative pattern of data, fail to make valid comparisons, and rely on biased heuristics that maintain their original explanation (Amsel & Brock, 1996; Chinn & Malhotra, 2002; Schauble, 1990; Schulz, Goodman, Tenenbaum, & Jenkins, 2008).

Yet, on some occasions profound changes in theories can be seen to occur. Despite the apparently conservative bias toward retaining one’s theory, theory change in response to anomalous evidence is a well-documented practice in numerous scientific communities and cultures (Kuhn et al., 1988; Piaget, 1974; Posner, Strike, Hewson, & Gertzog, 1982). Scientists and other individuals advance their understanding by acquiring new evidence and adjusting or overturning theories that no longer seem to be a good formulation of what is empirically observed. At the same time, the influence of anomalous evidence might be more incremental than the documented taxonomy of responses might suggest. An apparent failure of anomalous evidence to elicit theory change does not mean it has made no impact on the person’s belief in her theory, nor is it necessarily due to faulty judgment or stubborn refusal to accept that one is wrong. According to Lakatos, a philosopher of science and mathematics, scientists explicitly “red flag” anomalous findings to set them aside and attend to them later (Lakatos, 1970). This helps to maintain stability in scientific research by highlighting surprising discoveries to more strategically focus work to resolve them in the future (Kulkarni & Simon, 1988). Sufficient decreases in theory confidence may be necessary to reach the point at which one makes an explicit theory change, while smaller decreases could be insufficient. Conceptual change, and specifically theory change, is a latent construct that cannot be directly measured, but is effectively observed only through self-report and other quantifiable correlates (Kulikowich, 2007). Measuring the more incremental effects of evidence on confidence is perhaps necessary to more clearly observe the gradual processes that may underlie theory change.

Understanding the gradual nature of theory change in response to anomalous evidence requires experimental methods and measures that are appropriately sensitive. Even when people do not exhibit a theory change, it is possible that their theory confidence, that is their degree of certainty that their theory provides an accurate causal explanation for a phenomenon, has been impacted by their encounter with anomalous evidence. Typically, theory confidence should be negatively impacted by anomalous evidence, but these effects would only sometimes be sufficient to cause theory change. Likewise, these effects might make theory change more likely later when more anomalous evidence is encountered. Thus, the three experiments presented in this article focus on testing the conditions under which adults change from one coherent explanatory theory to another, using an experimental paradigm that includes measures of confidence change, in addition to binary measures of theory adherence. The examination of confidence change allows for better sensitivity for detecting the effects of anomalous evidence. In addition, several other factors that may contribute to theory change, including type and amount of evidence encountered, as well as the presence of an alternative theory, are also manipulated. The sections below outline why these factors are important to consider.
1.1. Perception of anomalous evidence

First and foremost, anomalous evidence must be recognized as being problematic for one’s accepted theory before any effort is made to resolve the discrepancy (Chinn & Brewer, 2001). People can distort the meaning of evidence to support their own hypotheses (Dunbar, 1993), or engage in biased confirmatory testing procedures when working to acquire evidence (Klayman & Ha, 1987). This effect is well documented in the context of belief polarization stemming from confirmation bias (Kelly, 2008). Even when anomalous evidence can be accurately recalled, there can be a failure to recognize that the data contradict one’s theory (Kuhn et al., 1988). Some researchers suggest that people often reject data that contradict their theories because they are biased toward scrutinizing the data more severely to find flaws in it (Edwards & Smith, 1996; Kunda, 1990).

It is likely that perceiving the anomalous evidence as convincing is necessary to elicit theory change (Chinn & Brewer, 1993a). In one study, belief was only decreased in an accepted theory about the causes of dinosaur extinction when participants determined that anomalous evidence was both valid and inconsistent with their accepted theory (Chinn & Brewer, 1993b; Mason, 2000). Some categories of the Chinn and Brewer taxonomy of responses to anomalous evidence suggest that the step at which people judge whether or not evidence is convincing is an important and even necessary antecedent to change. In each of the following experiments, participants are asked to judge the convincingness of the anomalous evidence to understand the role of such perceptions.

Additionally, the amount of anomalous evidence that one encounters is likely to have consequences for theory change, as sufficiently large amounts of anomalous information are probably more persuasive than much smaller amounts. Perhaps, if enough anomalous evidence accumulated in a sufficiently short time, it could cause incremental changes in theory confidence, pushing one closer and closer to making a theory change. To test these hypotheses, the amount of anomalous evidence given to participants was manipulated in the three experiments presented in this article.

1.2. Replicating versus converging anomalous evidence

When people encounter empirical data or evidence of some kind, they are thought to construct a model that integrates many elements of that evidence (Chinn & Brewer, 2001). Surprisingly, there is noticeably less literature on how people evaluate data than on how they evaluate theory, as both would seem equally important to conceptual change research. There is substantial research on how people evaluate covariation data, but it does not account for the possibility of questioning the quality of the data on methodological grounds (Amsel & Brock, 1996; Kuhn et al., 1988; Shaklee & Elek, 1988). One important consequence is that data are implicitly assumed to be valid in this literature. This is perhaps not a reasonable assumption and clouds the understanding why people often avoid making theory changes when they encounter anomalies (Chinn & Brewer, 2001).

The extent to which anomalous evidence can be disregarded or blocked is a major factor that could determine whether theory change or confidence change occurs. If new
anomalous evidence of the same exact methodological type, or replicating evidence, as that which was already resolved is encountered, then it would be efficient to use the same strategy to dismiss the new anomalous evidence that was successful the first time.

A different result may be seen when a body of anomalous evidence, defined as convergent evidence, is composed of methodologically different types of evidence that all contradict a specific theory. Individually, one piece of evidence might not cause an impact because it can be ignored, rejected, viewed with uncertainty, excluded from domain relevance, held in abeyance, or reinterpreted. But anomalous evidence may become more difficult to defend against, and more difficult to dismiss, when data from multiple methodological types challenge the person’s theory. By manipulating evidence type (replicating vs. convergent), this hypothesis is tested in the present work.

