# Chapter 1: Introduction – Broad Overview of Research

It is 2018; an ordinary American is standing in a foot of snowfall after an unexpectedly long cold snap (National Oceanic and Atmospheric Administration, Snowfall Extremes – Missouri, 2024). He communes with his neighbors over beer and they all bond over grousing about how cold it’s been recently. On that very same day, the president of the United States gives his thoughts on climate change:

“There is a cooling, and there’s a heating. I mean, look, it used to not be climate change, it used to be global warming. Right . . . That wasn’t working too well because it was getting too cold all over the place. (Donald J. Trump, ITV, January 2018)”

After vindication by a trusted authority, seeing the ‘obvious’ evidence around themselves, and reaching a consensus with their immediate peers, this individual feels justified in disbelieving the existence of global warming. In reality, the overwhelming majority of scientific evidence indicates that regardless of cold weather and the opinion of politicians, global warming is indeed a real phenomenon (Davis et al., 2020). While the scenario is somewhat contrived, the above story is a distressingly common example of how misinformation can drive polarized beliefs.

Providing information regarding best practices, scientific consensus, and state-level policy priorities in general can directly improve quality of life for ordinary people. However, as our example above shows, dissemination of up-to-date information can be objectively worthless if the message itself does not adequately lead to belief change for the behavior at hand. The world experienced this firsthand during the COVID-19 pandemic; large amounts of misinformation were spread regarding what medicine could work (e.g., hydroxychloroquine and ivermectin), what the cause of the disease was (e.g., natural origin, Chinese laboratory leak), and what were best practices to reduce disease spread (e.g., cloth masking vs N95 masking). This general swell of misinformation contributed to COVID-19 vaccine hesitancy, which was estimated to result in at least 232,000 preventable deaths. Furthermore, a significant majority of those who refused vaccination had been exposed to persuasive misinformation through social media and other outlets (Jia et al., 2023; Lee et al., 2022). Changing beliefs also has common application with regards to political policy; Universal Health Care (UHC) has been shown to result in better overall population health outcomes, yet the United States is the only country in the developed world that still lacks Universal Health Care (UHC) for its citizens (Alspaugh, 2021). Improving generally negative attitudes towards UHC in the United States (36% of Americans support UHC) would increase the likelihood of UHC adoption, and in doing so, improve population health (Pew Research Center, 2020).

Belief change is a multifaceted process, and prior literature lists many factors affecting openness to belief change for a given topic. Individual differences in (a) deontological and (b) utilitarian orientation broadly shape priorities and how issues are interpreted; (c) social consensus (e.g., the level of agreement on an issue amongst friends, family, peers, or other in-groups) consistently influences individuals; and (d) differences in attitudes due to core beliefs about what is fundamentally right or wrong (e.g., abortion should be legal, due to the core belief that women should have full bodily autonomy), which we define here as Moral Conviction, ‘inoculates’ individuals against changing their beliefs. Assessing how each of these individual factors interact and influence belief change broadly has real and direct implications for public health and safety.

# Chapter 2: Review of the Literature

## Attitude/Belief Formation

The broader literature of attitude formation directly informs the process of belief change. Historically, attitude formation was one of the cornerstones of early psychological research, originating with Thomas & Znaniecki (1918) and Jung (1923). Attitudes represent an evaluative integration of cognition and affect in relation to a ‘subject’, which can be an object, person, or an abstract idea (Crano and Prislin 2006; Albarracin and Shavitt, 2018). Prior literature theorized that attitude was a fixed memory, preserved in amber until recollection was necessary; Another alternative perspective was that attitudes were constructed ‘in-the-moment’, based on ephemera, such as individual mood or outside temperature (Schuldt et al., 2011). Contemporary research however indicates that attitudes are in fact a composite of both elements, the structure of which allows attitudes to both maintain consistency and flexibility, as appropriate (Albarracin et al., 2005).

