Running head: INFERENTIAL REASONING AND ATTITUDINAL ENTROPY

An Inferential Reasoning Perspective to Attitudinal Entropy

Pieter Van Dessel, Sean Hughes, Ian Hussey, and Jan De Houwer

Ghent University, Belgium

**Author Note**

PVD, SH, IH, JDH, Department of Experimental Clinical and Health Psychology, Ghent University, Henri Dunantlaan 2, Ghent, Belgium. PVD and IH are supported by a Postdoctoral fellowship of the Scientific Research Foundation, Flanders (FWO-Vlaanderen). JDH is supported by Methusalem Grant BOF16/MET\_V/002 of Ghent University. Correspondence concerning this article should be sent to Pieter.vanDessel@UGent.be

An Inferential Reasoning Perspective to Attitudinal Entropy

Dalege, Borsboom, van Harreveld, and van der Maas (this issue) describe a novel framework for the conceptualization of attitudes by drawing on principles from statistical mechanics. Most importantly, they apply the idea that systems can be characterized by randomness (i.e., entropy) to the field of attitude research. We applaud their initiative because the attitudinal entropy framework provides an interesting new perspective to theoretical questions and empirical findings and offers new research directions. It is also a timely effort given the increase in popularity of predictive processing theories of human cognition that emphasize entropy as an important factor in human cognition (see Metzinger & Wiese, 2017, for an overview). Specifically, these theories assume that people strive to minimize entropy by building a mental model of the world that allows them to respond to the environment in an optimal fashion (Friston, 2010). It has been noted however, that the conceptualization of entropy in this framework is implausible and requires more work (e.g., Otworowska, Van Rooij, & Kwisthout, 2018). Moreover, application of the principle of entropy to social cognition is lacking.

In a first effort to link predictive processing theories into the field of attitudes, we recently delineated an inferential model of evaluative stimulus-action effects (Van Dessel, Hughes, & De Houwer, 2018). This model describes the inferences that underlie evaluative learning on the basis of stimulus-based actions (e.g., repeated approach or avoidance of a stimulus) and outlines how these inferences might arise based on predictive processing principles. Though this model focuses on evaluative stimulus-action effects it can easily be generalized to explain evaluative behavior in general. Specifically, evaluative responding is considered to result from inferences about (the value of) action outcomes. These inferences are learning-, context, and goal-dependent, and reflect the (automatic) application of inference rules to activated information on the basis of a person’s belief network. This belief network can be seen as a generative model of the world that is continuously updated on the basis of available information. Entropy might play a role here such that model updating is based on the weighing of the extent to which integration of new information increases entropy (which is unfeasible) compared to other factors such as the inferred value of the information.

The attitudinal entropy framework of Dalege and colleagues bears resemblance to our inferential model not only because it integrates the idea of entropy reduction as a determinant of attitudes but also because it assumes that implicit and explicit attitudes are based on a single type of mental process which incorporates inferential reasoning principles. Indeed, assessment of entropy (conceptualized as consistency of attitude elements) seems to require the validation of relational (i.e., propositional) information. Moreover, the authors note that “weights between attitude elements generally arise based on inferences” (p.12). This framework therefore diverges from attitude models that distinguish between two types of processes underlying attitudes: e.g., automatic and controlled (e.g., Fazio, Sanbonmatsu, Powell, & Kardes, 1986), or associative and propositional processes (e.g., Gawronki & Bodenhausen, 2006). This accords with recent recommendations to explore alternatives to dual-process theories of human cognition (e.g., Melnikoff & Bargh, 2018) which is especially relevant in attitude research where these theories are highly prevalent (without clear support by empirical findings: see Corneille & Stahl, 2018).

Importantly, however, there are also several important distinctions between our inferential reasoning model and predictive processing models on the one hand and the attitudinal entropy framework on the other hand (which might also explain why the authors do not explicitly refer to these models). Below, we highlight the distinctions that we consider most relevant for evaluation and refinement of this novel theoretical framework.

