



PEARL

Modeling causal conditional reasoning data using SDT: caveats and new insights.

Trippas, Dries; Verde, Michael F.; Handley, Simon J.; Roser, Matthew E.; McNair, Nicolas A.; Evans, Jonathan St B.T.

Published in:
Front Psychol

DOI:
[10.3389/fpsyg.2014.00217](https://doi.org/10.3389/fpsyg.2014.00217)

Publication date:
2014

Link:
[Link to publication in PEARL](#)

Citation for published version (APA):

Trippas, D., Verde, M. F., Handley, S. J., Roser, M. E., McNair, N. A., & Evans, J. S. B. T. (2014). Modeling causal conditional reasoning data using SDT: caveats and new insights. *Front Psychol*, 5(0). <https://doi.org/10.3389/fpsyg.2014.00217>

All content in PEARL is protected by copyright law. Author manuscripts are made available in accordance with publisher policies. Wherever possible please cite the published version using the details provided on the item record or document. In the absence of an open licence (e.g. Creative Commons), permissions for further reuse of content should be sought from the publisher or author.



Modeling causal conditional reasoning data using SDT: caveats and new insights

Dries Trippas*, Michael F. Verde, Simon J. Handley, Matthew E. Roser, Nicolas A. McNair and Jonathan St. B. T. Evans

Faculty of Health and Human Sciences, School of Psychology, Cognition Institute, Plymouth University, Plymouth, UK

*Correspondence: dries.trippas@gmail.com

Edited by:

David E. Over, Durham University, UK

Reviewed by:

Henrik Singmann, Albert-Ludwigs-Universität Freiburg, Germany

Evan Heit, University of California, Merced, USA

Keywords: causal conditionals, reasoning, signal detection theory, belief bias, normative models

In deductive reasoning, people are asked to infer the truth of an argument's conclusion given a set of premises. Research into the processes underlying deduction has focused on examining how well people discriminate between logically valid and invalid arguments, and how irrelevant factors such as one's prior beliefs interfere with the ability to reason logically (Evans et al., 1983). This normative approach to validity has traditionally informed both practice and theory in the literature. However, its critics argue that "normativism" often leads investigators to biased or misleading interpretations of phenomena (Elqayam and Evans, 2011).

Formal modeling of deductive reasoning has often been successful by taking the traditional, normative approach. A case in point is the application of signal detection theory (SDT; Macmillan and Creelman, 2005) to the investigation of belief bias in syllogistic reasoning (Dube et al., 2010). In the SDT model, deductive judgments are based on strength of evidence; an argument is judged to be valid if its strength exceeds a criterion value. Because the choice of criterion is independent of the ability to discriminate between classes of arguments, the SDT model makes it possible to isolate response bias from accuracy. Dube et al. examined these two factors using ROC curves, which plot hits against false alarms at several levels of confidence. Hits and false alarms were defined in normative fashion as responding "valid" to logically valid and logically invalid conclusions, respectively.

Their analysis of ROCs led them to argue two significant points. First,

contrary to prevailing theories of belief bias, conclusion believability can affect response bias without affecting the quality of reasoning. Second, the curvilinear shape of the ROCs is consistent with the distributional assumptions of SDT. The latter is a key test because finding linear rather than curvilinear ROCs would be problematic for the model. The curvilinear ROCs obtained in syllogistic (see also Dube et al., 2011; Trippas et al., 2013; but see Klauer and Kellen, 2011) and other forms of reasoning (Heit and Rotello, 2010, 2014) are similar to those widely observed in memory and perception (Pazzaglia et al., 2013). This consistency across domains strengthens the case for the usefulness of the SDT approach. It also leads to an expectation of similar findings in other areas of reasoning. Below, we describe findings from conditional reasoning that violate this expectation in a surprising yet enlightening way.

