

Specially Motivated, Feminine, or Just Female:  
Do Women Have an Empathic Accuracy Advantage?

Sara D. Hodges, Sean M. Laurent, & Karyn L. Lewis

University of Oregon

Chapter to appear in J. L. Smith, W. Ickes, J. A. Hall, & S. D. Hodges (Eds.), *Managing interpersonal sensitivity: Knowing when—and when not—to understand others*. NY: Nova Science Publishers.

Our local newspaper recently printed a cartoon in which one teenage female character asked her young adult brother why he's glum and he responds "It's nuthin'." In the next frame, with only his laconic answer to go on, she summarizes his current career, social and romantic woes. In the final frame, he asks "How do you DO that?" to which she responds with a shrug, "I'm female."

The funny pages and aphorisms are full of stereotypes about women's "intuition" and what seems at times, perhaps particularly to men, women's uncanny ability to read others' thoughts. This chapter asks whether this reputation is deserved, and if so, wherein lies the source of women's skill. Greater aptitude? Greater motivation? Or some combination of the two? Two decades of research using Ickes' (1993; 2003; Ickes, Stinson, Bissonnette, & Garcia, 1990) paradigm for measuring empathic accuracy—the ability to accurately infer other people's thoughts and feelings – has provided quite a bit of empirical data that address this question.

The results of this research, however, have not been generous in providing simple answers. A rich mix of motivation and gender role expectations, combined with a messy set of outcomes that defy simple clarity, must be considered to understand women's "sometimes" advantage at empathic accuracy. No wonder "women's intuition" has a gauzy mystique to it. When we pull back the veil on empathic accuracy, the female advantage is considerably less impressive than the stereotypes would have us believe. And yet, as it turns out, there is something to the myth. After reviewing the literature linking women and empathic accuracy in order to write this chapter, our conclusions, which (spoiler alert!) we will give away up front and later elaborate upon, have been whittled down to three qualified statements:

1. Questioning or challenging a person's interpersonal sensitivity motivates greater empathic accuracy from women, but not from men, which sets this particular manipulation apart from other incentives that seem more gender-neutral.
2. Communion, the habitual motivation to attend to and make connections with others that is a central feature of the feminine gender role, is predictive of empathic accuracy—but only some of the time. That lukewarm endorsement gets a slight boost when one notes that communion is one of few individual difference variables whose ability to predict empathic accuracy has been replicated. On the other hand, as often as not, there is no discernible effect of communion—and even when there is, it is very small.
3. Even in the absence of explicit gender-specific motivation and higher communion scores, women probably hold a slight general edge over men in terms of their empathic accuracy performance, although this advantage is nothing approaching that which would merit the cartoon vignette that opened this chapter.

The first two statements make explicit reference to motivation, which for this chapter we define as the desire and drive to accomplish a certain goal or to create a state consistent with an ideal. In simplest terms, our first statement above refers to a motive created by a situation and the second refers to a motive originating within the person. The last statement is agnostic in terms of motivation: When it comes to the small (sometimes vanishingly so) amount of variance in empathic accuracy that is explained by people's sex (after controlling for their psychological femininity and masculinity), the female advantage can be explained in terms of a combination of biology, socialization, experience, cultural roles, and expectations, all of which in turn may be

mediated by some form of motivation. However, mapping those pathways is beyond the scope of the chapter (not to mention, we believe, beyond the scope of currently available data).

### **Empathic Accuracy – Hybrid Offspring of Interpersonal Accuracy and Empathy**

Empathic accuracy, as we will use the term in this chapter, is a specific skill that could be considered a union of the broader construct of interpersonal accuracy and the broader construct of empathy. The former constitutes a big tent encompassing topics related to people's ability to infer and make judgments about psychological characteristics associated with a target person (including that target person's traits, emotions, thoughts, and intentions). Interpersonal accuracy is operationalized as a greater correspondence between these judgments and some criterion that is considered representative of the target person's "true" characteristics. Such criteria may take the form of the target person's self-report; others' (including experts') ratings of the target person; or behavioral criteria that have been empirically linked to certain psychological constructs (e.g., speech errors and disfluency may indicate anxiety; performance on a particular test may indicate verbal intelligence).

