EMPATHIC ACCURACY AND THE USE OF STEREOTYPES IN INFERRING THE THOUGHTS AND FEELINGS OF OTHERS

By
BLAKE WAYSON LOCHER

A THESIS

Presented to the Department of Psychology
And the Honors College of the University of Oregon
In partial fulfillment of the requirements
For the degree of
Bachelor of Science

June 2009

An Abstract of the Thesis of

Blake Locher	for the degree of	Bachelor of Science
In the Department of Psychology	to be taken	June 13, 2009

Title: EMPATHIC ACCURACY AND THE USE OF STEREOTYPES IN INFERRING THE THOUGHTS AND FEELINGS OF OTHERS

Approved: [redacted]
Professor Sara D. Hodges

Past research has suggested that perceivers construct a schema of a target person in order to accurately infer that person's thoughts and feelings (Myers & Hodges, 2009). The current study was designed to test whether one result of having such schemas is greater accuracy in inferring thoughts and feelings that are stereotypic of a salient target group. Participants (145 undergraduates) inferred the thoughts and feelings of a woman discussing her experiences as a new mother. We hypothesized that perceivers would have greater empathic accuracy for thoughts and feelings that were more stereotypical of the salient target group (e.g., new mothers). Results supported our hypothesis: even when controlling for inferential difficulty, thoughts that were more stereotypical were easier to infer. This effect was moderated by participants' emotional stability, as measured by a Big Five personality inventory. Additional effects of participant variables are also discussed, along with directions for future research.

ACKNOWLEDGEMENTS

The author expresses sincere appreciation to Professor Sara Hodges and Karyn Lewis for their committed assistance in the design, implementation and analysis of the study on which this manuscript is based. Their amazing dedication and insight proved to be both invaluable and inspiring. In addition, special thanks are due to all members of the Hodges Social Cognition Lab of the Department of Psychology for their input throughout the year. I would also like to thank the friends and family who supported me throughout the entire thesis process.

Introduction

Men and women today are continually faced with the challenges and rewards of navigating the social world. Every time we pick up the phone or step outside our own front doors, we encounter the possibility for mutual understanding, respect, and friendship as well as the risk of failing to "connect" with those around us and the damage such confusion can cause to a relationship. Thus the study of empathic accuracy, or the ability to accurately infer what others are thinking and feeling (Ickes, Stinson, Bissonette, & Garcia, 1990), examines one measure of how well individuals understand one another and has the potential to be a very rewarding and important area of research.

Ickes (2003) has aptly characterized empathic accuracy as "everyday mind reading." While some of other people's thoughts may be verbalized and thus transparent, more often these thoughts and feelings must be inferred from imperfect and incomplete cues. This ability to infer what another person is thinking or feeling is one component of empathy, which is considered multifaceted (Myers & Hodges, 2009). It is important to note that empathic accuracy is not the same as empathic concern. Empathic concern is the extent to which one feels sympathy, worry, or compassion for another while empathic accuracy is the actual skill of understanding how that other person feels or what he is thinking. Indeed past research has even suggested that these two components of empathy may be unrelated (Myers & Hodges, 2009).

Past researchers (Ickes, 2001; Ickes, Stinson, Bissonnette, & Garcia, 1990; Klein & Hodges, 2001; Marangoni, Garcia, Ickes, & Teng, 1995) have designed an effective methodology for assessing empathic accuracy. The basic premise of this paradigm –

spearheaded by Ickes and colleagues – is comparison of a target's actual reported thoughts and feelings with the inferences about those thoughts and feelings made by a perceiver. In this method of assessment, people are video recorded while talking about various personal topics ranging from discussing an academic setback to engaging in a therapy session. Immediately after being recorded, these individuals, or "targets," watch the recording of themselves. Whenever the targets remember having a specific thought or feeling *at the time of the initial recording*, they pause the tape and record both the content of the thought they had and the time at which these thoughts occurred. Perceivers then watch the video, and the recordings are stopped at the same points where the targets had paused them. The perceivers are then prompted to infer what the target in the video was thinking or feeling. This procedure occurs for all the thought/feeling points on a videotape and the perceivers' inferences are then compared to the actual thoughts of the targets to obtain an index of empathic accuracy.

Following variations on this procedure, research has revealed many interesting aspects of empathic accuracy. Most importantly, work by Marangoni, Garcia, Ickes, and Teng (1995) as well as Ickes et al. (1990) revealed that empathic accuracy is a fairly consistent trait across subjects. In other words, these researchers discovered that, with some regularity, some people are generally more accurate than others: perceivers that are accurate tend to remain that way across multiple targets. The same trend applies to those that are empathically *in*accurate. Due to this within-perceiver reliability of accuracy scores, researchers have considered the possibility that empathic accuracy may be predictable by individual difference traits of the perceivers. Understanding which individual difference traits are related to empathic accuracy is a fruitful avenue for

research because it may help shape the definition of the concept and serve as a foundation for developing methods for improving people's empathic abilities.