Understanding attitude measurement is the next step when attempting to understand attitude/belief formation. Historically, self-report scales have been employed to measure attitudes, using numeric responses on single items or aggregates. These explicit measures of attitude are both popular and effective for measuring attitudes that people are willing and able to accurately report (Himmelfarb, 1993; Greenwald & Banaji, 1995). However, many attitudes exist in which motivated response bias precludes explicit measurement; the Implicit Association Test (IAT) developed by Greenwald and colleagues (1998) directly addresses these issues. The core principle of the IAT assumes that attitudinal evaluation by a subject will manifest itself as differential response time, ostensibly representing a ‘true’ measure of attitude unrelated to social desirability or conscious awareness. As measurement has improved, contemporary research finds that attitude support is bipolar, and not two ends of a single continuum; Information processing is more effortful at high levels of ambivalence, as compared to strong love or strong hate (Van Harreveld et al., 2004). Furthermore, accurate measure of attitudes necessitates awareness that attitudes are multifaceted, with attitude importance, attitude accessibility, attitude commitment, and attitude certainty all separate elements that are jointly evaluated to assess the broader concept of ‘attitude strength’ (Bizer & Krosnick 2001; Holland et al., 2003).

## Attitude Change

Research on attitude change has historically been aggregated under two broad umbrellas, attitude change based on the effects of persuasion, and attitude change based on social influence (O’Keefe, 2016). In a broad sense, persuasion here means influencing based on the strength of detailed argumentation, irrespective of the source or context of the interaction. In contrast, social influence relies on appeals about the position of the source (e.g., from the head of the center of disease control). This aligns with the contemporary consensus behind the dual process theory of belief change (an adaptation of the dual process theory of Kahneman & Tversky, 1974), which posits that under conditions of low motivation or lack of ability, attitude judgements are based on heuristics and shortcuts (System 1 thinking, e.g., Mom’s always right!) and in conditions of high motivation and care, judgement is based on systematic assessment (System 2 thinking) of the information (Chen & Chaiken 1999). Prior research indicates that attitude judgements based on System 2 thinking have significantly more confidence, while those based on System 1 thinking were less resistant to change and less stable (Petty & Wegener 1999); Kassin & Kiechel (1996) found that in a reaction time task relying either on system 1 or system 2 thinking, false accusations of negligence were convincing (e.g., the subject admitted that they did not ‘hit the button’ even if they did), but only when individuals were relying on System 1 thinking.

An alternative model for attitude change labeled the ‘Unimodel’ posited by Kruglanski and Thompson (1999) claims that both cues/heuristics and message argumentation are parts of a larger category of information, defined as ‘persuasive evidence’. Thus, the content of the information itself, and not the route of processing that is important. The Unimodel states that differing information contents (e.g., is this a heuristic, or a detailed argument?) are analogous to whipped cream in a can versus whipped cream in a tub; the distinctions between them are irrelevant insofar as it relates to how ‘persuasive evidence’ works to change minds. However, recent studies indicate that the Unimodel has not been shown to sufficiently explain attitude change beyond the dual process models (Hedhli, 2022). The dual process model has been shown to be more predictive than a unimodal framework in several studies examining direct practical applications of advertising, retail experiences, and branding (Maheswaran, Mackie,and Chaiken 1992; Richard and Chebat 2016). This literature indicates that the source of the cue in an advertisement (e.g., famous football player) is an especially influential cue for persuasion in conditions of low cognitive capacity; likewise, that the persuasive function of a ‘brand name’ significantly increases when the recipient is highly engaged, but is significantly blunted when the recipient is more passive.

Attitude change can also occur on a societal level, where generational changes reflect commensurate changes in attitudes. For example, political polarization has increased significantly for 12th graders in the 2010s as compared to prior decades (Twenge et al., 2016), or the “Obama Effect” from 1992-2008 wherein election surveys indicated that amongst white participants, belief in the intelligence and work ethic of Black Americans significantly increased (Welch & Sigelman, 2011). In general, attitude change can originate from social pressures (either individual peers or society more broadly) or from information describing the attitude object (persuasive, fact based, argumentation).

## Utilitarian and Deontological Orientation

Utilitarian and deontological orientation influences how individuals perceive issues, as well as concepts of ‘right’ and ‘wrong’, which directly impact openness to attitude change as well as the effectiveness of persuasion (Brady and Wheeler, 1996). Utilitarian reasoning can be defined as ethical judgement based on outcomes, not intentions. In contrast, Deontological reasoning can be defined as ethical judgement based on whether or not behavior adheres to a preconceived set of ‘rules’, this includes concepts like ‘rights’, ‘ideals’, and explicitly recorded law.