**Attitudes and evaluative behavior**

In the attitudinal entropy framework, an attitude is defined as the mental configuration of attitude elements (i.e., beliefs, feelings, and behaviors towards an attitude object). Attitudes are considered to directly relate to evaluative behavior such that global evaluations are modeled as the sum of node states (representing the valence of attitude elements: positive/negative). This conceptualization of attitudes has several important limitations.

First, because this framework focuses on the conceptualization of a mental construct (attitudes), it is difficult to determine its usefulness in the explanation and prediction of actual behavior (arguably the most important characteristic of good psychological theories). Indeed, predictions are often made at the mental (cognitive) level rather than at the behavioral (functional) level which makes it impossible to test these predictions outside the current framework (e.g., prediction 2: the size of the attitude network should predict the strength of the mere thought effect).

Second, it seems problematic to assume a more or less direct link between attitudes and evaluative behavior. It is well established that evaluative responses are determined by many (contextual) factors. For instance, one important aspect of evaluation is that it can occur under different conditions of automaticity (e.g., fast or unintentional: Moors & De Houwer, 2006) as found in studies that have used implicit evaluation measures. In their discussion of these studies, the authors assume that one aspect of automaticity (i.e., reduced attention to the stimulus) can explain relevant findings (because it changes the sum of node states). It should be clear, however, that attention to the attitude object is not necessarily reduced in implicit attitude measures (De Houwer Teige-Mocigemba, Spruyt, & Moors, 2009). In fact, attention to different aspects of the attitude object (e.g., specific attitude elements) might produce distinct effects (see Gawronski, Cunningham, LeBel, & Deutsch, 2010).

In our inferential model of evaluative stimulus-action effects (Van Dessel et al., 2018), we model evaluations (rather than attitudes) which we define as behavioral responses. This ensures that there is no conflation between the behaviors that need to be explained and the mental constructs that are used to explain these behaviors (see De Houwer, Gawronski, & Barnes-Holmes, 2013, for an in-depth discussion). The attitudinal entropy framework might benefit from this clear distinction between functional and cognitive levels of explanation, to allow for clear, testable predictions about the moderation of evaluative responses by specific contextual variables.

**The conceptualization of (attitudinal) entropy**

Dalege and colleagues describe entropy as a characteristic of attitudes (i.e., its uncertainty). Boltzmann entropy is formalized as the proportion of attitude elements in distinct states at a certain point in time whereas Gibbs entropy is formalized as the consistency of these configurations over time. Thinking about (or attention to) the attitude object is assumed to reduce Gibbs entropy. This formalization has merits because it is more elaborate compared to other (e.g., predictive processing) theories and it allows simulation of human evaluative behavior in a nodal network with a good fit to past findings and potential for novel predictions.

Importantly, however, even though consistency detection lies at the basis of the attitudinal entropy theory, it is not specified how this process occurs. Attitude elements are modeled as nodes that can only be switched on or off and are thus stripped from any relational content, making it difficult to see how consistency between attitude elements could be determined. The assumption that only the (momentary) valence of attitude elements is compared in this process is unfeasible given that it is not specified how the valence of attitude elements (not only beliefs but also behaviors and feelings) is determined. Moreover, studies show that the relational content of information about attitude objects influences evaluation (e.g., Van Dessel, De Houwer, & Smith, 2018) and important moderators of evaluation (e.g., diagnosticity of available information: see Cone & Ferguson, 2015) necessarily involve the validation of relational content of information.

Though Dalege and colleagues seem to assume that inferential reasoning processes determine attitude formation (as noted above), it is unclear why inferential reasoning processes are not incorporated in their calculation of entropy. At the very least, this calculation should take into account that evaluative responses not only depend on the amount of thinking about an attitude object but also the direction of thinking. The attitudinal entropy framework might benefit from a clearer specification of inferential reasoning processes that elucidates how people make specific inferences (in the context of information validation). Integration of basic principles from inferential reasoning models (e.g., context-dependence of inferences: Van Dessel et al., 2018) might allow for a more encompassing computation of attitude consistency. This might facilitate explanation of several findings that were not discussed by the authors yet do not readily fit within the attitudinal entropy framework (e.g., that implicit evaluations are sometimes found to predict certain behavior better than explicit evaluations: Banaji & Greenwald, 2013).