1.3. Availability of an alternative theory

Theory change might be more likely to occur when there is a combination of dissatisfaction with one’s accepted theory and a readily available alternative theory (Davies, 1997; Lord, Ross, & Lepper, 1979; Posner et al., 1982; Thagard, 1992). Without such an alternative theory, one is potentially left to identify or construct an appropriate mental model to explain the anomalous evidence. In much of the conceptual change literature, learning is framed as overcoming a misconception with an accurate mental model. However, when people must find a new theory, this presents another barrier to theory change occurring, even when their confidence in their own theory has decreased. Consequently, being provided with a coherent competing theory that would fit the anomalous evidence and organize it into a compelling explanation against one’s original theory should facilitate decreases in confidence and lead to more theory changes (Chi, 2008; Chinn & Brewer, 2001; Koslowski, Marasia, Chelenza, & Dublin, 2008; Posner et al., 1982; Thagard, 1992).

By clearly providing the alternative theory, people only need to assess the anomalous evidence and decide if it justifies changing to the alternative that they have just been reminded of, so it could be expected that presenting an alternative theory should increase the rate of theory change. This hypothesis is also tested in the research reported here.

1.4. Hypotheses

Three primary hypotheses were tested in this set of experiments. First, Experiment 1 set out to test the Incremental Change Hypothesis, which states that the amount of anomalous evidence that people are exposed to will have an incremental effect on participants’ confidence in their accepted theory. Participants who exhibit larger decreases in confidence in their accepted theory will be more likely to indicate that they have changed theories.

Experiment 2 was designed specifically to test the Convergent Evidence Hypothesis. The Convergent Evidence Hypothesis states that participants who are exposed to convergent anomalous evidence will decrease theory confidence more, and make more theory changes, than participants who are given replicating anomalous evidence.
Experiment 3 was designed to test the Alternative Theory Hypothesis that predicts that participants who are explicitly presented with an alternative theory that provides a clear explanation for the anomalous evidence will experience larger decreases in theory confidence and more theory change.

2. General method

All three experiments utilized the same basic procedures and any differences will be described below. All methods and procedures were approved by the Institutional Review Board.

2.1. Participants

Participants were recruited from a psychology department subject pool based on their responses to mass-testing questionnaires administered at the beginning of the semester. To qualify for the experiment, respondents had to provide answers that expressed adherence to one theory and nonadherence to the competing theory in one of three domains. The three domains were relationships, athletic skill, and cancer with the competing theories: birds of a feather versus opposites attract (relationships), natural born athlete versus practice makes perfect (athletics), and genetically inherited cancer versus cancer caused by environmental carcinogens (cancer). Participants indicated their belief in each theory via categorical questions as well as their degree of confidence in each theory on a 100% scale. They were asked to answer the same questions again when they were presented to the experiment. Only participants who were consistent in their responses between mass testing and the experimental session pretest were included in the analyses presented here (14% of participants were dropped for inconsistency).

2.2. Materials and design

The materials used in all experiments consisted of theory texts (Appendix A) and evidence packets of brief passages describing empirical relationships (Appendix B, see full materials in the Supplementary Materials file). Each theory text was a brief passage under 600 words that described the base assertions of the theory. The evidence packets began with the statement: “The evidence that you read about in this experiment might or might not be familiar to you. All of the information was taken from popular media and publications that you may read during the course of your academic studies or leisure.”

The evidence packets contained 12 brief (3–5 sentences) evidence passages describing an individual case, survey, or controlled study which were edited from actual published media reports. All specific identifying information was removed so that factors pertinent to the quality of the information source did not influence participants. Pilot testing of these materials demonstrated no significant differences between these three types of evidence in the level of convincingness raters perceived. Each evidence passage was either
supportive or not supportive of the participant’s accepted theory and the ratio of anomalous versus supportive evidence, referred to here on as anomalous evidence ratio, in the evidence packet was varied between subjects. For all studies, there were five levels of anomalous evidence ratio, ranging from five to nine pieces of anomalous evidence, from five to one pieces of supportive evidence, and each including two neutral passages that were theoretically tangential (5-5-2, 6-4-2, 7-3-2, 8-2-2, and 9-1-2).

A between-subjects design was utilized in all three experiments. Each participant in these three experiments received materials for only one domain (relationships, athletics, and cancer). They were also randomly assigned to one of the five levels of anomalous evidence ratio. Experiment 1 utilized only replicating evidence so that each passage was a specific instance of the same basic finding. Experiment 2 randomly assigned participants to receive either replicating or convergent evidence packets, which will be described under Experiment 2 methods. Experiment 3 used only convergent evidence packets, but had an additional manipulation of providing the competing theory to those randomly assigned to that condition which will be described under Experiment 3 methods.

2.3. Procedure

The experiments were conducted in small group sessions of 5–10 participants. After completing the informed consent process, participants were told to read the instructions and begin the experiment as soon as they were ready. Participants were instructed to read a theory text and indicated whether or not they believed the theory. Although participants’ attention was not drawn to it, the theory text represented a theory they had indicated belief in on the mass-testing questionnaire administered at the beginning of the academic semester. They rated their confidence in the theory on a 100% scale and then went on to read the evidence packet, which included 12 short-evidence passages. They were randomly assigned to receive one of five ratios of anomalous-to-supportive evidence. After reading all passages, participants were asked to judge if the total body of evidence (a) supported, (b) contradicted, or (c) provided equal support and nonsupport to their theory. They also separately rated the convincingness of the evidence that supported their theory and the evidence they perceived as anomalous on a 7-point scale. Finally, participants again answered the questions about which theory they believed, and rated their confidence in their theory on the 100% scale. When each experimental session was completed, participants were thanked and debriefed.

From these measurements, three key dependent variables were derived and were used in the analyses of all three experiments.

1. Theory change—the presence of an explicit theory change, as indicated by a difference in answers about theory belief before and after reading the evidence packet. This kind of explicit theory change, from here on referred to as “theory change,” is a dichotomous (yes vs. no) outcome and represents when a participant’s indicated belief in the theory changed from the one time to another.
2. **Confidence change**—the change in participants’ theory confidence ratings from before to after reading the evidence packet. This variable, referred to here on as “confidence change,” is defined as the incremental changes in confidence in one’s theory as rated on a 100% scale separate from indications of belief. Confidence change was computed by subtracting the pre-reading confidence rating from the postreading confidence rating—negative values indicate a decrease in confidence.

3. **Differential convincingness**—of interest was the difference between participants’ convincingness ratings for supporting and anomalous evidence. This variable was computed by subtracting the overall convincingness score of the anomalous evidence from the overall convincingness score of the supportive evidence. Results were scaled so that higher values for differential convincingness indicate that anomalous evidence was more convincing relative to supportive evidence.

