Insensitivity to Gradations in Warmth Traits Constrains Beliefs About Others’ Potential for Improvement

Nadav Klein

To cite this article: Nadav Klein (2015) Insensitivity to Gradations in Warmth Traits Constrains Beliefs About Others’ Potential for Improvement, Basic and Applied Social Psychology, 37:6, 348-361, DOI: 10.1080/01973533.2015.1111213

To link to this article: http://dx.doi.org/10.1080/01973533.2015.1111213

Published online: 23 Nov 2015.

Submit your article to this journal

View related articles

View Crossmark data
Insensitivity to Gradations in Warmth Traits Constrains Beliefs About Others’ Potential for Improvement

Nadav Klein
University of Chicago

Do people believe that others can improve their traits and personalities? Three experiments find that the answer depends on the trait in question. People believe that others’ potential for improving warmth traits is systematically lower than others’ potential for improving competence traits (Experiments 1a–1c). Consequently, shortcomings related to one’s warmth are considered harder to overcome than shortcomings related to one’s competence (Experiment 2). This asymmetry is partly rooted in people’s relative inability to detect gradations in warmth-related behaviors (Experiments 3–5). Enabling participants to more easily distinguish between moderate and high levels of warmth also increased their beliefs that other people can improve warmth traits. Beliefs about potential for improvement depend on how easily a trait can be evaluated.

Improvement often begins with believing that it is possible. Accordingly, believing that another person can improve in the future may increase the chances that this person indeed improves. Perhaps more poignantly, believing that another person cannot improve may create barriers for this person’s improvement (Spencer, Steele, & Quinn, 1999; Steele & Aronson, 1995). In politics, for example, a person who does not believe that her political opponents have the capacity to improve their seemingly intolerant dispositions may escalate conflicts rather than try to bridge differences (Chambers, Baron, & Inman, 2006; Waytz, Young, & Ginges, 2014). In the workplace, a manager who does not believe that her employees can improve their moral character may not foster the environment necessary to improve ethical conduct (Bazerman & Gino, 2012). In romantic relationships, failing to believe that one’s partner can improve interpersonal shortcomings may lead to a long-term deterioration of the relationship (Hui, Bond, & Molden, 2012; Kammrath & Peetz, 2012; Overall, Fletcher, & Simpson, 2010). More generally, if people believe that others’ potential for improvement is limited, then others’ failures may seem indicative of stable and enduring shortcomings. The present research examines people’s beliefs about others’ potential for improvement, and whether there are personality traits people are systematically unlikely to believe that others can improve.

IMPROVING WARMTH AND COMPETENCE

A productive starting point for investigating judgments of others’ potential for improvement is the distinction between warmth and competence, because people commonly judge others based on these two fundamental dimensions (Fiske, Cuddy, & Glick, 2007; Fiske, Cuddy, Glick, & Xu, 2002; Judd, James-Hawkins, Yzerbyt, & Kashima, 2005; Leach, Ellemers, & Barreto, 2007; Willis & Todorov, 2006; Wojciszke & Abele, 2008, 2014). Warmth encompasses traits related to the treatment of other people—moral and interpersonal traits such as kindness, sincerity, and trustworthiness. Competence encompasses traits related to the ability to achieve one’s goals and succeed in activities that require skill—traits such as confidence, independence, and intelligence. Because of the centrality of warmth and competence in social judgments, I focus on these two dimensions in assessing beliefs about potential for improvement.

At least two lines of research provide guidance about how people might assess the possibility of improving warmth and competence. The first includes experiments measuring beliefs about personality structure, which
find that people believe moral and interpersonal traits are more essential to one's identity than other types of traits (Dennett, 1978; Haslam, Bastian, & Bissett, 2004; Parfit, 1984; Strohminger & Nichols, 2014). Children appear to hold similarly essentialist views about interpersonal traits such as aggression (Giles, 2003). If people believe that moral and interpersonal traits are inherent to the structure of personality, then improving traits such as warmth, trustworthiness, and kindness may also mean fundamentally changing one's personality. In contrast, improving traits such as competence or confidence may not interfere with one's basic identity. This logic suggests that people may believe that improvement in warmth traits is more difficult than improvement in competence traits.

Another, seemingly more distant, line of research also raises a similar hypothesis. Research on evaluability suggests that the ease or difficulty with which people can identify gradations in a given attribute can profoundly affect judgments (Ariely & Loewenstein, 2000; Hsee, Loewenstein, Blount, & Bazerman, 1999; Hsee & Zhang, 2010; Parducci, 1965; Stevens, 1975). Evaluability refers to the ability to easily distinguish between lower and higher degrees of a given stimulus. For example, most people can more easily distinguish between low and high scores on the SAT test than between low and high scores on the TOFEL exam because of greater familiarity with the former and the relative obscurity of the latter. More generally, the ability to distinguish between different degrees of a given attribute varies greatly depending on the attribute being judged. This affects judgments, because highly evaluable attributes are judged differently from attributes that are difficult to evaluate. For example, one study finds that preferences for job candidates depended on people's ability to easily evaluate the attributes of these candidates (Hsee, 1996). Candidate attributes that were not easily evaluable (e.g., an ability to write "KY" computer programs) had a smaller effect on judgments than candidate attributes that could more easily be evaluated (e.g., GPA scores), because people could not evaluate the difference between high and low levels of "KY" programming ability.

More broadly, the more easily people can evaluate an attribute, the more likely their judgments are to closely track the magnitude of this attribute (Hsee et al., 1999). For example, if SAT scores are indeed easily evaluable, then people's judgments of relatively low SAT scores will be consistently lower than judgments of higher SAT scores. In contrast, if TOFEL scores are indeed difficult to evaluate for most people, then their judgments will not differentiate between low, moderate, and high scores on this test. At the extremes, a maximally evaluable attribute will lead to judgments that perfectly correspond to its magnitude, and a completely inevaluable attribute will lead to identical judgments regardless of its magnitude.