Causal conditionals are a form of deduction prevalent in everyday life. Consider the proposition: "If healthy foods are cheaper, then more people will eat healthy foods." Four types of conditional inferences are possible: modus ponens (MP; "Healthy foods are cheaper, therefore more people will eat healthy foods"), modus tollens (MT; "Fewer people eat healthy foods, therefore healthy foods are not cheaper"), affirmation of the consequent (AC; "More people eat healthy foods, therefore healthy foods are cheaper"), and denial of the antecedent (DA; "Healthy foods are not cheaper, therefore less people eat healthy foods").

From a normative point of view, MP and MT are valid and AC and DA are invalid inferences. Theories differ as to how people determine validity in these problems. According to mental model theory (Johnson-Laird and Byrne, 2002), people construct an initial mental model of the conditional (e.g., $p \rightarrow q$) which may then be fleshed out by considering additional models ($\neg p \rightarrow q$; $\neg p \rightarrow \neg q$). According to the suppositional account of the conditional (Evans et al., 2003, 2005; Evans and Over, 2004, 2012), people evaluate the subjective probability of a conditional by hypothetically supposing p and then assessing the conditional probability of q given p , $P(q|p)$. This relation between the natural language conditional and the conditional probability, $P(\text{if } p \text{ then } q) = P(q|p)$, can be used in a Bayesian/probabilistic model of conditional inference (Oaksford et al., 2000; Oaksford and Chater, 2009, 2013).

What these theories have in common is that there is no fundamental difference in how people process affirmation (MP + AC) and denial (MT + DA) inferences. This makes an SDT analysis straightforward and no different to that taken with the study of belief bias in syllogistic reasoning. For our case study, we analyzed aspects of a data set collected as part of a larger project under the direction of the fourth author of this paper¹. This study examined the influence of belief in causal conditional problems (e.g., believable: "If oil prices continue to rise, then UK

¹This research was supported by the award of an ESRC project grant RES-062-23-3285.

petrol prices will rise”; unbelievable: “If global temperatures rise, then less arctic ice will melt”). Hits were defined as “valid” responses to MP and MT and false alarms were defined as “valid” responses to AC and DA. This produced the ROCs seen in the top panel of **Figure 1**. The results are similar in some respects to those reported by Dube et al. (2010) for syllogisms: believability had no effect on accuracy (ROCs for believable and unbelievable items fall on the same curve) but seemed to affect response bias (confidence criteria for believable items are shifted to

the right)². However, there is a surprising difference: in contrast to the curvilinear ROCs observed with syllogisms, conditionals produced linear ROCs. A linear regression of the ROC (collapsing over believability) provided a good fit, $R^2 = 99.9\%$. Adding a quadratic component did not improve the fit, $p = 0.78$. Taken at face value, this result suggests that conditional reasoning requires a profoundly different

²Note that the current data pattern does not necessitate a criterion shift, but that it is also consistent with a symmetrical distribution shift. For more discussion on this issue, see, e.g., Verde et al. (2010).

model than the one that has seemed so successful when applied to other forms of reasoning, not to mention other cognitive tasks.

A different picture emerges when we depart from the strictly normative approach and consider separately how people respond to affirmation and denial conditionals. In the bottom left panel of **Figure 1**, plotting MP (hits) against AC (false alarms) yields typically curvilinear ROCs. Linear regression (collapsing over believability) provided a fit, $R^2 = 96\%$, that was significantly improved by the