Many highly polarized beliefs are rooted in utilitarian and deontological values, and thus are especially important when considering belief change (Tseng, 2021). For example, during the COVID-19 pandemic, healthcare professionals were forced to adopt utilitarian policies (e.g., mandatory vaccination, mandated isolation), which many found unacceptable. This has resulted in ‘moral injury’ arising from the conflict between individual deontological moral judgement and organization/country level utilitarian moral judgements (Akram, 2021). In the UK during the pandemic, the government traded off dignity and togetherness (a deontological value) for increased safety and reduced burden of disease (a utilitarian value), however, this resulted in significant ‘moral injury’ when doctors were forced to let patients die alone without seeing their families due to COVID-19 ‘no vistors allowed’ policies. In another context, arguments for the use of capital punishment are often based on deontological/utilitarian reasoning (Steiker, 2006). A deontological argument would be that killing an innocent is never an acceptable trade-off, thus the death penalty is unacceptable. Alternatively, that if the death penalty can prevent—through incapacitation of the offender or general deterrence—the loss to murder of even one innocent life, then it is a morally justified (or even required) penal response (e.g., any numbers of killers should die, if it saves one innocent). Conversely, a utilitarian argument would be that since the death penalty is a waste of resource (monetary cost), it is morally unjust to use it.

## Social Consensus

Social consensus directly relates to attitude formation and change in many ways. First and foremost, even if an individual does not intuitively hold a given attitude or belief, conforming to the majority opinion is extremely typical (Asch, 1956; Deutsch M, 1955). The effect of social consensus is also magnified under conditions of ambivalence; the experience of ambivalence itself motivates the search for corrective information, increasing susceptibility to social consensus (Hodson et al., 2001). Social consensus reliably impacts attitude formation and change in topics as broad as climate change, racial stereotyping, and weight discrimination (Goldberg, 2019; Stangor, 2001; Farrow, 2009). Higher amounts of perceived social consensus are associated with greater agreement with the consensus opinion, which remains a strong predictor even after controlling for demographic variables and other individual differences. Conversely, when a social consensus does not exist (new circumstances, or lack of agreement), individual judgement reigns (e.g., is it unethical to use AI to write letter to grandma?). The explicit lack of social consensus also can affect attitude formation and belief change; exposure to minority dissent (e.g., a dissenting attitude with low social consensus) increases critical thinking and divergent problem analysis regarding one’s original belief (De Dru & West, 2001). Being part of a social consensus also has inherent appeal; the process of having shared ‘negative attitudes’ (e.g., we both hate the Yankees), increases familiarity between people (Weaver & Bosson, 2011).

What normally would be the benefits of social consensus (familiarity, belief change, unity, etc.) become extremely hazardous when it results in propagation of misinformation. For example, stereotypes about the existence of “death panels” removing healthcare for the elderly or sick in the Affordable Care Act, has been seen as plausibly true even though thorough research has shown that description as factually wrong (Frankford, 2015; DiJulio, Firth, and Brodie 2014). Misinformation due to social consensus can occasionally even outweigh expertise; A survey of 9,972 otolaryngologists, conducted in 2013, found that 40 percent of the surgeons who are Republicans believed that the ACA created death panels, a percentage that stands in great contrast to the finding that only 8 percent of Democrats shared that belief (Rocke et al. 2014). Another recent example was from then presidential candidate Donald J. Trump (2015) speaking on the emerging social consensus amongst republicans about vaccination:

Autism has become an epidemic. Twenty-five years ago, 35 years ago, you look at the statistics, not even close. It has gotten totally out of control. … Just the other day, 2 years old, 2 and a half years old, a child, a beautiful child went to have the vaccine, and came back, and a week later got a tremendous fever, got very, very sick, now is autistic.”

Scientific consensus is clear in that vaccines do not cause autism; however, this has not resulted in broad social consensus, with roughly 6% of Americans believing that vaccines cause autism and over 50% being “unsure” about the presence or absence of a relationship (Newport et al., 2015). The dangers here are apparent, as increased social consensus (in some subgroups) that vaccination is related to autism has directly resulted in a 1.7x increase in US refusal to vaccinate children (Smith et al., 2008). Considering that the effect of social consensus can (depending on context) be harmful, understanding how to increase resilience against social influence is worthwhile. Interestingly, prior research has been done on the interaction between social consensus and deontology indicates that higher levels of deontological orientation results in less conformation to social consensus (Pincus, 2014). However, Pincus was not able to directly manipulate the level of social consensus, thus this interaction has not been directly empirically tested.