Women have well-documented advantages when it comes to other measures of interpersonal accuracy. They notice and remember details of other people's appearance more than men (Hall & Schmid Mast, 2008) and they are consistently better at reading nonverbal cues (Hall, 1978; Hall, 1984; McClure, 2000), seeming only to lose their edge over men when trying to nonverbally decode the truth from deceptive messages (DePaulo, 1992; Hurd & Noller, 1988; Rosenthal & DePaulo, 1979). However, with empathic accuracy as our focus, this chapter will concentrate only on a narrow segment of interpersonal accuracy, looking exclusively at accuracy

in inferring other people's thoughts, using the target's self-reported thoughts as the accuracy criterion.

Defining empathy—empathic accuracy's other progenitor—is more complicated than defining interpersonal accuracy, as the concept has been defined frequently and inconsistently, probably in part because empathy is a multidimensional construct made up of independent strands (Davis, 1983; Hodges & Myers, 2007; Ickes, 2003). Anecdotally, when we ask people to name the “most empathic” person they know and to tell us why, their answers describe a person who is not only able to read their thoughts (empathic accuracy), but who also feels for them (empathic concern) and communicates that feeling (the empathy is perceived). Thus, quintessentially empathic people are “triple threats” who possess all three qualities; however, there is ample evidence that these three strands of empathy (empathic accuracy, empathic concern, and perceived empathy) do not always co-vary together (Fernandez-Duque, Hodges, Baird, & Black, in press; Hodges, Kiel, Kramer, Veach, & Villanueva, 2010).

Belief in women's superior empathic accuracy may stem from the belief that they outscore men in other empathic domains—domains that might intuitively seem related to empathic accuracy but may turn out not to be. For example, in 1983, Eisenberg and Lennon reviewed the literature on emotional empathy and found that females' primary advantage over men was on self report measures of empathy, and not when behavioral observations or physiological measures were used. Apparently, it was more important for women than men to portray themselves as empathic, something Eisenberg and Lennon attributed to the fact that caring about and attending to others is a fundamental component of the female gender role (Helgeson, 1994; Spence & Helmreich, 1978).

Consistent with Eisenberg and Lennon's review, women also habitually outscore men on Davis's widely used self-report scale for measuring individual differences in empathic tendencies, the Interpersonal Reactivity Index (IRI; Davis, 1980). Clearly, women are asking the world to believe that they *feel* more empathy. Additionally, along with women's advantage at reading nonverbal messages (discussed above), women are also better at *sending* nonverbal messages (Wagner, MacDonald & Manstead, 1986), raising the possibility that part of their "empathic mystique" is that when they do feel empathy, they are perceived as more empathic because they are better able to communicate their empathy nonverbally. Thus, perceptions of a female advantage in empathic accuracy may be driven in part by women's desire to *appear* more understanding.

Given these various female advantages in domains related to both interpersonal accuracy and empathy, one could not be blamed for hypothesizing that women might outperform men on empathic accuracy—the ability to infer another person's thoughts. Measuring this highly specific skill has been made even more specific by the fact that several research labs studying empathic accuracy use the same paradigm developed by Ickes. In this paradigm (see Ickes, 2001; 2003 for reviews), a target person is video recorded either while describing some personal event or while in dialogue with another person. After the recording is completed, the target person then views the video and is asked to report any time he or she remembers having had a thought or feeling during the initial recording. The exact time that the thought or feeling occurred is noted, and the target person provides a description of the contents of his or her mind at that time.