Subsequent research thus attempted to shed light on the relationship between perceiver traits and empathic accuracy. Many individual difference variables that would intuitively seem likely to be positively related to empathic accuracy are, in fact, not. For example, individual differences in empathy might predict empathic accuracy. One measure used to assess the several dimensions of empathy is Davis's (1980) Interpersonal Reactivity Index. Part of this self-report measure assesses perspective taking: the ability and tendency to adopt another's point of view. The IRI also evaluates differences in empathic concern: the extent to which one reports feeling compassion for another. These capacities initially seem like they should aid empathic accuracy, but rather multiple studies have reported that higher scores on self-reported perspective taking and empathic concern are in fact either unrelated to or, in the case of perspective-taking, *negatively* correlated with empathic accuracy (Ickes et al., 1990; Laurent & Hodges, 2009; Myers & Hodges, 2009).

Investigation into sex and gender differences has yielded different results. Many studies suggest that men and women do not significantly differ in their empathic accuracy abilities (Ickes et al., 1990; Marangoni, Garcia, Ickes, & Teng, 1995) while others have found distinct differences in ability between the two sexes under certain circumstances (Ickes, Gesn, & Graham, 2000; Klein & Hodges, 2001) as well as differences attributed to gender (Laurent & Hodges, 2009). Klein and Hodges (2001) have suggested that motivation to be empathically accurate serves as one possible explanation for differences between men and women in ability on empathic accuracy

tasks. Their conjecture corroborated previous findings by Ickes and colleagues (1990), where higher motivation to be empathically accurate was in fact positively related to empathic accuracy. The source of motivation in various studies has ranged from monetary reward for accuracy (Klein & Hodges, 2001) to physical attractiveness of, and perceivers' interest in, a target (Ickes et al., 1990).

Stable individual traits of the perceiver are not the only differences that have been investigated in relation to empathic accuracy. Interpersonal variables have also been found to play a key role in inferring others' thoughts. For example, Marangoni et al. (1995) researched the effects of perceivers' familiarity with a target. The investigators compared the accuracy of perceivers' inferences after only a few minutes of watching a target to the accuracy after seeing the entire videotape. They found that, in general, empathic accuracy improved with increased exposure to a target (Marangoni et al., 1995). This suggests that determinants of one's empathic abilities may not be solely attributable to stable individual differences within the perceiver.

Another interpersonal variable that seems intuitively related to empathic accuracy is similarity between perceivers and targets. This is to say, having "walked a mile in another's shoes" seems like it should aid in accurately inferring another's thoughts and feelings. Surprisingly, however, similarity between perceivers and targets has been found to have little effect on empathic accuracy. When inferring thoughts and feelings of women discussing their experiences as new mothers, perceivers who had never had children, perceivers who were pregnant, and perceivers who were themselves new mothers had generally equivalent empathic accuracy scores (Hodges, Kiel, Kramer, Veach, & Villanueva, under revision). In other words, the similar experience of raising

a newborn did not aid new mother perceivers in any considerable way. In addition to the similarity of new motherhood, research has investigated whether sharing the experience of parental divorce or alcoholism aids empathic accuracy (Hodges, 2005). Results again lack any support for the notion that shared experience leads to greater empathic accuracy. In fact, a qualitative analysis of a sample of studies on clients and therapists revealed that only about one in four studies demonstrate any evidence of a therapist's similarity to a client – either demographically or through shared experience – aiding therapy outcome (Hodges, 2005).

The study by Hodges et al. (under revision) was unique in that, along with the standard comparison of targets' thought/feelings and perceivers' inferences, the procedure included a comparison of both groups' responses to the Maternal Attitudes Questionnaire (MAQ). The MAQ was first developed by Warner, Appleby, Whitton, and Faragher (1997) and in the Hodges et al. study, it was used to assess perceptions of the motherhood-related experiences for both targets and perceivers. Analysis of MAQ responses from the target about her own experience, from the new mother perceiver about her own experience, and from the new mother perceiver about the *target's* experience revealed an interesting finding: New mother perceivers' inferences about targets' maternal attitudes more closely resembled their own experiences than the experiences of the target (Hodges et al., under revision). In other words, new mother perceivers relied on their own attitudes to infer the attitudes of other new moms. This strategy of projection resulted in greater accuracy when the two women's responses (i.e. those of the target and those of the perceiver) on the MAQ were similar, but hurt accuracy when they were different. The researchers suggested that holding an accurate

stereotype of a new mother, and 'projecting' those experiences onto the new mother target, could aid accuracy by providing a similar basis for inference. Thus, if a perceiver (be she a new mother or not) had the right idea of what the stereotypical new motherhood experience was like, that perceiver would be more accurate when inferring the thoughts and feelings of a new mom, as long as the new mother's experience fit the stereotype to some extent. While mothers can be classified as either more stereotypical or more idiosyncratic as a whole, there also may exist variation in how much individual thoughts and feelings within a particular target adhere to the stereotype. This suggests that when studying empathic accuracy, it would be better to take into account stereotypicality *at the level of the thought/feeling*, rather than just at the level of the target.