More relevant for the present research, warmth traits may be less evaluable than competence traits. Therefore, detecting gradations in behaviors that exhibit warmth may be more difficult than detecting gradations in behaviors that exhibit competence. For example, differentiating the competence of two students with close scores on a math exam may be easier than differentiating the warmth of two donors who gave slightly different amounts of money to a charity organization. Similarly, differentiating the competence of two employees may be easier than differentiating between these employees' warmth.

In turn, if warmth traits are less evaluable than competence traits, this may affect beliefs in the possibility of improvement in these types of traits. Improvement means moving from one level in a given attribute to another, higher level. Therefore, strong beliefs that improvement is possible in an attribute may partly depend on being able to identify fine degrees of that attribute. As gradations within a given attribute can more easily be identified, the next level that would constitute improvement over the present state becomes clearer. This would tend to increase beliefs in the possibility of improvement in that attribute. In contrast, if identifying fine degrees of a given attribute is difficult, then identifying the next level that would constitute improvement over the present state may accordingly be more difficult. This would tend to reduce beliefs in the possibility of improvement in that attribute.

### Warmth May Be Less Easily Evaluated Than Competence

Existing research provides some support for the hypotheses just outlined. Support for the hypothesis that warmth traits may be relatively inevaluable comes from recent studies finding that evaluations of warmth exhibit diminishing sensitivity to varying degrees of prosocial behavior, whereby increasingly larger magnitudes of warmth-related behaviors result in increasingly smaller changes in evaluations (Klein & Epley, 2014; Klein, Grossmann, Üskül, Kraus, & Epley, 2015). For example, one experiment (Klein & Epley, 2014, Experiment 3) finds that giving $0, $1, or $2 out of a $6 endowment to another person is evaluated more positively as the giving amount increases, but giving $3 or more out of $6 is evaluated similarly regardless of the amount given. Other experiments (Klein et al., 2015) replicate this result across seven countries, including Western countries such as the United Kingdom and Denmark, as well as Eastern countries such as China and Russia. These findings suggest that people may be relatively...
insensitive to magnitude when evaluating warmth-related behaviors, which in turn may carry over to people’s beliefs about improving warmth traits.

Support for the possible connection between evaluability and beliefs in the potential for improvement comes from at least two lines of research. The first was discussed earlier and focuses on the broad effects of evaluability on judgments. In this regard, others’ potential for improvement is an unexplored domain that may be affected by evaluability. The second focuses on goal setting and finds that setting specific rather than abstract goals can increase beliefs and motivation for improvement (Kurman, 2003; for a review, see Locke & Latham, 2002). This research broadly suggests that the ability to evaluate one’s progress toward a goal increases one’s beliefs in the potential for improvement. This idea raises the possibility that a similar process might operate in beliefs about others’ potential of improving traits.

OVERVIEW OF EXPERIMENTS AND ANALYTICAL APPROACH

Experiments 1a–1c measure people’s beliefs about the possibility that others can improve various warmth and competence traits over the lifetime. Experiment 2 then tests a consequence of beliefs about others’ improvement, namely, that personal shortcomings related to one’s warmth are considered harder to overcome than personal shortcomings related to one’s competence.

Experiments 3–5 test one mechanism underlying beliefs in others’ potential for improvement. Specifically, these experiments test whether evaluations of warmth traits are less sensitive to gradations in warmth-related behaviors than evaluations of competence are sensitive to gradations in competence-related behaviors. In other words, people are better able to distinguish between different degrees of competence than to distinguish between different degrees of warmth. In turn, this relative inability to finely evaluate warmth-related behaviors contributes to people’s weaker beliefs that warmth traits can be improved.

In all experiments, the results are presented in terms of effect sizes and interpreted in two ways. First, a brief qualitative assessment of the size of each effect is presented according to Cohen (1988), wherein effect sizes are divided into three categories: small, medium, and large. Second, and more important, I highlight comparisons of effect sizes for the experimental conditions most relevant to the hypotheses of each experiment.

**Experiment 1a: The Possibility of Improvement**

Experiment 1a asked participants to judge whether people can improve warmth and competence traits. Experiment 1a also tested whether people believe that warmth or competence traits are more important for success in life. I predicted that although people would believe that warmth and competence traits are similarly important for success in life, people would also believe that it is less possible for others to improve warmth traits compared to competence traits.

**Method**

Participants (N = 58; 64% women; \( M_{age} = 21.55, SD_{age} = 4.13 \)) were undergraduate students recruited in a research lab in a Midwestern university for nominal payment. I used a 2 (trait type: warmth vs. competence) × 2 (judgment type: improvement vs. importance) mixed design. In the improvement conditions, participants read that the purpose of this research was to understand aspects of life that can be improved. Participants were then asked whether certain traits can be improved over one’s lifetime on a scale ranging from 0 (almost impossible) to 100 (entirely possible). The instructions emphasized that the judgments pertained to sustained improvements over time, rather than temporary and fleeting improvements.

In the importance conditions, participants read that the purpose of this research was to understand what aspects of life are important for success. Participants were then asked whether, in thinking about the various situations a person encounters in life, certain traits are important to have. Responses were made on a scale ranging from 0 (not at all) to 100 (very important).

All participants rated five warmth and five competence traits. Following Fiske et al. (2002), Study 1, the five warmth traits were warmth toward other people, sincerity, being a good-natured person, tolerance, and being a caring person, and the five competence traits were being a competent person, confidence, intelligence, being a competitive person, and independence. Presentation of traits was randomized within each trait type.

**Results**

I created a warmth index by averaging the five warmth traits (\( \alpha = .84 \)) and a competence index by averaging the five competence traits (\( \alpha = .69 \)). An analysis of variance (ANOVA) on trait type and judgment type with repeated measures on the first factor revealed medium effect for trait type, \( \eta^2_p = .06 \); a small-to-medium effect for judgment type, \( \eta^2_p = .04 \); and a medium-to-large effect size for the interaction, \( \eta^2_p = .11 \). As Figure 1 shows, warmth traits were rated as equally important for success in life (\( M = 72.40, SD = 17.44 \)) as competence traits (\( M = 70.73, SD = 14.64 \)), evidenced by the small effect size of this comparison (\( d = .08 \)). However, improving warmth
traits was rated as less possible ($M = 60.57, SD = 20.56$) than improving competence traits ($M = 70.86, SD = 12.05$) with a large effect size ($d = .85$).

