**Belief Change Under Conditions of Moral Conviction**

Moral convictions are attitudes that are perceived as grounded in fundamental distinctions between right and wrong. Moral conviction, while often correlated with strength of belief, is fundamentally distinct from strong but non-moral attitudes/beliefs. This is because they are perceived as a universal and objective truth, which is relatively resistant to influence from equals or superiors. Thus, morally convicted beliefs are more difficult to change as compared to beliefs grounded in preference or social convention (Skitka et al., 2021). Our research question is fundamentally, how can we change attitudes that are backed by moral conviction.

Given that morally convicted beliefs are hard to change, one plausible solution would be to first ‘demoralize’ the belief, and then attempt change. However, there is mixed evidence that this is effective. Moral conviction has been successfully reduced by framing arguments using pragmatic or economic counterarguments (Kodapanakkal et al., 2022, Kutlaca, 2013), or by emphasizing the personal or hedonic benefit of a counter position (Bastian et al, 2015; Feinberg et al, 2019). In contrast, unsuccessful reduction of moral conviction has been attributed to choosing topics that are already highly politicized and polarized (e.g., COVID-19 vaccination), framings that are dependent on belief in authority influence, or flawed psychometric measurements of moral conviction itself (Aignesberger et al., 2023; Fenzi et al., 2022; Brannon et al., 2019).

For our first study, we plan on testing several methods of moral conviction reduction (pragmatic/economic argument, and a personal/hedonic benefit argument), across an array of contemporary issues that falls along the spectrum of political belief and polarization (support for universal health care, capital punishment, and desire to exercise). Additionally, we plan on using an expanded item for measuring moral conviction that is an adaption of work by Skitka et al., (2021), which is novel insofar as it will measure perceptions of objectivity and universality in belief, which has been assumed but has not been directly assessed. Our goal is to directly reduce moral conviction on these stances, any change in behavior or belief would be a useful, but incidental benefit.

As our final goal is change in attitudes, after ‘demoralization’ we would like to leverage the effects of social influence. One of the strongest findings in psychology is that people conform towards the consensus group opinion (Asch, 1956; Deutsch M, 1955). However, one aspect of morally convicted beliefs that sets them apart from simple ‘strongly held’ beliefs, is that they appear to be independent of normative/majority influence (Skitka et al., 2005). In preliminary work, we were able to successfully manipulate social consensus by presenting false survey results that were assumed to be real, presenting either an artificially high or artificially low level of agreement with a position. We plan to directly test this interaction by attempting to reduce moral conviction and then seeing if that increases conformation towards the societal consensus.

According to the domain model of attitude moralization, the process of attitude moralization is a two-part process. The first part is the shift from seeing an initial attitude as being grounded in preference or opinion (e.g., Coke vs. Pepsi) to instead being seen as having moral significance. This initial step is labeled as “Moral Recognition”. Moral recognition can occur when activities previously seen as non-moral (e.g., eating meat) become connected to already pre-existing moral beliefs (i.e., ending life is immoral, eating meat requires the ending of life, thus, eating meat is immoral). This process is defined as ‘Moral Piggybacking’. Moral recognition can also occur when previously unknown moral objections to one’s preferences are made salient (e.g., learning that a brand of luxury goods is produced by slave labor). Finally, moral recognition can occur when individuals are induced to attach strong emotions (e.g., disgust or anger) to the act or concept that is being moralized (i.e., eating meat being moralized through the viewing of videos that show animal suffering inflicted by the meat industry).

The second process is the increase in moralization of already moralized attitudes, wherein a relatively less moralized attitude becomes more moralized. One aspect in which moral amplification differs from moral recognition is that people who hold weakly moralized attitudes are likely already aware of societal norms that exist both in support and opposition to their beliefs. Persuasive arguments framed using specifically moral language, centered on perception of harm, rights, and liberties have also been shown to be effective for moral amplification. Additionally, as has already been seen in the process of moral recognition, changes in attitude-specific emotions (e.g., happiness, excitement, anger, and disgust) predict parallel changes in extremity of attitude moralization. While the two concepts of moral amplification and recognition are theoretically distinct, many of the psychological factors affecting one also affect the other.