The actual measurement of empathic accuracy occurs when the same video recording is shown to perceivers – the people whose empathic accuracy is being measured. The video recording is stopped at the same time points that the target indicated having had a thought, and

perceivers are asked to guess what the target was thinking at that point. These inferences are then compared to the target's reported thought descriptions and scored for accuracy. Ickes' original paradigm (see Ickes, 2001) used trained coders to score accuracy; subsequent studies have also introduced the use of targets themselves as coders (Hodges, 2004; Hodges, 2009; Hodges et al., 2010). In measuring empathic accuracy, the "standard stimulus paradigm" (Ickes, 2001) may be used, with many participants inferring thoughts from the same target. Alternatively, in studies where the stimuli are dyadic conversations that the participant just took part in, the "unstructured dyadic interaction paradigm" (Ickes, 2001) is used, with each participant's empathic accuracy being measured with a different stimulus video.

### **Motivation made for a woman**

Initial results by Ickes and his colleagues showed no differences between men and women for empathic accuracy, suggesting that this was *not* a domain in which women were more accurate or empathic than men. Indeed, a string of nine studies (cited in Ickes, Gesn, & Graham, 2000) all showed no significant advantage for either sex, with non-significant nods in favor of one sex or the other occurring roughly equally. However, a slight change in methodology produced a surprising change in the pattern of results: When the reporting form that perceivers used to write down their inferences also included a question asking perceivers to indicate how well they thought they did at inferring the target's first thought and then repeated this same question again after all subsequent inferences, women significantly outscored men in terms of accuracy. Ickes, Gesn, and Graham (2000) theorized that the addition of the performance check ("How well do you think you did?") triggered something for women that it didn't trigger for men.

Specifically, it highlighted that empathic accuracy, a skill associated with the female gender role, was being measured, which in turn motivated women, but not men, to do better.

Consistent with this interpretation of Ickes et al.'s results, Klein and Hodges (2001) found that highlighting the "empathic nature" of inferring other people's thoughts improved women's accuracy. Participants were asked to watch a target discussing her difficulties getting into graduate school. (This target is affectionately known by the researchers who use her as "GRE Woman," because her quantitative score on the Graduate Record Exam presented the biggest hurdle to getting in to graduate school.) When participants were asked to rate how much empathic concern they felt for the target before inferring her thoughts, women were found to be significantly more empathically accurate than men. However, if participants rated how much empathic concern they felt for the target after trying to infer her thoughts, then women were not more empathically accurate than the men. Similarly, when participants were told that the empathic accuracy task measured empathy, women were marginally more accurate than men, whereas when participants were told that the task measured cognitive ability, women's and men's empathic accuracy scores were comparable. In other words, implicitly or explicitly connecting performance on the empathic accuracy task to familiar empathy constructs appeared to communicate to women that this was a task that they should try to do well on.

A third study sealed the deal, unambiguously challenging women to show their skills: Thomas and Maio (2008) found that when women's stereotypic advantage over men in the domain of social intuition was questioned, women's empathic accuracy improved, relative to a control condition. Notably, a second study demonstrated that telling men that they were potentially deficient in interpersonal intuition did not improve men's performance, once again

suggesting that there appears to be a special category of empathy-relevant motivators that work for women only.

This trio of findings should not be interpreted as indicating that men *cannot* be motivated to be more empathically accurate (see also Lewis, Smith, & Hawkinson; and Thomas, Legood, & Lee; both this volume). In Klein and Hodges' (2001) second study, the researchers once again asked all participants to rate their empathic concern for a target before inferring the target's thoughts, the manipulation that increased empathic accuracy for women in their first study. However, in this second study, they also manipulated whether the participants were promised monetary payments proportional to their empathic accuracy. Empathic accuracy improved across the board in the payment condition; and, notably, men's performance increased such that women's accuracy was no longer significantly higher than men's. In contrast, in a control condition with no monetary incentives, women were once again more accurate than men, replicating the results from Study 1.

As further indication that men can at times be motivated to be more accurate, Thomas and Maio (2008) found that suggesting to college-aged men that interpersonally sensitive men were more sexually successful with women improved the men's accuracy. So, with the promise of payoffs such as money and sex, men can ratchet up their empathic accuracy. However, at least in terms of using money as an incentive, this probably does not make them any different from women. (Women's money-motivated performance increased less than men's in Klein and Hodges' 2001 study, but this may in part have been due to the fact that women were already closer to ceiling; Thomas and Maio's 2008 paper did not test whether women could be motivated to be more accurate if *they* were told the skill made them sexually attractive to *men*.)