All members of society hold stereotypes, or widespread beliefs about the personal attributes of people in specific groups (Hilton & von Hippel, 1996; Mackie & Smith, 1998). Although stereotypes can be harmful, they are also sometimes accurate. Such stereotypes, when used as a mental set, may be a key to understanding empathic accuracy. When faced with a situation that requires inferring others' thoughts and feelings, people may tend to rely on mental representations or schemas of a target. As evidence of this, Gesn and Ickes (1999) presented some participants with unaltered videotapes of targets and some with videos that spliced the sequence of the target's discussion. Their study demonstrated that, when cumulative knowledge about a target is available, perceivers develop schemas and then make inferences about the target in a schema-consistent fashion. For example, if after watching a few minutes of a target a perceiver decided that the target was kind and generous, the perceiver would tend to

believe that the target had unselfish thoughts throughout the course of the video. When the videotapes were shown in their unaltered form, this cumulative knowledge tended to help perceivers' empathic accuracy when a target's thoughts were commensurate with the schema, but tended to hurt empathic accuracy when targets' thoughts were schema-inconsistent.

Schema consistency also affects information processing, which could, in turn, affect empathic accuracy. Past research (Hastie & Kumar, 1979; O'Sullivan & Durso; 1984; Stangor & McMillan, 1992) suggests that people tend to have better memory for schema-incongruent information and Hilton and von Hippel (1996) posit that this trend exists because schema-incongruent information is more effortful to process. Thus, if targets strayed from a perceiver's schema for new mothers, the perceiver's empathic abilities could suffer from increased processing effort when making inferences.

The current study was designed to investigate the effects of these 'prepackaged' schemas – stereotypes – on empathic accuracy. The current experiment tested the hypothesis that participants would have greater empathic accuracy for thoughts and feelings that were stereotypical of a new mother, whereas thoughts and feelings that are more idiosyncratic and less stereotypical would be harder for participants to infer. In addition, this study also explored the possible interactive effects stereotypes have with various individual difference variables. Past research has revealed the complexity of the relationships between perceiver traits and empathic accuracy yet still left large holes in our understanding of what leads one to be empathically accurate. Thus the current study analyzed many previously-investigated traits further. In addition, this study included other perceiver traits – such as personality measured on the Big-Five personality factors

– in its analysis in an effort to cast a wider net and perhaps fill in more of the hole in the current understanding of empathic accuracy.

In order to test our main hypothesis, we administered a previously established empathic accuracy task. All participants inferred thoughts and feelings that varied in stereotypicality. Upon completion of the study, participants' empathic accuracy scores were considered specifically in relation to the stereotypicality of the targets' thoughts as well as in relation to individual differences. It was also hypothesized that perceivers would have greater accuracy on thoughts and feelings that were less difficult to infer (Marangoni et al., 1995). In order to investigate whether any empathic accuracy benefits were due to more stereotypical thoughts and feelings being simply easier to infer, analysis also included a measure of inferential difficulty. This allowed for investigation into whether thought/feeling stereotypicality could explain variation in accuracy scores beyond what could be accounted for by inferential difficulty.

Past investigations into empathic accuracy have typically computed an accuracy score for participants by aggregating their performance across inferences. Such calculations, however, preclude the possibility of investigating the effects of individual thought/feeling characteristics on empathic accuracy. The current study utilized hierarchical linear modeling in order to predict empathic accuracy using both individual characteristics of perceivers and the nature of the thoughts and feelings of the targets.

Methods

Participants

Participants were 157 students (49 male) from the University of Oregon who received partial credit for an introductory psychology course requirement in exchange

for participation. Data from 11 subjects were dropped due to equipment malfunction or experimenter error and data from one subject was dropped because subject left the room mid-way through the procedure. Ages ranged from 18-38 years (M = 19.2). The sample was racially diverse as follows: 80.3% of participants self-identified as White, 7.6% as Asian, 3.8% as Hispanic, 3.2% as Native American, and 2.5% as Black. Four participants (2.5%) listed their ethnicity as "other." Non-white participants were generally evenly distributed across conditions. Academically, the sample contained students from a wide range of departments on campus: 25% of participants had not yet declared a major, 19.1% were psychology majors, 8.9% were human physiology majors, and 8.3% were business majors, with remaining participants studying a mix of subjects. Women were evenly distributed across the six video stimulus conditions (described below), with a range of 16 to 21 participants per condition. Men were also generally evenly distributed across conditions one through five, ranging from 8 to 11 participants per condition for five of the conditions, but only 4 men completed the experiment in condition six.

Materials

Individual Difference Measures[1]. Prior to the empathic accuracy task, all participants completed several self-report measures in order to analyze how individual differences may be related to empathic accuracy. These included the Interpersonal Reactivity Index (Davis, 1980), a self-report empathy scale that is composed of four sub-scales: perspective-taking (e.g. "When I'm upset at someone, I usually try to "put

[1] Only individual difference measures that were found to be related to empathic accuracy are explained here. A full list of the measures administered to participants also includes: the Balanced Inventory of Desirable Responding (BIDR; Paulhus & Reid, 1991), the Personality Attributes Questionnaire (PAQ; Spence & Helmreich, 1978), and Rosenberg Self-Esteem Scale (RSE; Rosenberg, 1965).

myself in his shoes" for a while"), fantasy tendencies (e.g. "After seeing a play or a movie, I have felt as though I were one of the characters"), empathic concern (e.g. "I am often quite touched by things that I see happen"), and personal distress (e.g. "I tend to lose control during emergencies"). All 28 items are rated on a 5-point scale that ranges from "does not describe me very well" to "describes me very well."

