ARTICLE IN PRESS

Global Environmental Change xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Global Environmental Change

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

What makes for compelling science? Evidential diversity in the evaluation of scientific arguments☆

Arthur Kary*, Ben R. Newell, Brett K. Hayes

University of New South Wales, Sydney, Australia

ARTICLE INFO

Keywords:
Reasoning
Science communication
Diversity effect
Best-worst scores
Discrete choice experiment
Climate change
Public health

ABSTRACT

Despite overwhelming scientific evidence, many members of the public remain skeptical about anthropogenic global warming. Hence, we examined how the presentation of factual scientific evidence affects lay evaluations of scientific claims. Taking inspiration from cognitive research on inductive reasoning, two studies examined the impact of evidential diversity on acceptance of claims in the domains of climate change and public health. Participants were presented with scientific claims based on competing evidence options and were asked to choose the best and worst form of evidence for each claim. The diversity of the available evidence was manipulated across three dimensions; geographical (evidence from two geographically near or far nations), socio-cultural (evidence from two culturally similar or dissimilar nations), and temporal (evidence drawn from two different periods or the same period). In both studies, diverse evidence on the geographical and socio-cultural dimension increased perceived support for scientific claims, but the relative impact of these dimensions differed between domains; geographical diversity had a larger effect on claims about climate change; socio-cultural diversity had a larger effect on claims about health. On the temporal dimension, recent non-diverse evidence (i.e. from the same recent period) increased perceived support for scientific claims more than diverse evidence. These results may have important implications for the communication of complex scientific evidence to a lay audience.

1. Introduction

The relevance of science to daily life and everyday decisions is growing. With increasing frequency, lay people are confronted with making decisions about complex socio-scientific issues such as whether they will support government regulation of carbon emissions, whether they should install solar paneling in their home, or whether they should reduce the amount of salt in their diet (Sinatra and Hofer, 2016).

Despite the increasing prevalence of socio-scientific issues, in many areas there remains a significant gap between the scientific and public understanding of key problems. A signature case is belief in anthropogenic global warming (AGW). Over the past decade there has been increasing agreement among scientists that carbon emissions from human activity contribute to global warming (Freudenburg and Muselli, 2010; IPCC, 2014). Over the same period however, public belief in AGW in countries such as the USA and Australia has remained static or even decreased (Brulle et al., 2012; Leviston et al., 2015; Smith and Leiserowitz, 2012).

Such cases highlight the importance of understanding how non-scientists interpret and evaluate scientific evidence. Despite the crucial importance of this issue, surprisingly little work has investigated the cognitive factors that affect how non-scientists evaluate the strength of scientific arguments (Corner and Hahn, 2009; Hahn et al., 2016). For example, although provision of scientifically accurate information about AGW can increase acceptance of the phenomenon and promote pro-environmental attitudes (Ranney and Clark, 2016), there have been few attempts to examine the cognitive mechanisms involved in the lay interpretation of such information. An understanding of these mechanisms is important if we wish to maximize the acceptance and impact of science in public policy debates.

A potentially useful approach to advancing our understanding of how lay people evaluate scientific data involves considering the implications of research on inductive reasoning in non-scientific domains. How people use evidence to draw more general conclusions is the central focus of research on inductive reasoning (Hayes et al., 2010). In studies of inductive reasoning, people are presented with new evidence that is assumed to be true (e.g., that lions and dolphins have "beta cells inside") and are asked to evaluate the strength or plausibility of some conclusion based on this evidence (e.g., that all mammals have beta cells). Such inferences are probabilistic, with belief in the conclusions

☆ This work was supported by Australian Research Council Linkage Grant LP120100224 to the second and third authors. We thank Jeremy Ngo and Manfred Ng for their help in manuscript preparation and Marilynn Brewer and Rachel McDonald for useful discussions during the course of this project.

* Corresponding author at: School of Psychology, University of New South Wales, Sydney 2052, Australia.
E-mail address: a.kary@unsw.edu.au (A. Kary).

https://doi.org/10.1016/j.gloenvcha.2018.01.004
Received 25 January 2017; Received in revised form 3 January 2018; Accepted 28 January 2018
0959-3780/

increasing or decreasing depending on the quality and quantity of evidence.

Laboratory studies of inductive reasoning have revealed a range of factors that affect perceptions of argument strength (see Hayes and Heit, 2013 for a review). In this respect, the literature on inductive reasoning is a potentially rich source of ideas about what types of evidence people find more or less convincing. However, in order to minimize the impact of prior knowledge on argument evaluations, the "evidence" and conclusions presented in these laboratory studies are usually not factual, often relying on invented or abstract properties (e.g., "has beta cells" or "has property P"). A major aim of the current studies was to examine whether a key factor identified in laboratory work as influencing inductive inferences, *evidence diversity*, also affects the way that people evaluate factual scientific evidence.

1.1. Evidence diversity in inductive reasoning

Philosophers of science have argued that other things being equal, a scientific theory is more strongly supported by diverse rather than non-diverse evidence (Bacon, 1620/1898; Hempel, 1966; Salmon, 1984; see Heit et al., 2004 for a review). Salmon (1984) for example, reviews the way that early 20th century scientists developed a range of methods for deriving Avogadro's number (6.02×10^{23}), the number of particles in a mole of any substance. These included Brownian movement, alpha particle decay, X-ray diffraction, black body radiation, and electrochemistry. The derivation of Avogadro's number was a significant scientific discovery as it provided support for what was, at the time, a controversial hypothesis: the existence of atoms and molecules. Salmon argued that the evidence taken from any one experimental technique alone would be unlikely to be viewed as convincing evidence in favor of atomic theory. The convergence of results based on diverse methods greatly strengthened the theory.