Experiment 1a finds that people believe that improvement in warmth traits is less possible than in competence traits but that the two types of traits are equally important for life success.

Experiment 1b: More Diverse Sample

The primarily undergraduate sample used in Experiment 1a raises the question of whether young persons have a full grasp of the intricacies and complexities that characterize a lifetime. Because young persons have experienced only a minority of the average lifespan, their beliefs of whether traits can or cannot be improved may differ from older persons’ beliefs. Experiment 1b therefore provided a near-exact replication of Experiment 1a while recruiting a sample with a broader range of ages.

Method

Participants ($N = 114$; 32.5% women; $M_{age} = 32.78$, $SD_{age} = 10.99$; age range = 18–72 years) were recruited from Amazon.com’s Mechanical Turk (M-Turk) for nominal payment. I used the same design as in Experiment 1a: 2 (trait type: warmth vs. competence) × 2 (judgment type: improvement vs. importance) mixed design. The procedure and dependent measures were also similar to Experiment 1a, except that here participants evaluated six warmth and six competence traits that were adapted from Fiske et al.’s (2002) Study 2. The six warmth traits were being friendly, trustworthy, warm, sincere, good-natured, and caring. The six competence traits were being skillful, efficient, competent, confident, capable, and independent. Presentation of traits was randomized within each trait type.

Results

I created a warmth index by averaging the six warmth traits ($z = .90$) and a competence index by averaging the six competence traits ($z = .80$). An ANOVA on trait type and judgment type with repeated measures on the first factor revealed a medium effect for trait type, $\eta^2_p = .069$; a medium-to-large effect for judgment type, $\eta^2_p = .096$; and an interaction with a small-to-medium effect size, $\eta^2_p = .032$. Warmth traits were again rated as equally important for success in life ($M = 76.61$, $SD = 16.35$) as competence traits ($M = 78.25$, $SD = 11.46$), evidenced by the small effect size of this comparison ($d = .09$). However, improving warmth traits was again rated as less possible ($M = 64.91$, $SD = 19.78$) than improving competence traits ($M = 73.33$, $SD = 14.89$) with a medium effect size ($d = .45$). Including age as a covariate in the ANOVA does not meaningfully alter these results, suggesting that experiencing more of life does not seem to affect beliefs in others’ potential for improvement over the lifetime.

Experiment 1c: Directly Ranking Warmth and Competence Traits

The previous experiments asked participants to separately rate each trait, thereby relying on the assumption that rating scales provide evenly spaced intervals and meaningful end-points. This may not be the case in judgments of varied personality traits. The conceptual intervals and end-points for “confidence,” for example, may not be comparable to “trustworthiness.” To address this concern, Experiment 1c asks participants to directly trade off warmth and competence traits by ranking them based on potential for improvement. This method elicits judgments without relying on rating scales.

Method

Participants ($N = 204$; 58.9% women; $M_{age} = 38.36$, $SD_{age} = 13.65$) were recruited from Amazon.com’s M-Turk for nominal payment. I used a 2 (trait type: warmth vs. competence) within-subjects design. Participants read that they would see a list of personality traits and be asked to think about whether it was possible to improve these traits over a person’s lifetime. Participants were then shown a list of six warmth and six competence traits (the same ones used in Experiment 1b) and asked to rank them from the least possible trait
to improve to the most possible trait to improve. The 12 traits were presented on the screen in random order.

Results

Note that higher numbers represent lower rankings (traits that are less possible to improve). I averaged the rankings of the six warmth traits and the rankings of six competence traits. Warmth traits were ranked lower on average ($M = 7.12$, $SD = 1.85$) than competence traits ($M = 5.88$, $SD = 1.85$), suggesting that participants considered warmth traits to be less possible to improve than competence traits. The effect size of this difference was $d = .36$, a small-to-medium effect.

Reviewing each trait separately, all of the warmth traits were ranked lower than five of the six competence traits. The one exception was “intelligence,” which was ranked lower ($M_{intelligence} = 8.00$, $SD_{intelligence} = 3.92$) than all of the six warmth traits. This result was in contrast to Experiment 1a, in which intelligence was not an outlier. Overall, direct rankings strongly tended to replicate rating scales. Warmth traits were considered less possible to improve than competence traits when participants directly traded off the two types of traits.

Experiment 2: The Difficulty of Improvement

Beliefs about others’ potential for improvement may affect beliefs about the probability that others can achieve valued goals. Experiment 2 tested this possibility by asking participants to evaluate the likelihood that a person who was denied promotion to a management position would be able to improve shortcomings related to either warmth or competence before reapplying. Experiment 2 also attempts to replicate previous findings from existing research suggesting that warmth traits are viewed as more essential to personality than competence traits to better understand whether these “essentialist” beliefs are related to beliefs about others’ potential for improvement.

Method

Participants ($N = 76$; 46% women with two participants failing to note their gender; $M_{age} = 32.00$, $SD_{age} = 11.19$) from Amazon.com’s MTurk participated for nominal payment. I used a 2 (shortcoming: warmth vs. competence) between-subjects design. Participants read about “Marti,” who has been working in the same company for 5 years before deciding to apply for an internal promotion to a management position. Participants read that after a lengthy application process, Marti was told that he would not be promoted and was given feedback that either centered on his strengths in competence and weaknesses in warmth, or vice versa. Specifically, the feedback was (with the experimentally manipulated text is in square brackets):

We were really impressed with your competence [interpersonal qualities] for the job. You had a good level of skills and independence [warmth and caring attitude]. However, your interpersonal qualities [competence] were [was] lacking compared to our requirements. You did not have the warmth, caring attitude, and kindness [skills and independence] that we were looking for and that are necessary for the job.