Participants also completed the Big 5 Mini-Markers (Saucier, 1994): a 40-item questionnaire consisting of adjectives that measure the Big 5 personality traits. For example, the adjective "creative" assesses openness to experience, "practical" measures conscientiousness, "talkative" assesses extraversion, "warm" measures agreeableness, and "fretful" measures emotional stability. Participants rate themselves on a 9-point scale for each of the 40 different adjectives from "extremely inaccurate" to "extremely accurate."

Empathic Accuracy Task. This study used the standard empathic accuracy paradigm established by Ickes and colleagues (1990). Targets were recorded during an interview talking about life experiences. Then, targets immediately watched their interview, recording what their thoughts and feelings at the time of the interview were and when they happened. Participants watched these videos and were asked to infer what the target was thinking or feeling at the same moments previously denoted by the targets.

Participants in this study were randomly assigned to infer the thoughts and feelings of one target. Targets in this experiment were the same new mothers used by Hodges et al. (under revision). These women had all given birth to their first child within two to four months prior to the time of the interview. Their discussions in the

interviews all consisted of describing their lives as new mothers, including topics such as how their lives had changed and whether their expectations of motherhood were being met. At the time of the initial study using these new mothers as targets (Hodges et al., under revision), twenty new mothers completed the process of being videotaped and recording their thoughts and feelings.

Six of these new mother targets were chosen to be used in the current study. Videos ranged in time from roughly two and a half minutes to roughly eight minutes and targets reported between six and nine thoughts and feelings. The principle investigators for the study reviewed the thoughts recorded by each mother and selected three targets that had reported more stereotypical thoughts and feelings and three targets that reported less stereotypical thoughts and feelings.

Procedure

Participants were run by one of five experimenters (1 male, 4 female) and were randomly assigned to condition. Conditions varied only by which new mother participants saw and each condition involved viewing only one new mother. Students participated either alone or in groups of two or three. However, those participating in groups did not interact with each other. All instructions and stimulus materials were presented on individual computers and participants could see only their screens alone. After obtaining informed consent, participants first completed demographic information and then completed the other self-report measures of individual differences, which were presented in a counterbalanced order. Participants then watched a videotape of a new mother describing her experience following the birth of her first child. After having seen the interview once in full, participants were then presented with the same video

that was stopped at the points at which the target reported a thought or feeling. At the end of each segment, participants were prompted to record what they inferred the new mothers to be thinking or feeling at the time the segment ended.

Following the inference task, a subset of participants (n = 97) also coded for stereotypicality the actual thoughts and feelings reported by the targets. These participants coded the stereotypicality of roughly fifteen of the statements provided by the various new mothers used in the study, including the thoughts and feelings of the specific target they saw, in addition to other targets' thoughts and feelings. Participants would not have been able to discern which thoughts belonged to the specific target they saw, as they were presented in random order amongst thoughts/feelings reported by other new mothers used in the study. Participants were asked how characteristic each statement was of a new mother and rated the thought/feeling on a 4-point scale (1=*uncharacteristic*, 2=*neither characteristic nor uncharacteristic*, 3=*somewhat characteristic*, 4=*very characteristic*). For example, participants typically rated the thought "I feel overwhelmed at times of being a mom" as highly stereotypical and the thought "About how sad I was that first week to have a totally different schedule than the rest of the world. But now it's gotten better" as low in stereotypicality. The thoughts and feelings of each target were rated by 31 to 38 participants. The inter-rater reliability of the participants' ratings was calculated for each separate target and Cronbach's alpha scores ranged from .74 to .96. The average reliability across all targets was also acceptable (Cronbach's α = .84). Thus, the stereotypicality of each thought/feeling statement was calculated by taking the average ratings provided by the participants. At the end of the experiment, participants were debriefed, thanked for their time, and

allowed to leave.

Empathic Accuracy Coding. Four independent raters assessed participants' empathic accuracy by comparing the inferred guesses of the participants to actual thoughts of the new mothers. They rated the similarity between the thoughts on a three-point scale by giving either a 0 (*essentially different content*), 1 (*similar, but not exactly alike, content*), or 2 (*basically the same content*). The participant's accuracy score for each inference was averaged across the raters. The inter-rater reliability of the four judges' ratings was calculated separately for each target and the Cronbach's alpha scores ranged from .77 to .88. The average reliability across all targets was also acceptable (Cronbach's $\alpha = .83$). Thus the four judges' ratings were averaged for each inference. Finally, the average thought/feeling empathic accuracy ratings were divided by 2 to create a 0 to 1 scale.

Thought/Feeling Difficulty. A previous set of independent coders (n=7), rated the difficulty of inferring each thought/feeling reported by the new mothers. At each stopping point in the stimulus videos, rather than inferring the new mothers' thoughts and feelings, these raters were presented with what the new mothers had reported thinking and rated how difficult it would be to correctly infer each thought/feeling based on what they had seen in the video. Raters coded for difficulty on a 3-point scale (1=*very difficult to infer*, 2=*somewhat difficult to infer*, 3=*easy to infer*). The inter-rater reliability of the judges' ratings was calculated for each separate target and the Cronbach's alpha scores ranged from .75 to .88. The average reliability across all targets was also acceptable (Cronbach's $\alpha = .80$). Thus the judges' ratings for each inference were averaged.