Research on inductive reasoning shows that non-scientists view evidential diversity as important in evaluating the strength or plausibility of an argument. Although there are some interesting exceptions (e.g., Medin et al., 2003), the general finding is that conclusions based on diverse forms of evidence are viewed as more plausible or convincing than those based on more similar forms of evidence (e.g., Feeney and Heit, 2011; Kim and Keil, 2003; Osherson et al., 1990; Shafto et al., 2007). For example, people are more likely to endorse a general conclusion (e.g., that mammals have some property P) given evidence about diverse or dissimilar instances (e.g., lions and dolphins have property P) than evidence about very similar instances (e.g., lions and tigers have property P). Such sensitivity to evidential diversity emerges during the elementary school years (Heit and Hahn, 2001; Rhodes et al., 2010) and has been shown to affect evidence selection in hypothesis testing (López, 1995), memory for learned material (Hahn et al., 2005), and conceptual change during development (Hayes et al., 2003). Although there is some debate about whether such effects are normative (cf. Heit and Hahn, 2001; Lo et al., 2002; Wayne, 1995), there seems little doubt that evidence diversity is an influential heuristic for assessing the strength of an argument.

These findings lead to the prediction that factual scientific arguments supported by more diverse forms of evidence will often be viewed as more plausible and persuasive than arguments based on less diverse forms of evidence. Here we test this prediction in two experiments using arguments from two scientific domains: climate change and public health.

1.2. Assessing the diversity of complex scientific evidence

The question of what constitutes "diverse evidence" is almost certainly more complex in real scientific domains than has been the case in previous laboratory studies of inductive reasoning. In previous work, the problem of measuring evidential diversity has generally been addressed by ensuring that all argument premises and conclusions are drawn from a single conceptual domain (e.g., animals). Hence, instances can be compared on a single dimension (e.g., taxonomic similarity) and it is quite straightforward to establish the relative similarity or diversity of the instances used in argument premises.

In contrast, the evidence involved in scientific domains such as climate change and public health is notoriously complex. Climatic events and population-wide changes in health are the end result of the interaction between multiple causal systems. The scientific evidence that bears witness to these phenomena comes in a wide variety of forms including archival records, observations from geographically and socially distinct testing sites, and experimental findings.

The comparison of two or more types of evidence relating to actual scientific phenomena therefore is likely to involve consideration of multiple dimensions of similarity. For example, suppose we learned the following: a) over the first half of the 20th century there was an increase in the mean sea level around the Australian coastline, and b) records taken over the past 20 years have shown that mean sea level has increased along the coastline of West Africa. Unlike the evidence used in conventional induction research, these statements differ on at least three dimensions; temporal (data collected at different times), geographical (data collected from different parts of the world), and social (data collected from regions with different cultural and economic profiles). Each of these dimensions needs to be considered in assessing the diversity of this evidence and how it affects belief in a conclusion like "global sea levels are rising".

This example raises the question of how people respond to diversity on multiple dimensions. Extrapolating from previous studies of the diversity heuristic in arguments with a single conceptual dimension, one might predict that diverse evidence across multiple dimensions will have a cumulative effect on confidence in a scientific claim. In other words, diverse evidence across multiple dimensions provides more compelling evidence for a claim than diverse evidence on fewer dimensions.

An alternative possibility is that in complex scientific domains people are selective in the way they employ the diversity heuristic. Diversity on some dimensions may have a greater impact on the evaluation of scientific claims than others. Nonscientists understand that different causal factors operate in different scientific domains (Sloman et al., 2007). For example, an understanding of the different causal principles that operate in the domains of everyday physics, biology and psychology emerges relatively early in development (Carey, 1995; Keil, 2003). Hence, the dimensions of diverse evidence that are perceived as most relevant for evaluating a scientific conclusion are likely to vary across domains. In the above example, it may be that people are more persuaded to believe that global sea levels are rising by the fact that increases have been detected in two very different locations and over different measurement periods. The social and cultural diversity of the testing sites may be seen as less crucial. In contrast, if the evidence and associated conclusions were concerned with a public health issue (e.g., disease prevalence, rates of immunization) then socio-cultural diversity may be a central consideration.

To take account of these complexities we developed a novel approach to studying evidential diversity in scientific domains. This involved systematically manipulating the diversity of scientific facts on three dimensions: temporal, geographical and social. Diversity on these dimensions was manipulated factorially so that people saw evidence items that were diverse on zero, one, two or all three dimensions. This allowed us to examine the separate impact of diversity on each dimension on belief in scientific conclusions in the domains of climate change and public health.

An additional justification for focusing on these three dimensions was that they are among the key factors that have been shown to influence subjective perceptions of "psychological distance" (see McDonald et al., 2015 for a review). Psychological distance refers to the extent to which something is perceived as distant from oneself. Previous work shows that the perceived "distance" of events (e.g., climate