Participants then read that Marti was told he could reapply for the same position within a year and that his application would be assessed based on his ability to improve upon his shortcomings. Participants were then asked how difficult it would be for Marti to improve upon his shortcomings ($4 = extremely difficult$) and how likely Marti was to succeed in his reapplication ($4 = extremely likely$). Finally, to assess whether improvement judgments are related to essentialism, participants were asked whether succeeding in the application required Marti to have to change who he was and his personality, on a scale ranging from $1$ (he would not have to change any part of who he is) to $9$ (he would have to change major parts of his personality).

Results

Beliefs about improvement. As Table 1 shows, participants believed that it would be more difficult for Marti to improve his shortcomings when he was lacking warmth ($M = 5.80$, $SD = 1.80$) than when he was lacking competence ($M = 4.58$, $SD = 1.93$; $d = .66$, medium-to-large effect). Participants also believed that Marti is less likely to be successful in his reapplication when he was lacking warmth ($M = 4.3$, $SD = 1.97$) than when he was lacking competence ($M = 1.36$, $SD = 1.73$; $d = .51$, medium effect). Thus, participants believed that overcoming a shortcoming related to warmth was more

<table>
<thead>
<tr>
<th>Shortcoming</th>
<th>Difficulty in Improving Specific Shortcomings</th>
<th>Likelihood of Succeeding in Reapplication</th>
<th>Degree of Necessary Personality Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lacking warmth</td>
<td>5.80 (1.80)</td>
<td>0.43 (1.97)</td>
<td>7.05 (1.68)</td>
</tr>
<tr>
<td>Lacking competence</td>
<td>4.58 (1.93)</td>
<td>1.36 (1.73)</td>
<td>4.22 (2.11)</td>
</tr>
</tbody>
</table>

Note. Standard deviations are in parentheses.
difficult than overcoming a shortcoming related to competence.

**Personality change.** Participants believed that success would require Marti to have to change more major parts of his personality when he was lacking warmth ($M = 7.05$, $SD = 1.68$) compared to when he was lacking competence ($M = 4.22$, $SD = 2.11$; $d = 1.51$, large effect). This result is consistent with previous research findings that people believe that interpersonal and moral traits are more central to one’s personality than other types of traits (e.g., Strohminger & Nichols, 2014) and suggests that the constrained beliefs in others’ potential for improvement in warmth traits may be multiply determined. The mechanism investigated in the present research (warmth traits may be more difficult to evaluate than competence traits) may not be the only one determining beliefs about others’ potential. I discuss other potential mechanisms further in the General Discussion.

In summary, Experiment 2 finds that participants believe that improving a deficit in warmth traits is more difficult than improving a deficit in competence traits. This result is a consequence of the asymmetry in beliefs about others’ potential for improvement: Shortcomings related to warmth traits were considered harder to overcome than shortcomings related to competence traits.

**Experiment 3: Evaluability of Warmth and Competence Behaviors**

One possible mechanism underlying the asymmetric beliefs in others’ potential for improvement relates to differences in how behaviors that exhibit warmth and competence are evaluated. In general, people rely on others’ behaviors to infer others’ traits (Epley & Dunning, 2000; Gilbert & Malone, 1995; for review, see Pronin, 2009). However, the translation from observed behaviors to inferred traits may not be straightforward. In particular, people may be more sensitive to gradations in another person’s competence-related behavior than they are sensitive to gradations in a warmth-related behavior. To test this possibility, Experiment 3 exposes participants to one of two behaviors—either a competence-related behavior (doing well on a math exam) or a warmth-related behavior (giving money to another person). To create a common unit for comparison, both of these behaviors were presented as measurable by dollars earned in the math exam or dollars given to another person. Each participant was exposed to one magnitude of one of these behaviors—either giving a certain amount of money to a stranger or earning a certain amount of money by succeeding in a math exam. After reading about one of these behaviors, participants were asked to evaluate the actor’s traits. If warmth traits are less evaluable than competence traits, then participants’ evaluations would be less sensitive to gradations in the amounts of money given to a stranger than to the amounts of money earned doing a math exam.

**Method**

**Pretest.** I first conducted a pretest to validate whether the behaviors used in the main experiment indeed reflect warmth and competence traits. Participants ($N = 61$; $49.2\%$ women; $M_{age} = 32.5$, $SD_{age} = 11.51$) from Amazon.com’s M-Turk were recruited for nominal payment. Participants read about two behaviors in counterbalanced order. The first behavior was coming to a research lab to do an experiment in which an actor would be given $6 and would have the opportunity to give to a stranger however much of this money he or she desired. The second behavior was coming to a research lab to do an experiment in which an actor would be given a math exam with six questions and be awarded $1 for every correct answer.

Participants were asked to judge whether 10 traits are relevant to determining the outcome of each of these two situations. Five of the 10 traits were warmth traits (tolerant, warm, good-natured, sincere, caring), and the other five traits were competence traits (competent, confident, independent, competitive, intelligent). Presentation of the 10 traits was randomized within each trait type. Responses were given on scales ranging from −3 (not relevant at all) to 3 (highly relevant), with the midpoint of 0 (neither relevant nor irrelevant).

**Main experiment.** Participants ($N = 308$; $40.6\%$ women; $M_{age} = 28.70$, $SD_{age} = 8.99$) from Amazon.com’s M-Turk participated for nominal payment. I used a 2 (behavior: giving money vs. doing exam) × 7 (degree of behavior) fully between-subjects design. Participants read that “Larry” came to our lab to do an experiment. In the giving money condition, Larry was provided with $6 and had the opportunity to give to a stranger however much of this money he decided. In the doing exam condition, Larry took a math exam that included six questions. For every correct answer, Larry earned $1. Therefore, although in both conditions the total amount possible for Larry to walk away with was $6, the giving money condition involved a warmth-related behavior (giving money to another person) and the doing exam condition involved a competence-related behavior (doing well on an exam).

Participants then learned the amount of money Larry gave or earned. This amount was manipulated fully between-subjects to be $0, $1, $2, $3, $4, $5, or $6, thereby including the entire possible range using integer
intervals. Finally, all participants evaluated Larry’s warmth on five traits (tolerant, warm, good-natured, sincere, caring) and Larry’s competence on five traits (competent, confident, independent, competitive, intelligent) on scales ranging from 1 (not at all) to 7 (very much). These traits were the same as those used in the pretest.