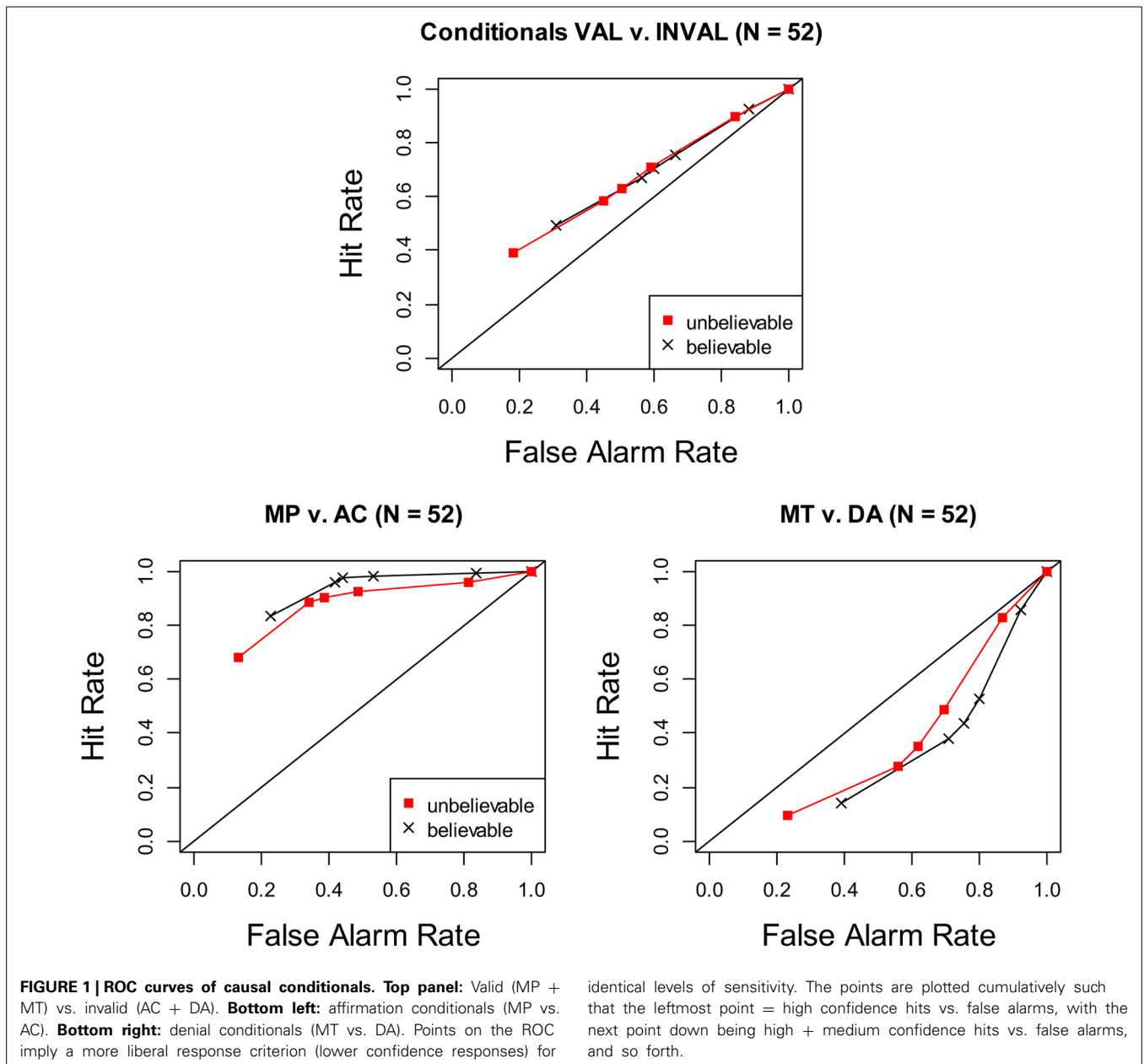


FIGURE 1 | ROC curves of causal conditionals. Top panel: Valid (MP + MT) vs. invalid (AC + DA). **Bottom left:** affirmation conditionals (MP vs. AC). **Bottom right:** denial conditionals (MT vs. DA). Points on the ROC imply a more liberal response criterion (lower confidence responses) for

identical levels of sensitivity. The points are plotted cumulatively such that the leftmost point = high confidence hits vs. false alarms, with the next point down being high + medium confidence hits vs. false alarms, and so forth.

