# Conclusions, Interpretations, and Recommendations

The purpose of this dissertation was to understand and change polarized beliefs. To do this, we created a set of experiments directly testing the effectiveness of increasing or decreasing social consensus on support for a variety of polarized topics (Study 1). We clearly showed that across a variety of topics, a social consensus manipulation can both increase and decrease support. We found significant support for our first Study 1 hypothesis (H1: high social consensus would lead to more positive support for highly polarized issues). There was a significant time x condition interaction, such that there was a greater increase over time in support for all three polarized issues that were manipulated (UHC, capital punishment, climate change) in our high social consensus condition as compared to our low social consensus condition. We also found mixed evidence supporting our second Study 1 hypothesis (H2: utilitarian and deontological orientation would be significant predictors of [topic] support). Deontological orientation was a significant predictor for support of UHC, but not for capital punishment or climate change. Furthermore, there was no effect of utilitarian orientation on final support for any of our topics.

We then executed on another series of experiments to understand the effects of moral conviction manipulation on an expanded set of polarized and non-polarized topics (Study 2). We found mixed evidence supporting our first Study 2 hypothesis (H1: the moral conviction manipulation would be a significant predictor of support for each topic, compared to the control condition). There was no main effect of our moral conviction manipulation on support for any of our four Study 2 topics (UHC, capital punishment, climate change, exercise). However, there was a significant interaction with our moral conviction manipulation and openness to belief change, such that being assigned to the ‘pragmatic’ moral conviction manipulation reduces the impact of ‘openness to belief change’ on support for [topic]. We did not find evidence supporting our second Study 2 hypothesis (H2: the two ‘moral’ interventions would increase, and the pragmatic and hedonic interventions would decrease, moral conviction relative to the control). There was no main effect of moral conviction manipulation on perceived moral conviction for any of our four topics.

Study 1 & 2 covered the impact of moral conviction and social consensus individually, and in doing so, laid the groundwork for Study 3, an experiment testing the joint interaction of social consensus and moral conviction on support for a polarized topic. We did not find evidence supporting our first Study 3 hypothesis (H1: High social consensus would lead to more support for issues). There was no main effect of our social consensus manipulation on support for any of our three Study 3 topics (UHC, capital punishment, usage of AI in the workforce). However, we did find a significant pre-post intervention increase in support for each topic in all conditions. Additionally, we did not find evidence supporting our second Study 3 hypothesis (H2: increased moral conviction will reduce the effect of social consensus and decreased moral conviction will increase the effect of social consensus). There was no significant interaction between the effects of our social consensus and moral conviction manipulations.

In light of the full set of results, some preliminary conclusions seem reasonable to draw. We were not successful at empirically replicating the finding that moral conviction inoculates individuals from the effects of social consensus (Skitka, 2021). One plausible explanation is that our moral conviction manipulations were improperly designed. While previous literature has indicated that framing arguments using moral terms (e.g., freedom, liberty, etc.) or centering on perceptions of harm (e.g., harmful, dangerous, contaminated, etc.) increases perceptions of moral conviction (Kodapanakkal 2021; Clifford, 2019), we were unable to successfully replicate this pattern in Study 2 or 3. Plausibly, this is due to the fact that the topics covered in prior literature were generally seen to be unpolarized, or at least not explicitly polarized (e.g., hiring algorithms, fish farming), whereas in Study 2 and 3, the majority of our topics were explicitly polarized (UHC, capital punishment, climate change, etc.).

Additionally, we were not able to find support for the existence of an interaction between social consensus and moral conviction. One conclusion is that this potentially stems from the difficulty we had in successfully manipulating moral conviction. Thus, after developing further improvements in moral conviction manipulation, we could plausibly replicate this study and see significant results. However, another explanation is due to the shared orientation of the arguments in Study 3; all of the essays were ‘in favor’ of the topic in question. A plausible conclusion is that the directionality of the persuasive arguments (e.g., all in favor) was the most significant factor, as that interpretation falls in line with the results we observed. We could investigate this further in future studies by presenting persuasive essays that are both in favor and opposition to the topic in question.

We were able to successfully manipulate perception of social consensus using a very simple experimental manipulation adapted from Kobayashi (2018). This bodes well for the generalizability of this method, as our intervention was successful even though the topics we chose (UHC, climate change, capital punishment), diverged significantly from the original set of topics Kobayashi chose to use (climate change, blood type personality, nuclear power, and whale research). Additionally, it is important to note that while the version of the manipulation that we used for this study explicitly relied on deceiving our participants as to the base rate of consensus for each of our topics, the result that there are substantive differences in support due to shared perceptions on what is popular or unpopular should generalize broadly. Deceiving the public in order to manipulate support for a topic would generally be seen as unethical, but increasing salience of the public as to how much consensus there actually is (i.e., just because the American public supports something in general, that doesn’t mean any given individual is aware of it) could be a low-cost intervention that leverages the strengths of social consensus effects.

One direction for future work on polarized belief formation and change that could be worth exploring is to delve deeper into using domain specific information. In a pragmatic sense, one approach would be to conduct extensive qualitative research with individuals that have polarized beliefs, so as to determine which shared traits or characteristics of the topic are seen with a moral lens. Domain specific pragmatic or moral arguments would plausibly be more effective at changing moral convictions, as compared to using the ‘general’ framework of arguments based on morality and harm (increasing moral conviction) or arguments based on economics and practical implementation (decreasing moral conviction).

Another direction for future study would be to explore different methods for obtaining evidence of actual revealed preferences, rather than relying on self-reported support for a topic score. This would significantly increase the external validity of the conclusions that could be drawn. In a practical sense, many beliefs do not lend themselves to easily revealed preferences, so this suggestion would not be viable for all topics.

Finally, one future goal for this research would be to see if the effect of social consensus on belief formation and change functions differently based on the type of social consensus manipulation. The manipulation we used in our research earlier was effective, but relatively impersonal. It would be very useful to research whether the effects of social consensus are greater in small or large group settings, where the relevant comparison group isn’t the nebulous concept of ‘Americans’ as a whole, but instead the social group immediately and physically around you. This seems especially pertinent, as this structure mimics actual human social dynamics (i.e., social consensus is assessed and formed through shared, in person experience), and thus has greater external validity.

In summary, we find that manipulating social consensus seems to affect support for a variety of polarized topics, such that greater perception of social consensus in favor of a topic is associated with increased support, and the obverse is associated with decreased support. Furthermore, we found that we were unable to successfully manipulate moral conviction for several polarized topics by framing arguments using moral terms (e.g., freedom, liberty, etc.) or centering on perceptions of harm (e.g., harmful, dangerous, contaminated, etc.). Given this, we could not find evidence in support of a significant interaction between the effects of social consensus and moral conviction on support for a polarized topic, as we previously theorized. Developing a better understanding of how to manipulate moral conviction is necessary to explore this potential interaction further.