Results

Pretest. I averaged the five warmth traits ($x = .80$) and the five competence traits ($x = .71$). Participants believed that competence traits were more relevant to succeeding in the math exam ($M = .79, SD = .93$), both compared to the 0 point and compared to the relevance of warmth traits to succeeding the math exam ($M = -1.93, SD = 1.26; d > 1.71$, large effects). Participants also believed that warmth traits were more relevant to the decision of giving money to a stranger ($M = 1.31, SD = 1.21$), both compared to the 0 point and compared to the relevance of competence traits ($M = -4.1, SD = 1.24; d > 2.18$, large effects). Thus, the pretest validated that doing an exam was considered more relevant to competence and giving money to another person was considered more relevant to warmth.

Main experiment. I averaged the five warmth ($x = .94$) and the five competence traits ($x = .82$). An ANOVA of behavior and degrees of behavior on warmth evaluations revealed main effects for behavior and degrees of behavior, $\eta_p^2 s > .08$, and the predicted interaction, $\eta_p^2 = .33$. Similarly, an ANOVA of behavior and degrees of behavior on competence evaluations revealed main effects for behavior and degrees of behavior, $\eta_p^2 s > .10$, and the predicted interaction, $\eta_p^2 = .39$. Finally, I conducted an ANOVA on only the “relevant” evaluations—evaluations of competence in the math test conditions and evaluations of warmth in the giving conditions. This ANOVA revealed medium-to-large main effects for behavior type and degrees of behavior, $\eta_p^2 s > .11$, and a medium effect for the interaction, $\eta_p^2 = .08$.

To better understand these interactions, I analyzed the giving money and earning money conditions separately. Because this experiment tests sensitivity to gradations in warmth and competence behaviors, the appropriate statistical analysis involves a linear regression of warmth and competence evaluations on the different degrees of behaviors. The slope of evaluations as a function of amount of money given or earned represents people’s sensitivity to gradations in warmth- and competence-related behaviors. Moreover, because existing research (Klein & Epley, 2014; Klein et al., 2015) finds that warmth evaluations exhibit diminishing sensitivity to gradations—greater magnitudes of warmth-related behaviors result in smaller and smaller changes in evaluations—I also added a quadratic term to measure whether participants’ evaluations exhibit diminishing sensitivity to gradations. Based on previous research, the key prediction is that this quadratic term will be sizable and negative in evaluations of warmth but negligibly small in evaluations of competence.

The effect of giving money on evaluations. I regressed warmth evaluations on the amount of money given and on a quadratic term (amount of money given squared). If participants become less sensitive to gradations for increasingly large giving amounts, then the quadratic term should be negative. Adding the quadratic term increased the proportion of variance explained by the regression model (measured by $R^2$) from 62% to 68%.

As Table 2 shows, the amount of money given to a stranger had a linear effect on warmth evaluations ($\beta = 1.34$), suggesting that warmth evaluations increased as more money was given. Importantly, the quadratic term was negative ($\beta = -.11$), indicating that increasingly larger giving amounts had increasingly smaller effects on warmth evaluations. Participants provided less sensitive evaluations of warmth in response to increasingly larger giving amounts.

I also regressed competence evaluations on giving amounts. Giving amounts had a very small linear effect on evaluations of competence ($\beta = -.03$) and a very small quadratic effect ($\beta^2 = -.01$). This regression model explained only 5.0% of the variance in competence evaluations, suggesting that giving money to another person was not seen as a sign of (in)competence.

The effect of earning money on evaluations. The amount of money earned in the math test had a large linear effect on competence evaluations ($\beta = .60$). In
contrast to the giving money conditions, however, the quadratic coefficient was negligible ($\beta^2 = -.01$). Moreover, adding the quadratic terms did not explain additional variance in the model, as 66.5% of the variance was explained without this term and 66.6% of the variance was explained with the quadratic term included. These results suggest that unlike warmth evaluations, participants did not exhibit diminishing sensitivity to gradations in competence for increasingly larger amounts of money earned; low and high amounts of money earned resulted in similar changes to competence evaluations.

I also regressed warmth evaluations on earned amounts in the math exam. Success in the exam had a small linear effect on warmth evaluations ($\beta = .18$) and a very small quadratic effect ($\beta^2 = -.01$). This regression model explained only 6.6% of the variance in warmth evaluations, suggesting that success in an exam was not seen as a sign of warmth.

In summary, Experiment 3 reveals an asymmetry in people’s sensitivity to gradations in behavior—evaluations were more sensitive to gradations in a behavior that exhibited competence than to a behavior that exhibited warmth. The pattern of the results is consistent with existing research (Klein & Epley, 2014) in that warmth evaluations exhibit diminishing sensitivity as warmth-related behaviors become increasingly large. In contrast, competence evaluations exhibit similar sensitivity to magnitude across low and high magnitudes of competence-related behaviors.

Experiments 4 and 5 test whether this asymmetric sensitivity to gradations partly explains the relatively weak beliefs in others’ potential for improvement in warmth traits found in Experiments 1a–2.

**Experiment 4: Evaluability Affects Beliefs About Improvement**

If people are relatively insensitive to varying degrees of warmth, this may partly explain why people believe that others’ warmth traits are relatively hard to improve. Improvement means moving from one level in a given attribute to another, higher level. Therefore, strong beliefs that improvement is possible in an attribute may partly depend on being able to easily identify gradations in that attribute. Experiment 4 tested this hypothesis by increasing the evaluability of both warmth and competence traits. If insensitivity to gradations affects beliefs about improvement, then enabling people to better differentiate between varying degrees of warmth should also increase their beliefs that improvement in warmth is possible. On the other hand, if competence traits can more easily be evaluated than warmth traits, then increasing evaluability should have a smaller effect on beliefs in the possibility of improving competence traits.