A more general inferential reasoning model of evaluations might be described that benefits from the elaborate conceptualization (and the specific method of computation) for entropy in the attitudinal entropy model. In this endeavor, we believe that it might be best to consider entropy not as a characteristic of an attitude (what would be the delineating factor of a configuration of attitude elements anyway?) but of a more general belief system. Modeling entropy might then allow estimation of the circumstances under which a person’s belief system might be updated. This could help elucidate how entropy might play a role in evaluative behavior and, by extension, behavior in general.

**Concluding Remarks**

The attitudinal entropy framework provides an intriguing new way to look at attitudes. It has heuristic and predictive value as it speaks to a wide variety of phenomena using a restricted set of concepts. It can simulate past findings and offers new research directions and questions. However, to provide a useful model of human behavior that survives the test of time, this framework might require (1) integration in other perspectives (e.g., predictive processing theory), (2) specification of inferential reasoning and its determinants (in the context of evaluative behavior), and (3) a clear distinction between different levels of explanation.

References

Banaji, M., & A. Greenwald. (2013). Blindspot: Hidden biases of good people. New York: Delcorte Press.

Corneille, O., & Stahl, C. (2018). Associative Attitude Learning: A Closer Look at Evidence and how it Relates to Attitude Models. *Personality and Social Psychology Review.*

De Houwer, J., Gawronski, B., & Barnes-Holmes, D. (2013). A functional- cognitive framework for attitude research. *European Review of Social Psychology, 24*, 252–287. doi:10.1080/10463283.2014.892320

De Houwer, J., Teige-Mocigemba, S., Spruyt, A., & Moors, A. (2009). Implicit measures: A normative analysis and review. *Psychological Bulletin, 135*, 347-368. doi:10.1037/a0014211

Fazio, R. H., Sanbonmatsu, D. M., Powell, M. C., & Kardes, F. R. (1986). On the automatic activation of attitudes. *Journal of Personality and Social Psychology*, *50*, 229–238. doi: 10.1037/0022-3514.50.2.229.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience,* 11:127–38.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: an integrative review of implicit and explicit attitude change. *Psychological Bulletin*, *132*, 692–731. doi: 10.1037/0033-2909.132.5.692

Gawronski, B., Cunningham, W. A., LeBel, E. P., & Deutsch, R. (2010). Attentional influences on affective priming: Does categorization influence spontaneous evaluations of multiply categorizable objects? *Cognition and Emotion, 24*, 1008-1025.

Melnikoff, D.E., & Bargh, J.A. (2018). The Mythical Number Two. *Trends in Cognitive Sciences, 22*, 280-293. doi: 10.1016/j.tics.2018.02.001.

Metzinger, T., & Wiese, W. (2017) *Philosophy and predictive processing*. Frankfurt am Main: MIND Group.

Moors, A., & De Houwer, J. (2006). Automaticity: A Theoretical and Conceptual Analysis. *Psychological Bulletin*, *132*, 297–326. doi: 10.1037/0033-2909.132.2.297

Otworowska, M., van Rooij, I., & Kwisthout, J. (2018). Maximizing entropy of the Predictive Processing framework. https://doi.org/10.31234/osf.io/5zam7

Van Dessel, P., De Houwer, J., & Smith, C. T.  (2018). Relational Information Moderates Approach-Avoidance Instruction Effects on Implicit Evaluation. ​*Acta Psychologica, 184*, 137-143. doi:[10.1016/j.actpsy.2017.03.016](http://dx.doi.org/10.1016/j.actpsy.2017.03.016).

Van Dessel, P., Hughes, S., & De Houwer, J. (2018). How Do Actions Influence Attitudes?   
An Inferential Account of the Impact of Action Performance on Stimulus Evaluation. Manuscript invited for revision at *Personality and Social Psychology Review*. Preprint available at: https://osf.io/kb3wq/