addition of a quadratic component, $R^2 = 99.99\%$, $p < 0.004$. Accuracy is defined by the distance of the ROCs from the chance diagonal. Contrary to the poor accuracy on display in the aggregate results in the top panel, people are quite sensitive to argument structure when affirmation is involved. In the bottom right panel of **Figure 1**, plotting MT (hits) against DA (false alarms) again yields typically curvilinear ROCs. Linear regression (collapsing over believability) provided a fit, $R^2 = 98\%$, that was significantly improved by the additional of a quadratic component, $R^2 = 99.99\%$, $p < 0.002$. People were sensitive to argument structure, but the position of the ROCs below the diagonal indicates that their treatment of denial arguments departed from the normative; MT are treated as *less* valid than AC.

Applying the SDT model in a normative fashion, as would seem reasonable given extant theories of conditional reasoning, produced results that contrast sharply with previous findings. The clearly linear ROC in the top panel of **Figure 1** is not only unlike the curvilinear ROCs observed with syllogisms but if taken at face value is problematic for the SDT model. It could be that there is something fundamentally different in the way that people reason about causal conditionals as compared to other types of problems. It seems to us more likely that the difference lies with affirmation and denial inferences; the latter do not seem to be treated in the normatively prescribed fashion. Once this is assumed, the ROC results become more sensible and fall in line with previous results (in a reanalysis of published and unpublished data sets, Heit and Rotello, 2014, have also reported curvilinear ROCs from MP plotted in the manner of **Figure 1**, lower left). This interpretation converges with Singmann and Klauer's (2011) finding, based on state-trace analysis, that affirmation and denial problems may depend on different processes.

Why use ROC analysis rather than simply examine the raw validity judgments? Interpreting the latter often relies on assumptions that may not be justified (Klauer et al., 2000; Dube et al., 2010). The main advantage of a formal model like SDT lies in its specification of assumptions. However, models can also produce

insights that are not obvious at first glance. A qualitative difference between affirmation and denial inferences is not necessarily predicted by extant theories. Moreover, various manipulations seem to exert a similar effect on both types of inferences (e.g., Cummins, 1995). Finally, it is interesting to note that the production of linear ROCs when performance is driven by multiple underlying processes has been predicted in theory (DeCarlo, 2002). These results may offer a case study of how this can occur in practice.