Existing research finds that more exposure to a stimulus increases the evaluability of that stimulus, as people become more familiar with it and increasingly able to compare between different gradations of it (Hsee et al., 1999; Morewedge, Kassam, Hsee, & Caruso, 2009). Experiment 4 relied on these findings and utilized a common way to facilitate comparisons between varying degrees of a given stimulus—asking participants to provide either single or joint evaluations (Hsee & Zhang, 2010). Participants either evaluated only one possible action that exhibited either warmth or competence (single evaluation) or evaluated multiple possible actions that exhibited either warmth or competence (joint evaluation). Evaluating multiple possible actions in a given domain encourages comparisons between different magnitudes of these actions and therefore should sensitize evaluations to varying gradations of these actions. If sensitivity to gradations affects beliefs about improvement, then evaluating multiple warmth-related actions should increase beliefs in the potential for improving warmth traits. In contrast, because people more easily identify gradations in competence, facilitating comparisons between varying degrees of a competence-related behavior should have a smaller effect on beliefs in the possibility of improving competence traits.

**Method**

Participants ($N = 244$; 43.9% women; $M_{age} = 31.71, SD_{age} = 10.09$) from Amazon.com’s M-Turk online panel took part for nominal payment. I used a $2 \times 2$ (behavior: giving money vs. doing exam) × 2 (evaluability: low vs. high) between-subjects design. Similar to Experiment 3, this experiment manipulated whether “Larry” gave money to a stranger or earned money by correctly answering math questions. Unlike Experiment 3, this experiment manipulated the evaluability of Larry’s actions. Participants in the low evaluability condition evaluated Larry based on one action in either the giving money or the math exam scenario: either the most generous or competent action (giving $6 or earning $6 out of the possible $6), the most selfish or incompetent action (giving $0 or earning $0 out of $6), or a moderate action (giving $3 or earning $3 of the $6). Participants evaluated only the relevant traits, either warmth when giving money to another person or competence when earning money by correctly answering exam questions.

Participants in the high evaluability conditions evaluated Larry on all of these three possible actions in either the giving money or the math exam scenario (giving or earning $0, $3, and $6). Presentation of the actions was made in random order. Thus, every participant in the high evaluability conditions evaluated three rather than one action, thereby receiving more exposure to, and directly comparing between, different actions in warmth and competence domains.
Competence evaluations were made only in the doing exam conditions on three traits (competent, confident, intelligent), and warmth evaluations were made only in the giving money conditions on three traits (trustworthy, warm, caring). Each warmth or competence trait was evaluated on a scale ranging from 1 (not at all) to 7 (very much).

After evaluating Larry’s action(s), participants provided their beliefs about whether Larry can improve upon either the three warmth traits in the giving money conditions or the three competence traits in the doing exam conditions. The scale for these measures ranged from 0 (virtually impossible to improve) to 100 (entirely possible to improve).

**Results**

**Evaluations in the low evaluability conditions.** I averaged the three warmth traits ($\alpha = .99$) and the three competence traits ($\alpha = .99$). Similar to the results of Experiment 3, warmth evaluations in the low evaluability conditions exhibited diminishing sensitivity to gradations in the amount of money given to another person. As Figure 2 shows, although participants evaluated giving $3$ as warmer ($M = 6.06, SD = .87$) than giving $0$ ($M = 2.82, SD = 1.48; d = 2.69, large effect), they did not evaluate giving $6$ ($M = 6.01, SD = .86$) as any warmer than giving $3$ ($d = .06$, negligibly small effect).

In contrast, competence evaluations were more sensitive to degrees of success in the math exam. Participants evaluated earning $3$ as more competent ($M = 4.46, SD = .85$) than earning $0$ ($M = 2.25, SD = .89; d = 2.59, large effect) and evaluated earning $6$ as more competent still than earning $3$ ($d = 1.67, large effect). Consistent with predictions and with existing research, evaluations of the warmth-related behavior in the low evaluability conditions were relatively less sensitive to gradations in that behavior than evaluations of competence were sensitive to gradations in the competence-related behavior.

**Evaluations in high evaluability conditions.** In contrast, evaluations of multiple actions were relatively sensitive to gradations in both the giving-money and earning-money scenarios. Participants evaluated giving $3$ as more warm ($M = 5.53, SD = .82$) than giving $0$ ($M = 2.72, SD = 1.14; d = 2.59$) and evaluated giving $6$ as warmer still than giving $3$ ($M = 6.49, SD = .55; d = .94,$ both large effects). Likewise, participants evaluated earning $3$ in a math exam as more competent ($M = 4.24, SD = .80$) than earning $0$ ($M = 2.59, SD = 1.25; d = 1.93$), and they evaluated earning $6$ as more competent still than earning $3$ ($M = 6.01, SD = .84; d = 1.70,$ both large effects).

In sum, when participants were able to directly compare between multiple outcomes, they became more sensitive to gradations. This effect was more pronounced in warmth-related behaviors, in which noticing gradations appears to be naturally more difficult for participants. These results are consistent with past research showing that facilitating comparisons increases the evaluability of otherwise difficult-to-evaluate stimuli (Hsee et al., 1999; Klein & Epley, 2014).

**Beliefs about improvement.** I averaged beliefs about the possibility of improvement in the three warmth traits ($\alpha = .94$) and beliefs about the possibility of improvement in the three competence traits ($\alpha = .77$). An ANOVA of beliefs about improvement on evaluability and trait type revealed a medium-to-large effect for evaluability, $\eta_p^2 = .037$; a medium effect for trait type, $\eta_p^2 = .020$; and a small effect for the interaction, $\eta_p^2 = .004$.

As Figure 3 shows, and replicating the asymmetry in beliefs in improvement observed in previous experiments, participants in the low evaluability conditions believed that competence traits are more possible to
improve \((M = 71.29, \ SD = 21.44)\) than warmth traits \((M = 60.20, \ SD = 29.18; \ d = .44, \) medium effect\). Consistent with the hypotheses, in the high evaluability condition, the warmth-competence asymmetry was reduced. Participants in the high evaluability conditions believed that competence traits are still directionally more possible to improve \((M = 78.04, \ SD = 14.05)\) than warmth traits \((M = 74.11, \ SD = 13.36)\) with a small effect \((d = .29)\). The difference between these effect sizes \((d_{\text{low-evaluability}} - d_{\text{high-evaluability}} = .15)\) is equal to one third of the effect associated with the warmth-competence asymmetry in the low evaluability conditions, suggesting that evaluability was responsible for some, but not all, of the constrained beliefs in others’ potential for improvement in warmth traits.