3. **Experiment 1**

Experiment 1 tested the *Incremental Change Hypothesis* that although theory change might be infrequent when participants are confronted with anomalous evidence, the influence can be observed as changes in theory confidence. A prevailing explanation for the failure to observe theory change is the assumption that anomalous evidence is wholly ignored or dismissed. However, the *Incremental Change Hypothesis* suggests that there is some influence of anomalous evidence, which will result in changes, typically decreases, in the confidence that participants have in their theory. If participants do encounter and process anomalous evidence, then varying the anomalous evidence ratio should alter participants’ theory confidence, even when failing to elicit theory change. One specific prediction consistent with this hypothesis is that decreases in theory confidence will be more common than actual theory change. Second, it is expected that the anomalous evidence ratio will be predictive of the decreases in theory confidence and theory change. Finally, decreases in theory confidence should be predictive of theory change. It is also predicted that the more convincing that participants view the anomalous evidence relative to the supportive evidence, as measured in the *differential convincingness*, the more participants will make theory changes and decrease in confidence in their theories.

3.1. **Method**

The sample of 107 participants included 49 in the relationships domain, 31 in the athletics domain, and 27 in the cancer domain.

3.2. **Results**

A manipulation check on the five-level factor of anomalous evidence ratio showed that the ratio affected the frequency of participants who reported that they *perceived nonsupport* for their accepted theory, $\chi^2(8) = 32.41$, $p < .001$. Across the five ascending levels
of anomalous evidence ratio, the percentage of participants indicating the evidence set overall did not support their theory was 11% (2), 19% (5), 57% (12), 60% (12), and 82% (15). This result suggests that amount of anomalous evidence matters as comparisons made between anomalous evidence ratio groups show that participants who got larger amounts of anomalous evidence tended to rate the complete body of evidence as more anomalous to their theories.

3.2.1. Theory change

The first important finding of Experiment 1 is that only 13% of the total subjects exhibited theory change, but that 54% exhibited decreases in theory confidence, which a McNemar test showed to be a significant difference, \( p < .001 \). These results support the Incremental Change Hypothesis by demonstrating that theory confidence change can be much more prevalent than theory change when participants are confronted with anomalous evidence.

Anomalous evidence ratio did not significantly affect the likelihood of theory change, 9% (2), 15% (4), 9% (2), 25% (5), and 6% (1), \( \chi^2(4) = 4.06, p = .40 \), nor were differences observed between the relationship, 14% (7), athletics 7% (2), and cancer domains 19% (5), \( \chi^2(2) = 1.96, p = .38 \).

3.2.2. Confidence change

The Incremental Change Hypothesis was further supported by the effects found for anomalous evidence ratio on confidence change (see Table 1). A 3 \( \times \) 5 (domain by anomalous evidence ratio) analysis of variance (ANOVA) showed that, as predicted, anomalous evidence ratio significantly impacted confidence change, \( F(4, 106) = 3.34, p = .01 \), such that decreases in theory confidence differed across levels of anomalous evidence ratio. The linear trend test was statistically significant, \( t(102) = 3.42, p = .001 \). There was no significant effect for domain, \( F(2, 106) = 1.04, p = .36 \). However, as shown in Fig. 1, the domain by anomalous evidence ratio interaction was significant, \( F(8, 106) = 2.12, p = .04 \), indicating that the confidence change differs between the domains across the five levels of anomalous evidence ratio. Examining the nature of this interaction in closer detail, there may be two reasons for this interaction. First, the athletics domain seems to show the least effect of anomalous evidence ratio. Second, participants in the 8-2 ratio in the cancer domain showed no decrease in confidence. This interaction may be spurious as it was not observed in the other experiments.

Table 1

Analysis of variance results for Experiment 1 (theory confidence change)

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>F</th>
<th>p</th>
<th>( \eta^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain (D)</td>
<td>2</td>
<td>1.043</td>
<td>0.356</td>
<td>0.023</td>
</tr>
<tr>
<td>Anomalous evidence ratio (AER)</td>
<td>4</td>
<td>3.340</td>
<td>0.013*</td>
<td>0.129</td>
</tr>
<tr>
<td>D ( \times ) AER</td>
<td>8</td>
<td>2.119</td>
<td>0.042*</td>
<td>0.159</td>
</tr>
</tbody>
</table>

Note. *Denotes \( p \) value significant at \( p > .05 \).
The predicted relationship between confidence change and theory change was found, $r_{pb} = -0.35, p < .001$, demonstrating that bigger decreases in confidence were more predictive of theory change, and providing additional support for the **Incremental Change Hypothesis**.

### 3.2.3. **Differential convincingness**

Examination of the participants’ convincingness ratings showed that supporting evidence ($M = 6.92, SD = 1.54$) was rated as more convincing than anomalous evidence ($M = 5.60, SD = 1.84$), $t(106) = 5.46, p < .001$. An ANOVA showed that neither domain, nor anomalous evidence ratio, nor their interaction had a significant effect on differential convincingness, $F_s < 1$ (see Table 2).

As expected, differential convincingness was correlated with theory change, $r_{pb} = 0.20$, $p = 0.044$. There was also a significant correlation between differential convincingness and confidence change, $r = -0.25$, $p = 0.014$, such that the larger the difference in convincingness ratings, the larger the decrease in theory confidence. The correlation between differential convincingness and confidence change remained significant when tested in a partial correlation that controlled for anomalous evidence ratio $r_{ab,c} = -0.21$, $p = 0.04$, suggesting that differential convincingness exerts a direct influence on confidence change.

### 3.3. **Discussion**

Findings from Experiment 1 supported the **Incremental Change Hypothesis** as participants demonstrated theory change relatively infrequently, while over half of the sample

---

**Table 2**

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
<th>$F$</th>
<th>$p$</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain (D)</td>
<td>2</td>
<td>0.443</td>
<td>0.643</td>
<td>0.010</td>
</tr>
<tr>
<td>Anomalous evidence ratio (AER)</td>
<td>4</td>
<td>0.465</td>
<td>0.761</td>
<td>0.020</td>
</tr>
<tr>
<td>D $\times$ AER</td>
<td>8</td>
<td>0.511</td>
<td>0.846</td>
<td>0.043</td>
</tr>
</tbody>
</table>

*Note.* *Denotes $p$ value significant at $p > .05$. 