Distinguishing these two concepts opens the door to future work that delineates which factors may be more applicable to one process or the other. Furthermore, as new technologies and societal changes develop, the process of moral recognition becomes increasingly salient. For example, moral recognition is pertinent when considering AI chatbots, as an entirely novel concept, the public in general has no preconceived notions with regards to its morality. This can be contrasted with contemporary topics that have been discussed for years, such as abortion, which lends itself more to studies related to further amplification or demoralization of an already morally weighty topic.

One goal of our research is to be successful at ‘demoralizing’ beliefs held with moral conviction. In the previous literature, while some researchers have been successful in ‘demoralization’, other researchers have been unsuccessful (Brannon 2019; Clifford 2017). Additionally, some research by Asadullah and colleagues (2019) unexpectedly found a null effect of moral conviction. The primary shared issue that these studies hold was an improper parameterization of the psychometric measurement of moral conviction itself.

For example, Brannon and colleagues (2019) were unable to reduce moral conviction as they defined it, on attitudes related to genetically modified organisms used as food. However, they also acknowledged that their null effect on moral conviction was likely due to the poor psychometric properties of the single-item measure they adapted from Skikta and colleagues (2005). In general, single item measures have significant concerns regarding both reliability and validity. Prior literature indicates that multi-item measures are superior in several important psychometric properties (Sarstedt 2007). Multi-item measures have greater reliability. This is due to combining numerous items, which averages out random error in measurement. Test-retest reliability is also improved. Multi-item measures have superior segmentation in results, leading to greater variability, which improves correlation with the goal criterion. Another practical advantage is that multi-item measures can use data imputation to address issues with item non-response.

Furthermore, prior research on moral conviction has significant concerns regarding adequate reliability and validity of the most popularly used measures, Skitka and colleagues single item screener (Skitka et al., 2005) and the follow-up multi-item screener developed later (Skitka et al., 2017). Primarily, there is undue reliance on the ‘face validity’ of the measures themselves as adequate justification for their usage. However, validity theory researchers clearly state that ‘the appearance of validity does not constitute scientific evidence’ (Mozier 1947; Royal 2016). Furthermore, the convergent validity of these items was tested by seeing how well it predicted a similarly ‘face valid’ measure of moral conviction (e.g., the degree that participants indicate their attitudes on an issue were connected to their fundamental beliefs of right and wrong). Given that the secondary measure was only ‘face valid’, it brings into question how confident the researchers can be regarding the convergent validity of their original item. It is important to note that the multi-item measure had reasonable discriminant validity from similar, but theoretically distinct topics (intolerance, policy acceptance, etc.), and test-retest reliability likewise high for the multi-item measure after 1 and 3 month intervals.

The combination of these issues indicate that successful experimentation on the concept of moral conviction requires a reliable and valid multi-item measure, and that prior literature has done an inadequate job assessing construct validity. One goal of our research is to assess and develop a more holistic moral conviction scale. Our initial step in developing this scale is clarifying the intended meaning and breadth of the theoretical construct of moral conviction, as recommended by Carpenter and colleagues (2012). We refer to the most unique aspects of moral conviction, which is the perception of universality and objective truth (Skitka et al., 2021). This aspect is what separates attitudes grounded in fundamental distinctions of right and wrong from strong but nonmoral attitudes and beliefs.

Piggybacking off of prior work on moral conviction, our scale will be an extension of the 3-item moral conviction scale developed by Skitka and colleagues, which indicated strong test-retest reliability and discriminant validity, by adding additional measures of universality and objective truth. Our measure of universality is taken from prior literature on measurement of subjective universality (Van Zomeran et al., 2012) and our measurement of ‘objective truth’ is adapted from prior literature on the psychosocial perception of objective truth (Goodwin and Darley 2010: Goodwin et al., 2008). Our plan is to follow the general procedure outlined by Scherer and colleagues (2016) in their development of another novel measurement, the medical maximizer-minimizer scale. We plan to test divergent validity by assessing this expanded construct of moral conviction against other facets of attitude perception, specifically, attitude importance and attitude certainty. We plan to test convergent validity by assessing brief choice vignettes directly relating to the hypothetical moral decisions associated with each issue. If possible, we would also like to assess test-retest reliability with our study sample, with a 1 to 3 month interval between assessment. Ideally, this work would result in a measure of moral conviction that is clearly operationalized, reliable, and valid, which should lead to more certainty in the success or failure of our moral conviction manipulations.