In summary, Experiment 4 replicates the asymmetry in beliefs about improvement; people across conditions believed that improving competence is more possible than improving warmth. In addition, Experiment 4 also finds that enabling participants to more easily identify gradations in a warmth-related behavior reduced the warmth-competence asymmetry in beliefs about improvement. It appears that being able to discern the “next level” that would constitute improvement compared to the present state can make beliefs about others’ improvement in warmth and competence traits more symmetric.

Alternative interpretations. Two alternative interpretations to the results of Experiments 3 and 4 deserve mention. First, one alternative interpretation would suggest that participants’ nonlinear evaluations of warmth behaviors may reflect a preference for fairness, whereby participants downgrade an actor who gives more than half of an amount of money because they prefer a fair outcome to an unfair outcome regardless of whether the actor violates fairness by being generous or selfish. However, this alternative explanation would not predict the results of Experiment 4, wherein asking participants to evaluate several possible actions created a difference between evaluations of a fair action and a generous action. Moreover, preferences for fairness cannot explain the change in beliefs about the potential for improving warmth traits that resulted from exposing participants to multiple actions in Experiment 4.

A second alternative interpretation would suggest that participants’ nonlinear evaluations of warmth behaviors in Experiment 3 are vulnerable to ceiling effects because of the bounded nature of the scale used in that experiment. However, ceiling effects cannot explain the results of Experiment 4, wherein exposing participants to multiple actions created a difference between evaluations of the moderate action (giving $3 out of $6 to another person) and the more generous action (giving $6 out of $6). If my metrics can detect differences in evaluations of multiple actions, then they are also theoretically capable of detecting differences in evaluations of single actions.

Experiment 5: Providing Range Information to Increase Beliefs in Improvement in Warmth

Experiment 5 provides another test of the evaluability mechanism. Existing research suggests that providing people with information on the highest and lowest possible values of an attribute can enable them to more easily place a given value within this range (Hsee, 1996; Hsee et al., 1999). For example, judging whether a given TOFEL exam score is high or low is easier when the range of possible scores is known than when this range is unknown. Experiment 5 provides participants with the range of donations (warmth-related behavior) or rewards from a performance task (competence-related behavior) and measures how this range information affects beliefs about the potential for improvement in warmth and competence traits. If evaluability is connected to beliefs about improvement, range information should increase beliefs about improvement in warmth traits more than in competence traits.

Method

Participants \((N = 209; \ 42.2\% \) women; \(M_{\text{age}} = 30.78, \ SD_{\text{age}} = 9.46\)) from Amazon.com’s M-Turk online panel took part for nominal payment. I used a 2 (trait type: warmth vs. competence) \(\times\) 2 (behavior type: warmth-related vs. competence-related) \(\times\) 2 (range information: given vs.
not given) mixed design. Participants read about a person who participated in a psychology study. To avoid gender-related confounds, the gender of this person was not mentioned. In the warmth-related behavior condition, participants read that the study gave this person the opportunity to donate money to charity and that this person decided to donate $20. In the competence-related behavior condition, participants read that the study provided this person with financial rewards for successfully performing a task that required efficiency, competence, and skill. Participants in that condition read that this person earned $20 in the study.

In the range conditions, participants also learned that $40 was the highest amount and $0 was the lowest amount donated or earned in the hypothetical study. Thus, the focal person donated or earned precisely half of the maximum amount. In the no range conditions, this range information was not provided.

All participants were finally asked to evaluate how possible it was for the focal person to improve three warmth traits (warmth, caring for others, kindness) and three competence traits (efficiency, competence, skill) on scales ranging from 0 (almost impossible) to 100 (entirely possible).

Results

I averaged judgments of the possibility of improving warmth traits ($z = .96$) and competence traits ($z = .92$). An ANOVA of trait type, behavior type, and range information with repeated measures on the first factor revealed a medium main effect for trait type, $\eta^2_p = .04$, whereby competence traits were generally judged to be more possible to improve than warmth traits (see Figure 4). The ANOVA also revealed an interaction between behavior type and range information, $\eta^2_p = .03$ (small-to-medium effect). All other effect sizes were relatively small, $\eta^2_p s < .011$.

Beliefs about improvement in warmth traits. When the focal person donated money, having range information increased participants’ beliefs that improvement in the warmth traits was possible ($M = 75.86$, $SD = 24.85$) compared to not having range information ($M = 63.78$, $SD = 22.45$) with a medium effect size ($d = .51$). Knowing the range of donations increased beliefs in the potential for improvement in warmth traits.

When the focal person performed a skill task, having range information had a much smaller (and negative) effect on participants’ beliefs that improvement in the warmth traits was possible ($M = 69.87$, $SD = 22.01$) compared to not having range information ($M = 72.52$, $SD = 15.48$; $d = -.14$). In a skill task less relevant for warmth traits, knowing the range of performance had a negligible effect on beliefs in the potential for improvement in warmth traits.

Beliefs about improvement in competence traits. When the focal person donated money, having range information increased beliefs that improvement in the competence traits was possible ($M = 75.80$, $SD = 20.33$) compared to not having range information ($M = 68.21$, $SD = 19.86$). This difference was associated with a small-to-medium effect size ($d = .37$). This effect size is roughly two thirds of that found for warmth traits ($d_{warmth} = .51$), suggesting that range information had a smaller effect on beliefs in improvement in competence traits compared to warmth traits. This result suggests that evaluability is a more important factor in influencing beliefs about the possibility of improving warmth traits than the possibility of improving competence traits. As can be seen in Figure 4, the smaller effect of range information on competence evaluations in the donation task can be attributed to the stronger baseline belief that competence can be improved compared to warmth.

When the focal person performed a skill task, having range information had a small (negative) effect on beliefs that improvement in competence traits was possible ($M = 76.73$, $SD = 17.55$) compared to not having range information ($M = 78.03$, $SD = 16.06$; $d = -.08$). Knowing the range of performance in a skill task had a negligible effect on beliefs in the potential for improvement in competence traits.