---

Fig. 1. Experiment 1: Interaction of domain by anomalous evidence ratio on confidence change.
showed decreased confidence in their theories. Providing anomalous evidence did not reliably lead to an explicit theory change, but its influence was detectable when measuring participants’ confidence in their theory before and after encountering anomalous evidence. While many participants did not decrease in their confidence, and thus might have dismissed or disregarded the anomalous evidence by some method, many more changed how they felt about their theory, even when they were not sufficiently compelled to change it. This finding is important because, while many studies and informal observations indicate that people are not influenced by anomalous evidence, this experiment has shown that an important change in thinking can occur without outward evidence of theory change.

Other important findings from Experiment 1 included the effect of anomalous evidence ratio on participants’ theory confidence, such that larger amounts of anomalous evidence caused a correspondingly larger decrease in theory confidence. Furthermore, participants’ decreases in theory confidence were themselves predictive of theory change. Larger drops in confidence predicted theory change. In combination, these findings suggest that over time people may be pushed closer to a theory change by experiencing greater doubt about their theory in the face of accumulating anomalous evidence. Furthermore, when participants found the anomalous evidence more convincing than the supporting evidence, as evidenced by higher scores on differential convincingness, they were more likely to decrease their confidence and make a theory change.

4. Experiment 2

The purpose of Experiment 2 was to test the Convergent Evidence Hypothesis by adding a factor to the same design utilized in Experiment 1 in the form of the randomly assigned dichotomous variable of evidence type which included the groups: (a) convergent evidence, consisting of qualitatively different types of evidence (e.g., survey, observational, case study), and (b) replicating evidence, which constitutes the same type of information being replicated by different sources.

The predicted effect was that the type of evidence would affect both theory change and confidence change such that participants should find it more difficult to ignore, stone-wall or dismiss a variety of different types (cases, surveys, controlled study) of anomalous evidence than replications of only one methodological kind. Convergent evidence should lead to greater decreases in theory confidence and more theory change than replicating evidence.

4.1. Method

The sample of 133 participants included 48 (23 replicating, 25 convergent) in the relationships domain, 44 (19 replicating, 25 convergent) in the athletics domain, and 41 (18 replicating, 23 convergent) in the cancer domain. As with Experiment 1, subjects participated in only one domain for which they qualified, were randomized to an anomalous
evidence ratio condition, and were randomized to the evidence type condition (replicating vs. convergent).

4.2. Results

Just as in Experiment 1, a manipulation check on the five-level factor of anomalous evidence ratio showed that the ratio affected the frequency of participants who reported that they perceived nonsupport for their accepted theory, $\chi^2(8) = 60.18, p < .001$. Across the five ascending levels of anomalous evidence ratio, the percentage of participants indicating the evidence set overall did not support their theory was 13% (3), 11% (3), 76% (19), 75% (21), and 88% (21). Participants who got larger amounts of anomalous evidence tended to rate the complete body of evidence as more anomalous to their theories.

4.2.1. Theory change

Just as in Experiment 1, only a relatively small percent (19% [14]) of participants changed their theory, while many more (58% [44]) exhibited decreases in theory confidence. A McNemar test showed these frequencies to be significantly different, $p < .001$.

The Convergent Evidence Hypothesis was supported as the evidence type manipulation did significantly affect the amount of theory change, $\chi^2(1) = 5.18, p = .02$, with 21% (15) of the convergent evidence group and only 7% (4) of the replicating evidence group explicitly changing their theory. Higher anomalous evidence ratio did not significantly predict the number of theory changes, 12% (3), 18% (5), 15% (4), 14% (4), and 12% (3), $\chi^2(4) = 0.58, p = .97$, and neither did domain, relationship 13% (6), athlete 11% (5), and cancer 20% (8), $\chi^2(2) = 1.35, p = .51$.

4.2.2. Confidence change

As shown in Table 3, and providing further support for the Convergent Evidence Hypothesis, both evidence type and anomalous evidence ratio had statistically significant effects on confidence change in a three-way (3 [domain] x 5 [anomalous evidence ratio] x 2 [evidence type]) ANOVA. As shown in Fig. 2, convergent evidence led to greater decreases in confidence than replicating evidence, $F(1, 103) = 5.14, p = .03$. Higher levels of anomalous evidence ratio led to larger decreases in confidence, $F(4, 103) = 5.09, p < .001$. Again, the linear trend was significant $t(52.5) = 3.85, p < .001$. There were no significant differences due to domain, nor did any interactions with domain approach significance, $F_s < 1.26$. However, the anomalous evidence ratio by evidence type interaction was marginal, $F(4, 103) = 2.36, p = .06$. This trend appears to be due to the replicating evidence group experiencing a drop in confidence for all ratios above 5-5, whereas the convergent evidence group experienced the largest decrease for the 9-1 ratio.

A significant point-biserial correlation between confidence change and theory change was found, $r_{pb} = -.24, p = .006$. This finding replicates the one found in Experiment 1, indicating that smaller decreases in theory confidence are often insufficient to cause theory change, but larger drops are predictive of theory change.
4.2.3. Differential convincingness

As was found in Experiment 1, supporting evidence \((M = 6.80, SD = 1.92)\) was rated as more convincing than anomalous evidence \((M = 5.54, SD = 1.90)\), \(t(132) = 4.75, p < .001\). As shown in Table 4, no significant differences were found in the differential convincingness of evidence across the three domains, \(F < 1\), across levels of anomalous evidence ratio, \(F(4, 132) = 1.75, p = .14\), or for evidence type, \(F < 1\), nor were there any significant interactions, \(Fs < 1.81\). The null result for evidence type is important because it suggests that convergent evidence need not be perceived as more convincing than replicating evidence to result in the greater decreases in theory confidence and more theory change. This lack of an effect for evidence type was consistent across all three domains, and not due to low power. Differential convincingness predicted theory change, \(r_{pb} = .17, p = .05\), and there was a significant correlation between differential convincingness on confidence change \(r = -.34, p < .001\). This correlation remained significant when tested in a partial correlation that controlled for anomalous evidence ratio, \(r_{ab,c} = -.30, p = .001\), thus replicating the direct effect for convincingness found in Experiment 1.