Results

Plan of Analysis

Past research on empathic accuracy has utilized aggregated empathic accuracy scores, computing one score of average accuracy across the thoughts and feelings for each participant. However, in aggregating empathic accuracy across different types of thoughts and feelings reported by targets, this method ignores substantial and potentially meaningful variation *within a target's thoughts and feelings* that may help explain the complicated relationship between individual differences and empathic abilities. Therefore, we chose to use hierarchical linear modeling in order to account for the nesting inherent in empathic accuracy data. In order to investigate what variables affect empathic accuracy at the level of the thought/feeling, a two-level hierarchical linear modeling structure was utilized with thought/feelings nested within participants. In other words, Level One consisted of participants' empathic accuracy score for each thought/feeling along with that thought/feeling's aggregated ratings of stereotypicality and inferential difficulty. Level Two consisted of variables at the level of the participant, specifically participants' individual difference variables.

An unconditional, or baseline, model that estimated the variability in regression equations across participants was established. This model was of little interest, but rather was used in order to compare whether or not subsequent explanatory models that included predictor variables could account for more of this variance. First, the unconditional model was compared to models including a variable that coded for order of inference (the first inference made, the second inference made, etc.) and for sex to ascertain if empathic accuracy increased linearly over the course of the videotape or

differed significantly for men and women. Subsequently, the baseline unconditional model was compared to a model including stereotypicality and difficulty at Level One (thought/feelings) to assess the legitimacy of including such variables in our calculations. Attempts to explain the variability of the slopes and intercepts of this more complicated model were made by including various individual difference measures at Level Two in our exploratory analyses.

All analyses were performed using the statistical software *HLM 6: Hierarchical Linear and Nonlinear Modeling* (Raudenbush, Bryk, & Congdon, 2000). The grouping variable (i.e., participants) had a sufficient sample size (n=143). Thus, full maximum likelihood estimation procedures were used for all the following models.

Establishing an Unconditional Model

A random coefficients model was used to test whether average empathic accuracy at the thought/feeling level differed across individuals. This unconditional model indicated that the estimated average thought/feeling empathic accuracy was different from zero (β_{00}=0.24, SE=0.01, $t(144)$=26.84, p<.001). The results of the unconditional model also indicated that there was significant variation in average thought/feeling empathic accuracy across individuals ($\chi^2(144)$=207.00, p=.001).

In order to determine if perceivers' accuracy increased with more exposure to a target, a second possible unconditional model that included a linear order term was tested. The slope (β_{10}= -0.0003, SE=0.004) associated with this order term was not significant ($t(144)$=-0.07, p=0.941). The average thought/feeling empathic accuracy did not increase linearly and this variable was excluded from all further models. In order to test whether men and women differed significantly in their empathic accuracy abilities,

a dummy coded sex variable (0=female, 1=male) was included in Level Two. The slope (β_{01}= 0.02, SE=0.02) associated with this term was not significant ($t(143)$=1.06, p=.291). Men and women performed equally well on average thought/feeling empathic accuracy. Thus, the sex variable was excluded from future models.

Testing Conditional Models

Models Testing: Effects of Thought/Feeling Characteristics[2]. In order to test our hypothesis and determine if increased thought/feeling stereotypicality resulted in greater empathic accuracy, a conditional model including the stereotypicality rating of each thought/feeling (grand mean centered) and the inferential difficulty of each thought/feeling (grand mean centered) at Level One was computed. The conditional model of interest included the difficulty and stereotypicality covariates, holding them fixed. In other words, we computed a model that assumed that inferential difficulty and stereotypicality affected each thought/feeling the same way. This model reached convergence in 13 iterations and proved to be a better fit of the data when compared to the baseline model ($\chi^2(2)$=71.21, p<.000). This model showed that the effect of thought/feeling difficulty (β_{20}= 0.13, SE=0.02) on empathic accuracy was significant ($t(1005)$=6.66, p<.001). More difficult thoughts were harder to infer and led to lower empathic accuracy. This model also showed that when difficulty was controlled for, the effect of thought/feeling stereotypicality (β_{10}= 0.04, SE=0.02) on empathic accuracy was marginally significant ($t(1005)$=1.89, p=.059), suggesting that individuals show better empathic accuracy on more stereotypical thoughts/feelings and that this boost is

[2] Conditional models 1 through 3 included Stereotypicality, Difficulty, and both factors together at Level One, respectively. In these models, the thought/feeling characteristics were computed to include random effects for difficulty and stereotypicality. However, these models could not reach convergence within the maximum number of allowed iterations (10,000) and thus their results could not be trusted.

over and above the effect of inferential difficulty. Examination of the variance components of the fixed effects showed that even with the addition of the difficulty and stereotypicality variables to the equation, there was still significant variation in average thought/feeling empathic accuracy ($\chi^2(144)=217.29$, $p=<.001$) left to be explained.