![Figure 4](image-url)

**FIGURE 4** Beliefs about improvement in warmth and competence traits as a function of having information on the possible range of values in warmth- and competence-related tasks in Experiment 5. *Note.* Error bars represent standard errors.
These results suggest that having range information reduced the asymmetry in beliefs about improvement in warmth in the donation scenario compared to competence in the skill task scenario. Without range information the difference in beliefs about improvement in warmth versus competence was $d = -0.22$, whereas with range information this difference was far smaller, at $d = -0.04$.

In summary, Experiment 5 increases the evaluability of warmth-related behaviors by providing the highest and lowest outcomes associated with these behaviors. Knowing this range increased beliefs in the potential for improvement in warmth traits more so than it increased beliefs in the potential for improvement in competence traits, thereby reducing the warmth-competence asymmetry in beliefs about improvement. This result suggests a role for evaluability in affecting beliefs about others’ potential for improvement.

**GENERAL DISCUSSION**

Believing that others can improve has a role in facilitat- ing professional and social success, but the present research finds that beliefs in others’ potential for improvement are asymmetric. A series of seven experiments finds that people believe that others are less capable of improving warmth traits than improving competence traits. Experiments 3–5 find that the relatively weak beliefs in the potential for improvement in warmth are partly explained by relative insensitivity to gradations in warmth. In general, people’s judgments depend on the ability to differentiate between low and high degrees of an attribute (Hsee et al., 1999; Hsee & Zhang, 2010; Morewedge et al., 2009; Parducci, 1965). Experiment 3 reveals that increasingly greater degrees of generosity led to progressively undifferentiated evaluations of warmth. Experiments 4 and 5 manipulated the evaluability of warmth- and competence-related behaviors and revealed that enabling participants to detect finer gradations of warmth also increased beliefs in others’ potential for improving warmth traits, thereby reducing the asymmetry in beliefs about improvement across warmth and competence traits. Increasing the evaluability of competence-related behaviors had a smaller effect on beliefs in the potential for improving competence traits.

**Barriers to Believing That Others Can Improve**

Understanding how to increase beliefs in others’ potential for improving warmth is a productive avenue for future research. One way to do so may be to frame warmth traits as interpersonal “skills.” If people think of traits such as kindness, generosity, and warmth as a set of competencies, the possibility of improving them may seem more plausible.

Another way to increase beliefs in the capacity of others to improve may be interventions that increase the evaluability of warmth traits. One way to do so may be to enable social comparisons in warmth-related behaviors, which in turn can increase the evaluability of warmth traits and therefore increase the belief that others can improve these traits. There are various ways to implement this idea in practice. For example, some companies and organizations allow employees to behave generously toward one another by giving peer-to-peer-bonuses. If these organizations also enabled employees to view one another’s bonuses, this could promote stronger beliefs in the possibility of improvement in warmth because social comparisons could more readily be made.

**Additional Mechanisms and Extensions**

Many psychological phenomena are multiply determined, and beliefs about others’ possibility for improve- ment are unlikely to be an exception. In addition to the evaluability mechanism explored here, at least three other mechanisms seem possible. The first was men- tioned earlier and relates to beliefs about the “essential” elements of personality (Strohminger & Nichols, 2014). If warmth traits are considered more essential to one’s personality, then improving them may imply greater change than improving competence traits.

Second, some behaviors that exhibit warmth toward others tend to involve “losses” to the actor. For example, giving away money or spending time with others requires transferring one’s resources to others. In contrast, competence-related behaviors often involve gaining resources, such as money, a job promotion, or recognition. This gain versus loss difference may be a productive avenue for future research, because beliefs about improvement may be connected to well-documented differences in judgments of gains and losses (i.e., Kahneman & Tversky, 1979).

Third, when people form impressions of others, warmth traits tend to be processed faster than com- petence traits (Abele & Bruckmüller, 2011). This faster and preferential processing of warmth traits may under- lie other differences in processing warmth and com- petence traits that interact with beliefs in the potential for improvement in those traits.

Just as additional mechanisms can underlie the effects reported here, extensions of the present findings are possible as well. One such extension involves investigat- ing the possible effects of evaluability on beliefs about decline—rather than improvement—in warmth and competence traits. One doubt about this prediction is that insensitivity to gradations in warmth tends to occur.
in relatively high levels of warmth-related behaviors. Generous behavior is difficult to differentiate, whereas relatively selfish behavior is easier to differentiate (Klein & Epley, 2014; Klein et al., 2015). This suggests that improvement judgments are more likely to be affected by evaluability compared to decline judgments. Nevertheless, understanding judgments of decline remains a potential avenue for future research.

The Potential for Improvement in Others: Beliefs Versus Reality

People’s relatively pessimistic beliefs about others’ potential for improvement in warmth traits raise the question of whether these beliefs conform to reality. Are warmth traits indeed systematically more difficult to improve than competence traits, as participants in the present experiments believe? Although the answer will remain unclear without experiments that test this question directly, two sets of findings appear to cast doubt on any actual differences in the potential for improving warmth and competence.

First, meta-analytic research finds consistent positive effects of psychotherapy on personal and social success, even in nonclinical populations (Lipsey & Wilson, 1993). Mental health depends in part on social elements such as a supportive social network and a feeling of belonging. Improvement in warmth traits is presumably required to develop these social support systems and thereby achieve better mental health. Therefore, the efficacy of psychotherapy suggests that improvement in warmth traits may be possible in reality, albeit with the help of trained therapists.

Second, existing research finds that people strongly believe that their personality has changed considerably during the period from the past to the present (Quoidbach, Gilbert, & Wilson, 2013). If people believe that change—and with it, improvement—was possible for them in the past, then people should also, at least rationally, believe that the same is possible for others in the future.

ACKNOWLEDGMENTS

I thank the Epley lab group for providing feedback on this research and the Basic and Applied Social Psychology editorial staff and three anonymous reviewers for their service and feedback on this article.

REFERENCES