## Moral Conviction

Moral conviction is a separate element of attitudes that impacts belief change and openness to persuasion, in a way that is psychologically distinct from other attitude constructs (e.g., attitudes that are strong or certain are not necessarily highly moralized). For example, Wright and colleagues (2008), found that individual differences in moral conviction uniquely impact variables such as social distancing. Fundamentally, moral conviction reflects foundational beliefs about what is ‘right or wrong’ about a given attitude object, and beliefs rooted in moral conviction are perceived as objective and universal (Morgan & Skitka, 2020). In practice, this means that differing levels of moral conviction consistently predict how much an individual believes that their attitude about an issue is ‘objectively true’ and ‘universally applicable in all cases’. Conversely, beliefs with low levels of moral conviction are viewed as subjective preferences where legitimate disagreement is acceptable (Skitka, 2010). An example of how this applies is that individuals make faster evaluations (as measured through the Implicit Association Test) about if a given behavior is universally right or wrong, when the behavior is first evaluated as ‘morally’ right or wrong, as compared to being pragmatically ‘good/bad’ or ‘pleasant/unpleasant’ (Van Bavel et al., 2012).

Notably, there is significant disagreement on what beliefs people hold with moral conviction (Wright et al., 2008). Relatively few topics are ‘universally’ viewed with moral conviction (e.g., rape, incest, executing the mentally disabled). It is instead more common for issues to only be held with moral conviction for a subset of the population (e.g., gun ownership for the NRA, vegetarianism for PETA). Conversely, there are few beliefs that are ‘universally’ viewed as nonmoral (e.g., choosing to exercise, taste in music, etc.). This indicates that for every individual, many of their beliefs should be viewed through the lens of moral conviction when attempting to affect attitude change.

Prior literature in the field of attitude formation reinforces the idea that moral conviction directly affects belief change. Carpenter and colleagues (2013) Functional Attitude Theory states that beliefs about attitudes that have high moral conviction and represents something about oneself (e.g., I love recycling, which makes me a good person) are more easily changed when emphasizing image-based considerations and downplaying the intrinsic qualities of the attitude object (e.g., recycling because you are environmentally conscious versus recycling to get some monetary compensation). Another important interaction between moral conviction and attitude change is the ‘inoculation’ of individuals against the effects of peer and authority influence. Individuals that feel strong moral conviction about a belief do so because of a ‘greater moral purpose’ underlying the structures of authority, rather than the authority themselves (Kohlberg 1976, Rest et al. 1999). For example, strength of moral conviction about physician-assisted suicide, instead of prior perceptions of supreme court legitimacy/fairness, predicted whether or not an individual believed a supreme court judgement about physician-assisted suicide was reasonable (Skitka, 2009). In another case, levels of moral conviction predicted resistance to peer influence with regards to accepting the use of torture to deter terrorism (Aramovich, 2012); people continue to uphold morally convicted viewpoints, even when explicitly challenged by peers or authorities.

Given how much moral conviction impacts attitudes, directly manipulating moral conviction is a promising avenue to increase belief change. While some evidence indicates that the degree of perceived moral conviction can change, the mechanisms through which it does so are debated. For example, historical evidence indicates that some things that were once considered preferences (cigarette smoking in the 20’s-30’s) can evolve into morally weighted judgements (smoking seen as an ‘uncouth’ habit), that can even have real consequences (e.g., public smoking being banned in many venues) as the societal perspective about the concept changes (Rozin, 1999). In comparison, experimental manipulation of individual perceptions of moral conviction have found some success when using framing effects that center on arguments containing harm, fairness, or disgust, or alternatively, framing issues as ‘rights’ necessary for society (Kodapanakkal, 2021; Clifford, 2017; Wisneski & Skitka, 2017). However, this evidence is somewhat mixed, as Clifford and colleagues (2017) were unable to reduce moral conviction on ‘food politics’ e.g., support for factory farming, genetically modified food, animal welfare)

## Need for Further Research

Given the increasing pace of public perceptions and human belief that necessitates change, either due to new information (e.g., discovery of a new drug, or best practice) or due to changing circumstances (e.g., a global pandemic), understanding how to affect shifts in polarized attitudes is becoming increasingly important. Prior qualitative and quantitative research in the fields of attitude formation, deontology/utilitarian orientation, social consensus, and moral conviction indicate several key features that can inform how to best change polarized attitudes. However, many of these relationships have only been assessed through association and have not been empirically tested under experimental conditions. Furthermore, in theory, many of these features should interact with each other in significant ways, but these interactions have not been investigated either. Therefore, I propose a series of studies to determine: 1) How social consensus can be used to change polarized beliefs, 2) How moral conviction affects belief change, and 3) How changing levels of moral conviction interact with the effects of social consensus.