Exploratory Models: Effects of Participant Variables. After establishing the effects of thought/feeling difficulty and stereotypicality, we were next interested in testing whether or not individual difference variables may moderate these effects. It is important to note that the inclusion of participant variables at Level Two was largely exploratory, and no a priori hypotheses were formed. Only models that revealed significant or near-significant effects are reported here.

The final model computed in this study was as follows (see Table 1 for summary of model statistics):

$$emp\ accuracy = \pi_0 + \pi_1(stereotypicality) + \pi_2(difficulty)$$

$$\pi_0 = \beta_{00} + \beta_{01}(empathic\ concern\ from\ the\ IRI) + r_0$$

$$\pi_1 = \beta_{10} + \beta_{11}(emotional\ stability)$$

$$\pi_2 = \beta_{20}$$

In this model the empathic concern subscale of the IRI ($\beta_{01}= 0.03$, SE=0.01) was found to be nearly significantly related to variation in average thought/feeling empathic accuracy ($t(143)=1.90$, $p=.060$). Individuals higher in empathic concern were marginally more empathically accurate. In addition, the emotional stability subscale of the Big-5 personality scale ($\beta_{11}= 0.04$, SE=0.02) was found to significantly interact with stereotypicality ($t(1003)=2.23$, $p=.026$). The effect of stereotypicality on empathic accuracy was different depending on an individual's level of emotional stability.

Specifically, individuals with greater emotional stability showed a greater boost in empathic accuracy from stereotypicality than individuals with lower emotional stability. This final computation also modeled the same general pattern for thought/feeling variables. Thought/feeling difficulty (β_{20}= 0.13, SE=0.02) was again significantly related to empathic accuracy ($t(1003)=6.78$, $p<.001$). Thought/feeling stereotypicality (β_{10}= 0.04, SE=0.02) was *nearly* significantly related to empathic accuracy ($t(1003)=1.92$, $p=.055$). This final model, including empathic concern and emotional stability, significantly improved the fit in comparison to the conditional model that included only difficulty and stereotypicality ($\chi^2(2)=9.19$, $p=.010$). However, examination of the variance components and deviance tests associated with the final model indicated that this model still did not explain all the variance in average thought/feeling empathic accuracy scores ($\chi^2(143)=214.15$, $p<.001$).

Discussion

The present study addressed two hypotheses that focused on predicting empathic accuracy from target variables. First, it was hypothesized that participants would be more empathically accurate on thoughts and feelings that were stereotypical of the salient target schema – the experience of becoming a new mother. This hypothesis was supported in that the relationship between thought/feeling stereotypicality and empathic accuracy was found to be marginally significant. Stereotypicality, then, is a thought/feeling characteristic that should be considered when predicting empathic accuracy since individuals tend to be more accurate on more stereotypical thoughts and feelings than they are on more idiosyncratic thoughts and feelings.

Second, it was hypothesized that participants would be more empathically accurate on thoughts and feelings that were less difficult to infer. Analysis revealed that thought/feeling difficulty did significantly account for some variation in the average thought/feeling empathic accuracy. On the 0-1 empathic accuracy scale, participants' scores increased by roughly .13 when the thought/feeling being inferred dropped by one unit on the difficulty scale. Intuitively this finding is not surprising, and difficulty is still clearly an important factor to account for when analyzing empathic accuracy. In our final model, inferential difficulty explained a large portion of variance, but additional unique variance was also explained by thought/feeling stereotypicality. The increase in empathic accuracy derived from a target's adherence to a pre-existing schema in the perceiver's mind was orthogonal to the increase in empathic accuracy derived from inferential ease.

This finding adheres to what is already known about how people use schemas to make inferences. It seems that rather than solely relying on schemas that have been constructed during the inference-making process as Gesn and Ickes (1999) demonstrated, perceivers also rely on pre-existing schemas when making inferences about a target's thoughts and feelings. Targets' adherence to such stereotypes in their thoughts and feelings made the inference-making process easier for perceivers beyond simply having thoughts and feelings that were easier to infer. This "boost" in empathic accuracy was small (.04), but nearly significant, suggesting that thought/feeling stereotypicality plays a role in determining whether or not a perceiver can accurately infer the thoughts and feelings of a target.

In light of the outcomes of thought/feeling characteristics, the current study also explored the effects of individual perceiver traits. Level Two of the final model revealed a significant interaction effect of emotional stability on how stereotypicality influenced thought/feeling empathic accuracy. Participants higher in emotional stability received a greater "boost" in empathic accuracy from thought/feeling stereotypicality than participants who showed less emotional stability. One explanation for this finding is that participants with low emotional stability were unable to draw upon a "new mom" schema to help them make more accurate inferences either because they were unable to access the schema or the schema did not exist for them. It is also possible that individuals low in emotional stability have a schema they can assess, but it is a less accurate "new mom" schema. When these individuals relied on them during the empathic accuracy task, they received little aid in making accurate inferences.