REFERENCES

- Cummins, D. D. (1995). Naive theories and causal deduction. *Mem. Cognit.* 23, 646–658. doi: 10.3758/BF03197265
- DeCarlo, L. T. (2002). Signal detection theory with finite mixture distributions: theoretical developments with applications to recognition memory. *Psychol. Rev.* 109, 710–721. doi: 10.1037/0033-295X.109.4.710
- Dube, C., Rotello, C. M., and Heit, E. (2010). Assessing the belief bias effect with ROCs: it's are-spone bias effect. *Psychol. Rev.* 117, 831–863. doi: 10.1037/a0019634
- Dube, C., Rotello, C. M., and Heit, E. (2011). The belief bias effect is aptly named: a reply to Klauer and Kellen (2010). *Psychol. Rev.* 118, 155–163. doi: 10.1037/a0021774
- Elqayam, S., and Evans, St. B. T. (2011). Subtracting 'ought' from 'is': descriptivism versus normativity in the study of the human thinking. *Behav. Brain Sci.* 34, 233–248. doi: 10.1017/S0140525X1100001X
- Evans, J. St. B. T., Barston, J. L., and Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Mem. Cognit.* 11, 295–306. doi: 10.3758/BF03196976
- Evans, J. St. B. T., Handley, S. J., and Over, D. E. (2003). Conditionals and conditional probability. *J. Exp. Psychol. Learn. Mem. Cogn.* 29, 321–355. doi: 10.1037/0278-7393.29.2.321
- Evans, J. St. B. T., and Over, D. E. (2004). *If*. Oxford: Oxford University Press.
- Evans, J. St. B. T., and Over, D. E. (2012). Reasoning to and from belief: deduction and induction are still distinct. *Think. Reas.* 19, 267–283. doi: 10.1080/13546783.2012.745450
- Evans, J. St. B. T., Over, D. E., and Handley, S. J. (2005). Suppositions, extensionality, and conditionals: a critique of the mental model theory of Johnson-Laird and Byrne (2002). *Psychol. Rev.* 112, 1040–1052. doi: 10.1037/0033-295X.112.4.1040
- Heit, E., and Rotello, C. M. (2010). Relations between inductive reasoning and deductive reasoning. *J. Exp. Psychol. Learn. Mem. Cogn.* 36, 805–812. doi: 10.1037/a0018784
- Heit, E., and Rotello, C. M. (2014). Traditional difference-score analyses of reasoning are flawed. *Cognition* 131, 75–91. doi: 10.1016/j.cognition.2013.12.003
- Johnson-Laird, P. N., and Byrne, R. M. J. (2002). Conditionals: a theory of meaning, pragmatics, and inference. *Psychol. Rev.* 109, 646–678. doi: 10.1037/0033-295X.109.4.646
- Klauer, K. C., and Kellen, D. (2011). Assessing the belief bias effect with ROCs: reply to Dube, Rotello, and Heit (2010). *Psychol. Rev.* 118, 155–164. doi: 10.1037/a0020698
- Klauer, K. C., Musch, J., and Naumer, B. (2000). On belief bias in syllogistic reasoning. *Psychol. Rev.* 107, 852–884. doi: 10.1037/0033-295X.107.4.852
- Macmillan, N. A., and Creelman, C. D. (2005). *Detection Theory: A User's Guide, 2nd Edn.* Mahwah, NJ: Erlbaum.
- Oaksford, M., and Chater, N. (2009). Précis of Bayesian rationality: the probabilistic approach to human reasoning. *Behav. Brain Sci.* 32, 69–84. discussion: 85–120. doi: 10.1017/S0140525X09000284
- Oaksford, M., and Chater, N. (2013). Dynamic inference and everyday conditional reasoning in the new paradigm. *Think. Reas.* 19, 346–379. doi: 10.1080/13546783.2013.808163
- Oaksford, M., Chater, N., and Larkin, J. (2000). Probabilities and polarity biases in conditional inference. *J. Exp. Psychol. Learn. Mem. Cogn.* 26, 883–889. doi: 10.1037/0278-7393.26.4.883
- Pazzaglia, A. M., Dube, C., and Rotello, C. M. (2013). A critical comparison of discrete-state and continuous models of recognition memory: implications for recognition and beyond. *Psychol. Bull.* 139, 1173–1203. doi: 10.1037/a0033044
- Singmann, H., and Klauer, K. C. (2011). Deductive and inductive conditional inferences: two modes of reasoning. *Think. Reas.* 17, 247–281. doi: 10.1080/13546783.2011.572718
- Trippas, D., Handley, S. J., and Verde, M. F. (2013). The SDT model of belief bias: complexity, time, and cognitive ability mediate the effects of believability. *J. Exp. Psychol. Learn. Mem. Cogn.* 39, 1393–1402. doi: 10.1037/a0032398
- Verde, M. F., Stone, L. K., Hatch, H. S., and Schnall, S. (2010). Distinguishing between mnemonic and attributional sources of familiarity: positive emotion bias as a case study. *Mem. Cognit.* 38, 142–153. doi: 10.3758/MC.38.2.142

Received: 30 November 2013; accepted: 25 February 2014; published online: 12 March 2014.

Citation: Trippas D, Verde MF, Handley SJ, Roser ME, McNair NA and Evans JSBT (2014) Modeling causal conditional reasoning data using SDT: caveats and new insights. *Front. Psychol.* 5:217. doi: 10.3389/fpsyg.2014.00217

This article was submitted to *Cognitive Science*, a section of the journal *Frontiers in Psychology*.

Copyright © 2014 Trippas, Verde, Handley, Roser, McNair and Evans. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.