**Table 3**

Analysis of variance results for Experiment 2 (theory confidence change)

<table>
<thead>
<tr>
<th></th>
<th>(df)</th>
<th>(F)</th>
<th>(p)</th>
<th>(\eta^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain (D)</td>
<td>2</td>
<td>0.220</td>
<td>0.803</td>
<td>0.004</td>
</tr>
<tr>
<td>Anomalous evidence ratio (AER)</td>
<td>4</td>
<td>5.089</td>
<td>0.001*</td>
<td>0.165</td>
</tr>
<tr>
<td>Evidence type (ET)</td>
<td>1</td>
<td>5.142</td>
<td>0.025*</td>
<td>0.048</td>
</tr>
<tr>
<td>D × AER</td>
<td>8</td>
<td>1.066</td>
<td>0.393</td>
<td>0.076</td>
</tr>
<tr>
<td>D × ET</td>
<td>2</td>
<td>1.261</td>
<td>0.228</td>
<td>0.024</td>
</tr>
<tr>
<td>AER × ET</td>
<td>4</td>
<td>2.359</td>
<td>0.058</td>
<td>0.084</td>
</tr>
<tr>
<td>D × AER × ET</td>
<td>8</td>
<td>0.781</td>
<td>0.620</td>
<td>0.057</td>
</tr>
</tbody>
</table>

*Note. *Denotes \(p\) value significant at \(p > .05\).*

![Fig. 2. Experiment 2: Main effect for evidence type on confidence change.](image-url)
4.3. Discussion

The results of Experiment 2 replicated the findings from Experiment 1 and provided additional support for the Incremental Change Hypothesis whereby encountering a higher ratio of anomalous evidence was seen to have significant effects on theory confidence. More anomalous evidence led to larger decreases in theory confidence, but that was not always enough to cause theory change.

Results of Experiment 2 also supported the Convergent Evidence Hypothesis in that convergent evidence had a significantly greater impact than replicating anomalous evidence on eliciting theory change and decreasing theory confidence. The results of Experiment 2 further indicate that anomalous evidence often causes incremental change in theory confidence, even when theory change does not happen. When anomalous evidence comes in multiple types (i.e., convergent evidence), it requires more work to dismiss through discounting, stonewalling, and abeyances, as it can call for multiple strategic efforts to address each one of the different challenges to one’s theory. Replicating evidence, on the other hand, can be sufficiently addressed with one argument. This would explain why the convergent evidence, though equal to replicating evidence in perceived support when considered as a whole, and equal to replicating evidence in differential convincingness ratings, had a greater influence on confidence and theory change in combination.

Experiment 2 also replicated the finding that differential convincingness is significantly correlated with negative confidence change independent of the effects of anomalous evidence ratio. Because this result is a correlational finding between two variables that are not separated in chronology, it makes it impossible to ascertain directional causality. Yet, it is plausible that participants’ view of the convincingness of the evidence had an impact on their confidence in their theory.

5. Experiment 3

The purpose of Experiment 3 was to test the Alternative Theory Hypothesis by presenting an alternative theory that provides a coherent explanation for the anomalous evidence. The hypothesis was that having a theory explicitly available that both contradicts the
accepted theory and provides an explanatory framework for anomalous evidence would lead to larger decreases in accepted theory confidence and push one toward theory change.

5.1. Method

The sample of 120 participants included 39 in the relationships domain (18 alternative theory, 21 control), 49 in the athletics domain (21 alternative theory, 28 control), and 32 in the cancer domain (15 alternative theory, 17 control).

Subjects in Experiment 3 were randomized to either the alternative theory condition which received both their own original theory and the one that they did not believe, or to the control condition that received their own theory and read a second text related to the domain. In the control condition, the second text was tangential to either of the theories (technology’s influence on relationships, historical changes in athletic equipment, and advancements in cancer detection). Another procedural difference was that participants in the alternative theory group were asked to rate their confidence in both their own theory and the alternative theory that was presented to them, both before and after they read the evidence. Otherwise, all procedures for Experiment 3 were the same as the first two experiments.

5.2. Results

The manipulation check indicated that participants perceived the differences in levels of anomalous evidence. The effect for anomalous evidence ratio on perceived support of evidence replicated the pattern and that was found in Experiments 1 and 2, \( \chi^2(8) = 62.68, p < .001 \), such that higher ratios of anomalous evidence were viewed as less supportive of the theory: 4%, 24%, 70%, 87%, and 95%.

5.2.1. Theory change

Just as in Experiments 1 and 2, there was a statistically significant difference between the rate of theory change (22%) and confidence change (75%). A McNemar test indicated a significant difference in these proportions, \( p < .001 \).

The primary purpose of Experiment 3 was to test the Available Alternative Theory Hypothesis, which predicted that participants receiving the alternative theory that provided an explanation for the anomalous evidence would experience more theory change and greater decreases in theory confidence.

Participants in the alternative theory condition did not make significantly more theory changes, 22%, from the control group, 21%, \( \chi^2(1) = 0.18, p = .89 \), nor did theory change differ by anomalous evidence ratio, 19% (5), 19% (5), 19% (5), 23% (6), and 19% (5), \( \chi^2(4) = 0.64, p = .959 \). Unexpectedly, in this study, theory change rate did differ by domain, \( \chi^2(2) = 6.46, p = .04 \), as a higher percentage (12% [38]) of participants in the cancer domain changed their theory than did those in the relationship (6% [15]) or athletics (8% [16]) domains.
5.2.2. Confidence change

A three-way (3 [domain] × 5 [anomalous evidence ratio] × 2 [alternative theory presentation]) ANOVA found a significant effect for the alternative theory presentation, $F(1, 120) = 4.41, p = .04$, on confidence change (Table 5). Participants presented with the alternative theory decreased in confidence ($M = -10.67, SE = 2.16$) more than the participants who were not presented the alternative theory ($M = -4.79, SE = 1.78$). The effect of the anomalous evidence ratio, $F(1, 120) = 1.52, p = .21$, was not significant, but as shown in Fig. 3 there was a statistically significant linear trend, $t(53.4) = 2.90, p < .01$. The effect of subject domain, $F < 1$, was not significant. None of the interaction terms approached significance, $Fs < 1.23$.

As in the previous experiments, larger decreases in theory confidence were associated with more theory change, $r_{pb} = -.36, p < .001$. The correlation between participants’ confidence in their accepted theory and confidence in the alternative theory was also significant, $r = - .43, p = .001$, demonstrating that as participants lost confidence in their theory, they increased in confidence in the alternative. (This last result could only be analyzed for participants in the alternative theory condition because the control group was not presented with the alternative theory.)