Further investigation, of course, would be necessary in order to understand why, exactly, perceivers low in emotional stability lack reliable "new mom" schemas. Past research may provide some direction for this investigation. Research has shown a distinct connection in adolescents between low levels of emotional stability and the presence of maladaptive schemas, which represent patterns of distorted thinking about the world (Muris, 2006). Muris (2006) posits that using such distorted thinking patterns results in dysfunctional perception. It is possible that such dysfunctional perception further leads to unreliable schemas about groups of people – new mothers in the instance of the current study.

Out of all the empathy subscales of the IRI, the only dimension of empathy that produced an even marginally reliable result was empathic concern. Independent of the

thought/feeling difficulty or stereotypicality (but after controlling for them), participants that self-reported higher empathic concern showed greater empathic accuracy when inferring the thoughts and feelings of the new mothers. This trend also makes intuitive sense. Perceivers who feel greater compassion for a target are likely more motivated to accurately understand what the target is thinking and feeling. As seen in past research on empathic accuracy (Ickes, et. al, 1990; Klein & Hodges, 2001), greater motivation tends to lead to greater empathic accuracy. In the current study, the effect of empathic concern on empathic accuracy was small (.03), which may explain why it was not found in past research. Plus, the added power of the current study's analyses may have helped tease out the effect. This effect was also found to be non-significant but remained close enough within the margins of significance to warrant further investigation into the empathic concern's effects on empathic accuracy.

Unlike past research (Marangoni, Garcia, Ickes, & Teng, 1995), our model showed no linear increase in empathic accuracy, suggesting that increased exposure to the target new mother did not aid perceivers' empathic abilities. However, in Marangoni, et al.'s (1995) study, perceivers watched video tapes that were twenty five to thirty minutes in length, and each provided thirty thought/feeling opportunities for perceivers to make inferences. All video stimuli in the current study lasted no more than ten minutes and provided only six to nine opportunities for perceivers to make inferences about their target's thoughts and feelings. It is likely that increased exposure to targets does indeed aid empathic accuracy, but the abbreviated nature of the perceiver-target exposure in this study prevented perceivers from fully "learning" about their targets and did not lead to this effect.

In addition, there was no significant difference in empathic accuracy performance between the two sexes. This finding was expected, however, considering the procedure of the empathic accuracy task. Females tend to out perform males at empathic accuracy only when women are more motivated in the inference task than men, which often occurs because of an indicator that some aspect of empathy is being measured (Ickes, Gesn, & Graham, 2000). In studies, this occurs typically when investigators ask perceivers to guess their own empathic accuracy (Ickes, Gesn, & Graham, 2000) or when investigators assess perceivers' sympathy towards targets prior to measuring their empathic accuracy (Klein & Hodges, 2001). Without these special circumstances in the current study, there was no cue to the empathic nature of the inference task and therefore women were not more highly motivated to be accurate than men.

While these findings are interesting, they should be understood in the light of the limitations of the study. In the HLM equation, the reliability estimate is the ratio of true parameter variance to total observed variance. In our final model, the reliability estimate was low (*reliability estimate*=0.31). This low value indicates that empathic accuracy scores in our study may not be completely reliable. It is likely that this effect is a result of the imperfect nature of measuring empathic accuracy. The empathic accuracy assessment paradigm relies on targets accurately recalling their thoughts and feelings during the interview and accurately recording them. In addition, perceivers must then be able to correctly verbalize what they infer targets' thoughts and feelings to be. These issues result in error variance accompanying true variance in scores, lowering the reliability estimate of the model's intercept.

In addition, trends involving stereotypicality may become significant in the presence of a more precise measure of thought/feeling stereotypicality. This measure was limited in the current study due to low variation amongst stereotypicality scores, ranging only from 1.92 to 3.68 on a 4-point scale. Using a scale that allows for more variation (perhaps one set on a scale utilizing more than four points) could make clearer the exact role that thought/feeling stereotypicality plays in empathic accuracy.

Finally, the sample of perceivers used in this study was largely female, young adult, Caucasian college students. These findings thus must be generalized to other groups with caution. In addition, participants completed the empathic accuracy measure with only one target. One cannot conclude from this study whether the effects of stereotypicality will display cross-target consistency. The effect of stereotypicality may vary depending on whether stereotypes are based on thoughts and feelings associated with an important experience (e.g. becoming a new mother or taking the Graduate Record Exam) or are based simply on characteristics of groups of people (e.g. fraternity brothers or Midwesterners). Future research would benefit from participants inferring thought/feelings from multiple targets that come from a variety of backgrounds and discuss different topics.

Currently, psychologists interested in how people understand others face many questions. There seem to be some clear predictors of empathic accuracy and important, influential factors that are necessary to account for when analyzing a perceiver's ability to infer the thoughts and feelings of another. However, in this study – and in the body of past research in general – the factors investigated in relation to empathic accuracy have not been able to fully explain what makes a perceiver empathically accurate. We

discovered effects of thought/feeling stereotypicality, inferential difficulty, as well as perceivers' levels of emotional stability and empathic concern, but there was still significant variation left to explain. There is still a clear hole in the understanding of which individual difference factors and thought/feeling characteristics predict empathic accuracy. Further research into these variables will undoubtedly shed more light on what leads a person to be a good perceiver.

References

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *Catalog of Selected Documents in Psychology, 10*, 85.