5.2.3. Differential convincingness

The supportive evidence ($M = 5.42, SD = 0.95$) was rated as being more convincing than the anomalous evidence, which was supportive of the alternative theory ($M = 4.02, SD = 1.09$) and, like in Experiments 1 and 2, the difference was significant, $t(120) = 11.26, p < .001$. As shown in Table 6, a three-way (3 [domain] × 5 [anomalous evidence ratio] × 2 [alternative theory presentation]) ANOVA found that anomalous evidence ratio significantly impacted the differential convincingness of the anomalous evidence, $F(4, 110) = 3.49, p = .011$. However, the presentation of the alternative theory did not have a significant effect on these ratings, $F(1, 110) = 1.91, p = .171$, nor did domain, $F < 1$. None of the interaction effects approached significance, $Fs < 1.21$.

Differential convincingness was a statistically significant predictor of theory change, $r_{pb} = .32, p < .001$, and there was an effect of differential convincingness on confidence change, $r = -.43, p < .001$. This correlation held up significantly in a partial correlation analysis.

Table 5
Analysis of variance results for Experiment 3 (theory confidence change)

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>F</th>
<th>p</th>
<th>$\eta^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domain (D)</td>
<td>2</td>
<td>0.021</td>
<td>0.980</td>
<td>0.000</td>
</tr>
<tr>
<td>Anomalous evidence ratio (AER)</td>
<td>4</td>
<td>1.515</td>
<td>0.205</td>
<td>0.063</td>
</tr>
<tr>
<td>Available alternative theory (AAT)</td>
<td>1</td>
<td>4.410</td>
<td>0.039*</td>
<td>0.047</td>
</tr>
<tr>
<td>D × AER</td>
<td>8</td>
<td>0.863</td>
<td>0.550</td>
<td>0.071</td>
</tr>
<tr>
<td>D × AAT</td>
<td>2</td>
<td>1.228</td>
<td>0.298</td>
<td>0.027</td>
</tr>
<tr>
<td>AER × AAT</td>
<td>4</td>
<td>0.285</td>
<td>0.887</td>
<td>0.012</td>
</tr>
<tr>
<td>D × AER × AAT</td>
<td>8</td>
<td>0.711</td>
<td>0.681</td>
<td>0.059</td>
</tr>
</tbody>
</table>

Note. *Denotes $p$ value significant at $p > .05$. 

---

that controlled for anomalous evidence ratio $r_{ab,c} = -.42$, $p < .001$, thus replicating the direct relationship between differential convincingness and confidence change found in Experiments 1 and 2.

5.3. Discussion

The results of Experiment 3 provided partial support for the Alternative Theory Hypothesis by showing that those who had the alternative theory made readily available had significantly larger decreases in theory confidence than those to whom it was not presented. However, providing the alternative theory failed to lead to significantly more theory change. Experiment 3 also provided additional support for the Incremental Change Hypothesis. Even though theory change remained less frequent in this study, effects of increasing amounts of anomalous evidence and the presence of an alternative theory could be seen in participants’ changes in confidence in their theories. For participants in the alternative theory condition, increases in confidence in the alternative theory were

![Fig. 3. Experiment 3: Main effect for presentation of alternative theory on confidence change.](image)

<table>
<thead>
<tr>
<th>Table 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analysis of variance results for Experiment 3 (differential convincingness)</td>
</tr>
<tr>
<td>df</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Domain (D)</td>
</tr>
<tr>
<td>Anomalous evidence ratio (AER)</td>
</tr>
<tr>
<td>Available alternative theory (AAT)</td>
</tr>
<tr>
<td>$D \times$ AER</td>
</tr>
<tr>
<td>$D \times$ AAT</td>
</tr>
<tr>
<td>AER $\times$ AAT</td>
</tr>
<tr>
<td>$D \times$ AER $\times$ AAT</td>
</tr>
</tbody>
</table>

*Note. *Denotes $p$ value significant at $p > .05$. 
significantly correlated with decreases in confidence in the accepted theory, thus showing that the theories competed directly.

The group presented with the alternative theory did not differ from the control group in their interpretation of the evidence’s convincingness. One might expect that having the alternative theory available would provide participants with a better explanatory framework for the anomalous evidence that they encountered, and that they would consequently find the evidence more convincing, but this was not apparent in the findings.

6. General discussion

The over-arching contribution of these three experiments is the demonstration of the gradual nature of the conceptual change process, and that certain elements of evidence and theories can influence people’s confidence in their theories, even when they do not waiver from explicit acceptance of the theories. The first experiment set out to test the *Incremental Theory Hypothesis*, or the idea that theory change happens gradually instead of all at once. Support for this hypothesis was found in that a greater number of participants experienced confidence change, with a smaller number explicitly changing in their acceptance of theories. There was a consistent finding that increases in anomalous evidence predictably impacted people’s confidence in their theory, even when they did not make an explicit theory change. The second experiment tested the *Convergent Evidence Hypothesis*, or the idea that qualitatively different forms of evidence would be more difficult to dismiss than replications of the same finding. Support for this hypothesis was found in that participants presented with convergent evidence experienced greater confidence change and theory change than those presented with replicating evidence. Support for the *Incremental Theory Hypothesis* and the effects of increased anomalous evidence on confidence were also found in Experiment 2. Experiment 3 explored the *Available Alternative Theory Hypothesis*, or that having an available alternative theory that explains the anomalous evidence would elicit more confidence and theory change. Although those presented with an alternative theory did not express more theory change, it was found that when confidence in the accepted theory decreased, confidence in the alternative theory increased. And, yet again, support for the *Incremental Theory Hypothesis* was found in Experiment 3.

6.1. Theory confidence and theory change

The consistent patterns across studies showing decreases in confidence in response to anomalous evidence make it clear that many participants did not simply ignore or summarily dismiss this evidence, but instead experienced a measurable change in their thinking (Chinn & Samarapungavan, 2009). However, it was only by using a measure of confidence in a theory (as opposed to belief revision or theory acceptance), and computing changes in confidence from before to after anomalous evidence was presented, that the influence of anomalous evidence could be evidenced. It is a clear contribution to the literature that the results from all three experiments provide support for the new
Incremental Change Hypothesis, in that larger amounts of anomalous evidence presented to participants predicted greater decreases in theory confidence. And as predicted, the amount of anomalous evidence and theory confidence change showed a linear trend. These results show that people are predictably and measurably influenced by anomalous evidence in a gradual fashion. These findings suggest that assessments of the accuracy of accepted theories are often updated after new anomalous data are encountered. This phenomenon would not have been detectable by only examining theory adherence and suggests that this new experimental framework by which to quantify the amount of change is a potentially useful one.

6.2. Replicating versus converging anomalous evidence

Support for the Convergent Evidence Hypothesis was obtained as convergent evidence elicited larger decreases in theory confidence and more frequently led to theory change than did replicating evidence. This novel finding suggests that more cognitive effort might be needed to resolve the challenges to the accepted theory presented by multiple kinds of anomalous evidence. But, without qualitative data, it is not clear if the participants in Experiment 2 used strategies that enabled them to ignore, reject, hold in abeyance, or reinterpret the anomalous evidence based on the methodologies used to collect it, as scientists might do during anomaly resolution (Chinn & Brewer, 1993a). Viewing these results through the scope of some previous work dedicated to how people create models from multiple sources would suggest that a stronger or more robust model can be constructed from evidence obtained through various different methodologies (Chinn & Brewer, 2001; Hemmerich & Wiley, 2002; Kuhn et al., 1988; Wiley et al., 2009). With a relatively small amount of research historically directed at evaluation of data, as compared to theory, this is an important finding that needs intensive follow-up investigation (Amsel & Brock, 1996; Chinn & Brewer, 1993a; Gorman, 1989; Kuhn et al., 1988; Lord et al., 1979).

What is clear is that participants in the replicating evidence group more commonly preserved their accepted theory and their confidence in it. This result makes a significant contribution to the literature as it demonstrates an observable effect for a diversity of anomalous evidence in being more persuasive than replication of any one finding. We can see this effect in the scientific world in that scientists do not change theories or lose too much confidence because of a single anomalous finding. Instead, as Lakatos describes, they red flag results and defer changes in theory until a diversity of data suggest that the current theory is wrong ((Lakatos, 1970). It is possible that use of this strategy is the reason why support for the Convergent Evidence Hypothesis was found.

6.3. Availability of an alternative theory

The third major hypothesis, the Alternative Theory Hypothesis, received partial support as presenting participants with the alternative theory did not uniformly lead to significantly more theory changes as predicted, but did cause larger decreases in theory confidence. In Experiment 3, decreases in theory confidence were negatively and significantly
correlated with increases in alternative theory confidence. When confidence in one’s accepted theory decreased, the confidence in the alternative theory, which provided an explanation for the anomalous evidence, increased. This suggests a more fluid, dynamic, and gradual system underlying conceptual change where different competing coherent mental models may appear more or less plausible based on new data. However, to better understand the time course and duration of this process that has only begun to be charted in this study, more extensive repeated measures and longitudinal designs are needed.

6.4. Implications for models of theory change and limitations

It has been argued that conceptual change approaches have generally put too much emphasis on belief revision as a function of sudden, radical change (Limón, 2001; Vosniadou, 2013; Vosniadou & Ioannides, 1998). As an alternative, one can conceptualize discovery as a gradual process (Chinn & Brewer, 2001; diSessa, 2006; Posner et al., 1982; Vosniadou & Brewer, 1994) during which conceptual structures are continuously enriched based on everyday experiences. The present approach argues that in addition to the need to accumulate evidence, an important undercurrent representing changes in confidence in reaction to that accumulating evidence may be a necessary precursor for conceptual shifts, radical restructurings, and explicit changes in beliefs or theories.

When the results of these experiments are viewed in the context of Chinn and Brewer’s taxonomy (1993a,b, 1998), they suggest that a minority of participants engaged in radical restructuring, acknowledging and accepting the anomalous evidence and explicitly changing their theories. This response was most common when participants were provided with anomalous evidence of varying methodological kinds. However, a larger proportion of the participants did not make a theory change, but did exhibit theory confidence changes. As noted above, more qualitative work is needed to better understand how participants regarded the anomalous evidence in these studies. Given the current methodology, it is not clear whether participants felt uncertainty about the evidence’s validity, excluded it as irrelevant to the domain, held it in abeyance, or somehow reinterpreted its ramifications. Some participants might have accepted the anomalous evidence and made peripheral theory changes, but retained their original theory, as the experiments were not sensitive to detect peripheral tweaks.

Similarly, these experiments are limited in that they presented only a cross-sectional look at the effects of different amounts of evidence on theory change. In order to better understand the nature of incremental change and the longer term effects of evidence type and alternative theories, it will be necessary to implement longitudinal designs in which change is measured over time and temporally diffuse encounters. These data suggest that changes in confidence can move one closer to making a theory change. In order to more thoroughly test this theory, it would be important to provide research participants with anomalous evidence, of different kinds, over many repeated experimental sessions. An extended longitudinal design is absolutely necessary to determine how confidence changes might affect theory change over time, whether multiple exposures to anomalous evidence work together to diminish one’s confidence in their theory and enhance the odds of a
theory change as one accumulates, or whether anomalous evidence may be forgotten from one encounter to the next. Engaging in such data collection efforts would allow one to create a more complete understanding of the relation between confidence and theory change, and it could also allow one to test for important mediators, like how participants perceive the evidence they are presented. It will also be important to explore these same questions in a classroom setting rather than the laboratory context used here.

One strength of this set of experiments is that they attempted to investigate and generalize conceptual change processes across three different domains, as opposed to more prototypical studies that examine change in only a single domain at a time. That said, the domains that were chosen did not represent topics for which students might have a “misconception” in any traditional sense. When alternative theories were presented, neither alternative was explicitly formulated as a misconception or naïve theory. These experiments included college-aged adults who had consistently expressed belief about the topic on more than one occasion prior to the experimental manipulation, suggesting their theory belief had some stability (Carey, 1999; Chi, 2008; Gopnik & Wellman, 1994). However, stable beliefs may not be entrenched to the same degree as naïve conceptions or misconceptions that are encountered in relation to understanding some scientific phenomena. More work is needed to see if confidence change can also be seen in domains where prior beliefs might be more entrenched than for the topics that were used in this study. Indeed, although the use of three different topic domains was intended to help demonstrate how the effects of the manipulations might generalize, there were some small differences in the pattern of results for the different domains in Experiment 1. These effects could have been due to differences in how people think about these three domains, and perhaps related to how strongly they believed their theories for each of them before entering the study. Future research is needed to comment on how different levels of entrenchment might alter the response to anomalous evidence for other topics.