Gesn, P. R., & Ickes, W. (1999). The development of meaning contexts for empathic accuracy: Channel and sequence effects. *Journal of Personality and Social Psychology*, 77, 746-761.

Hastie, R. & Kumar, P. A. (1979). Person memory: Personality traits as organizing principles in memory for behaviors. *Journal of Personality and Social Psychology, 37*, 25-38.

Hilton, J. & von Hippel, W. (1996). Stereotypes. *Annual Review of Psychology, 47,* 237-71.

Hodges, S. (2005). Is how much you understand me in your head or mine? *Other minds: How humans bridge the divide between self and others,* 298-309.

Hodges, S. D., Kiel, K. J., Kramer, A. D. I., Veach, D., & Villaneuva, R. (under review). Giving birth to empathy: The effects of similar experience on empathic accuracy, empathic concern, and perceived empathy. Unpublished manuscript, University of Oregon, Eugene.

Ickes, W. (2001). Measuring empathic accuracy. In J. Hall & F. Bernieri (Eds.), *Interpersonal sensitivity: Theory, measurement and applications* (pp. 219-241). Hillsdale, NJ: Lawrence Erlbaum.

Ickes, W. (2003). Everyday mind reading: Understanding what other people think and feel. Amherst, NY: Prometheus Books.

Ickes, W., Gesn, P. R., & Graham, T. (2000). Gender differences in empathic accuracy: Differential ability or differential motivation?. *Personal Relationships*, *7*, 95-109.

Ickes, W., Stinson, L., Bissonnette, V., & Garcia, S. (1990). Naturalistic social cognition:

Empathic accuracy in mixed-sex dyads. *Journal of Personality and Social Psychology, 59*, 730-742.

Klein, K. K. J. & Hodges, S. D. (2001). Gender differences, motivation, and empathic accuracy: When it pays to understand. *Personality and Social Psychology Bulletin, 27*, 720-730.

Laurent, S. M., & Hodges, S. D. (2009). Gender and empathic accuracy: The role of communion in reading minds. *Sex Roles, 60,* 387-398.

Mackie, D. M., & Smith, E. R. (1998). Intergroup relations: Insight from a theoretically integrative approach. *Psychological Review, 105*, 499–529.

Marangoni, C., Garcia, S., Ickes, W., & Teng, G. (1995). Empathic accuracy in a clinically relevant setting. *Journal of Personality and Social Psychology, 68*, 854-869.

Muris, P. (2006). Maladaptive schemas in non-clinical adolescents: Relations to perceived parental rearing behaviors, Big Five personality trait factors, and psychopathological symptoms. *Clinical Psychology and Psychotherapy, 13*, 405-413.

Myers, M. W., & Hodges, S. D. (2009). Making it up and making do: Simulation, imagination and empathic accuracy. To appear in K. Markman, W. Klein, & J.

Suhr (Eds.), *The handbook of imagination and mental simulation* (pp. 281-294). New York: Psychology Press.

O'Sullivan, C. & Durso, F. (1984). Effect of schema-incongruent information on memory for stereotypical attributes. *Journal of Personality and Social Psychology, 47*, 55-70.

Paulhus, D. & Reid, D. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology, 60,* 307-317.

Raudenbush, S., Bryk, T., & Congdon, R. (2000). *HLM 6: Hierarchical Linear and Nonlinear Modeling.* Scientific Software International.

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.

Saucier, G. (1994). Mini-Markers: A brief version of Goldberg's unipolar Big-Five markers. *Journal of Personality Assessment, 63*, 506-516.

Spence, J. T., Helmreich, R. L., & Stapp, J. (1974). The Personal Attributes Questionnaire: A measure of sex role stereotypes and masculinity-femininity. *JSAS: Catalog of Selected Documents in Psychology, 4,* 43-44.

Stangor C. & McMillan D. (1992). Memory for expectancy-congruent and expectancy-incongruent information: a review of the social and social developmental literatures. *Psychological Bulletin, 111*, 42–61.

Warner, R., Appleby, L., Whitton, A., & Faragher, B. (1997). Attitudes toward motherhood in postnatal depression: development of the Maternal Attitudes Questionnaire. *Journal of Psychosomatic Research, 43*, 351-8.

Table 1

Summary of HLM Analyses for Variables Predicting Empathic Accuracy

	Unconditional Model				Conditional Model			
	Coeff.	SE	*t*	*p*	Coeff.	SE	*t*	*p*
Fixed Effect								
Intercept	0.24	0.01	26.84	<.000	0.24	0.01	27.62	<.000
Empathic Concern	—	—	—	—	0.03	0.01	1.90	.060
Stereotypicality	—	—	—	—	0.04	0.02	1.92	.055
Emot. Stability	—	—	—	—	0.04	0.02	2.23	.026
Difficulty	—	—	—	—	0.13	0.02	6.78	<.000
	SD	df	χ^2	*p*	SD	df	χ^2	*p*
Random Effect								
Individual	0.24	—	—	—	0.23	—	—	—
Intercept	0.06	144	207.00	.001	0.06	143	214.15	<.000

Note. All the fixed effects are reported using the robust standard errors.