A final observation is that across studies, there was no explicit instruction to engage in theory comparison, nor was refutation of the participants’ original theory by anomalous evidence made explicit. Theories were determined to be flawed only if participants viewed them as such in light of their failure to account for the anomalous evidence. Participants were able to abandon their theory and express belief in another theory if they felt that it better explained the topic domain after having encountered the anomalous evidence. Such a change can be described within the framework defined by Chi (2008) as an abandonment of one mental model for another because the anomalous evidence is in conflict with the original theory, but not with the alternative theory. The experimental paradigm provided some control as to pitting two theories against each other by presenting anomalous evidence that could not be explained by one, but not by the other theory. However, as opposed to many studies that have done work on naïve conceptions (or misconceptions) and more advanced conceptions that are mutually exclusive, the alternative theories that were considered in this study were not necessarily diametrically opposed. Future work exploring different sets of theories that are in greater opposition, and instructional conditions that highlight the mutual exclusivity of the alternative theories, will help
to further explicate whether confidence change may also serve as a precursor to theory change in those contexts.

7. Conclusion

Overall, the results from these three experiments demonstrate that theory change and confidence change are distinct but correlated constructs, and that anomalous evidence often causes alterations in people’s thinking about their theories that are only sometimes sufficient to elicit a theory change. The results also demonstrate that the diversity of methodological types of anomalous evidence that are encountered can increase the prevalence of theory change and the magnitude of decreases in theory confidence. These findings offer evidence for theory change as the result of a gradual and incremental process of confidence change. Future experiments using longitudinal or qualitative methods will help to more fully demonstrate the mechanisms by which exposure to anomalous evidence results in theory change and belief revision.

Acknowledgments

The authors thank Thomas D. Griffin and Kenneth Zoll for argument, laughter, and discussion. Joshua A. Hemmerich passed away on March 25, 2015.

References


**Supporting Information**

Additional Supporting Information may be found in the online version of this article:
- **Appendix A** Theory passages.
- **Appendix B**. Evidence piece passages.

**Appendix A: Example theory passage**

**A.1. Natural born athlete***

The role that genetics play in athletic aptitudes and talents is indisputable. Genetic endowment is the most determinant factor of an individual’s level of achievement in athletics. Those who are born to be strong, fast, coordinated, and have high endurance, and recover quickly from injuries and fatigue will perform the best in their chosen sport.

One of the primary determinants of athletic performance is the composition of the person’s muscle tissue. There are two different primary types of muscle fibers, fast twitch and slow twitch. Simply put, fast-twitch fibers can contract approximately 10 times as fast as slow-twitch fibers, but fast-twitch fibers also fatigue, or “burn out” much more quickly. The average proportion of these fibers in a healthy human adult is roughly half and half, but it varies greatly by individual person.

Much work in human physiology indicates that people who possess a higher proportion of fast-twitch muscle fibers will perform better at sports that require short bursts of explosive speed and power such as sprinting, power lifting, and long jumping. Conversely, those who possess a higher portion of slow-twitch muscle fibers will perform better at sports requiring endurance and stamina such as long-distance running and swimming. The proportion of the types of muscle fibers that an individual possess is largely determined at birth.

The difference between these two populations of athletes, particularly professional and most specifically championship caliber, is obvious when individuals are examined from sports that require one type of muscular performance versus the other. Virtually all Olympic power lifters and sprinters have a very large number of fast-twitch muscle fibers at their disposal that are capable of firing powerfully but tire out relatively quickly. This kind of tissue composition allows for these competitors to explosively lift heavy weights above their heads and to sprint across a long stretch of track in only a few seconds. Endurance athletes, like the ones that perform well in the Boston Marathon, have been shown to possess a larger percentage of slow-twitch muscle fibers that do not contract as forcefully but also take a long time to fatigue.
This difference in biologically determined cellular composition is not the only genetically determined characteristic that determines an individual’s success in a particular sport. Characteristics such as height, body fat composition, bone structure, and lung capacity all help to determine the ability of an athlete in a particular sport. All of these characteristics are largely genetic. Tall parents have tall children who often do well at basketball. Parents with very light bone structures are more likely to have light children who make good mountain climbers. People who inherit good lung capacities and hydrodynamic bodies are the ones who becoming champion swimmers. Success in sports is clearly a matter of the quality and fortuitous combination of genes that an individual inherits from his or her parents.

Furthermore, as a child discovers that he is good at a particular sport, his interest and motivation to play that sport increases. Kids who have a natural ability to run fast will pick sports that require fast running to be successful. Youngsters who have good upper body strength usually choose to play sports that allow them to take advantage of such a gift. Likewise, coaches and recruiters look for the genetically gifted athletes that will help their team win. Children with natural physical attributes are often encouraged to pursue the sport that they are gifted for: “Junior’s getting rather tall, is he trying out for basketball this year?”

Scientists have evidence now for what people have intuitively known for a long time: that athletes are born, not made. Genetic inheritance is the mechanism by which one obtains his or her level of ability in sports. It is clear that natural endowment is the gift that makes champions.

Appendix B: Example evidence passages

B.1. Replicating

C1. Surveys conducted by the National Athletics Board indicated that people who had attained athletic achievements also had parents who had extensive athletic achievements. Furthermore, in most cases the people reported that they were naturally talented at their chosen activity, which was usually the same activity that their parents had excelled in. The vast majority of respondents claimed that their good performance in their chosen activity just “came naturally.”

C2. Surveys of college athletic coaches from several different sports indicated that talent and natural athleticism were the most important factors in making a player successful in their chosen sport. These inborn characteristics were the first things that the coaches considered when trying to recruit a player. Many coaches explained “You just can’t teach athletic ability; people are either born with it or not.” Many coaches agreed with the assertion that the most naturally talented players were the ones who achieved the most in competition.
B.2. Converging

A1. Studies of monozygotic (identical twins) separated at birth have shown that they both tend to have the same level of achievement in athletics, typically in the same sports. For example, a 17-year-old high school student who was an accomplished basketball star and was being recruited by several professional teams unknowingly has a brother in another state who is receiving the same acclaim and interest. This is the pattern found with most identical twins that have been separated.

A2. Many physiological researchers have observed that the greatest predictor of an individual’s athletic ability and aptitude for athletic performance is the ability of their parents. In a multigenerational study, parents were tested on several dimensions related to athletic ability, including running speed, endurance, upper body strength, and vertical jump. When their children were tested on the same dimensions many years later (when they reached the same age their parents were when tested), it was found that their performance was accurately predicted by the performance of their parents on the same physical tests years earlier.

B.3. Neutral

NA1. Sports physiologists have studied the role that nutrition plays in athletic performance and found that the food that one takes in has a significant effect on one’s performance. When runners were deprived of adequate carbohydrates for a 4-day period, they fatigued much sooner than they did under normal nutritional intake.

*All examples taken from Athlete Evidence, Natural Athlete.