Thus, men's empathic accuracy (and, as best we know, women's too) appears to be responsive to extrinsic motivators (like money and the promise of sex). For women, however, there appears to be a more intrinsic and uniquely gendered motivation to be empathically accurate—a motive which can be triggered by insinuations about concern for others (e.g., Ickes, Gesn, & Graham, 2000; Klein & Hodges, 2001) as well as broader threats to women's presumed advantage in this domain (Thomas & Maio, 2008). The three manipulations that have worked to improve women's performance can all be seen as “challenges” to women's ability to read others' thoughts, from the very subtle “How well do think you did?” (Ickes, Gesn, & Graham, 2000), to the slightly more directive, “This is a test of skills related to empathy” (Klein & Hodges, 2001), right up to the outright throw down, “There is reason to believe women are actually NOT so good at this” (paraphrasing Thomas & Maio, 2008).

The other thing that these three successful manipulations share is that all of them make some reference to components of communion, or a communal orientation, which forms the central core of the female gender role in modern gender theories. Communion has been defined as a “focus on others and forming connections” (Helgeson, 1994, p. 412), in contrast with agency, or a focus on the self, considered to be the core component of the male gender role. Thus, even though women may not always or even generally outperform men on empathic accuracy tasks, if their performance is linked in their minds to this central aspect of the female gender role, they do better.

### **Is being “feminine” motivation enough?**

The subtlety of some of the manipulations described above suggests that the triggers that bring the female gender role into awareness for women are quite sensitive. Indeed, it seems

reasonable that communion may never be far out of awareness for some women—and for that matter, some men—for whom communion is a central part of the self-concept. In other words, for people of both sexes who habitually focus on other people and connecting with those people, there may be a chronic motivation to accurately infer others' thoughts in order to better understand and “connect” with them. Do people who score high on individual difference measures of communion do better at inferring other people's thoughts? If so, it would constitute a rare victory in the so far not-very-successful campaign to find individual difference measures that consistently predict empathic accuracy (see Ickes et al., 1990; Ickes, Buysse et al., 2000).

Unfortunately, as seems always to be the case with research linking individual differences to empathic accuracy, communion's role is complicated and inconclusive. The story starts out in straightforward way: Laurent and Hodges (2009a) tested the communion hypothesis, measuring communion using the femininity scale of the Bem Sex Roles Inventory (BSRI; Bem, 1974) while controlling for socially desirable responding. The participants in their study first completed the empathic accuracy task using the same target that had been used in Klein and Hodges (2001)—the woman who had trouble getting into graduate school because of her test scores (a.k.a. “GRE woman”). Later, the same participants then completed a bundle of individual difference measures including the BSRI and the Marlowe-Crowne Social Desirability Scale (Crowne & Marlowe, 1960).

Laurent and Hodges' results supported the communion hypothesis: After controlling for social desirability, more communal participants (as well as female participants, after controlling for communion) were more empathically accurate. Communion would seem to be a more desirable trait for women than men; however, men, not women, showed greater correlations between the self report communion variables and social desirability. In other words, in this study,

for those who reported high levels of concern for others, for men, a bigger part of that concern was apparently concern for looking good. However, for both sexes, the “pure” part of communion—i.e., that part unrelated to social desirability concerns—was correlated with greater empathic accuracy. Similar but less robust results were found using the expressiveness subscale of Spence’s Personal Attributes Questionnaire (PAQ; Spence, Helmreich, & Stapp, 1974). This subscale, which is also a measure of traits associated with the female gender role, correlated .69 with the BSRI femininity/communion subscale.

Further support for the communion hypothesis comes from another study in our lab (Laurent, Hodges, & Lewis, unpublished data). In this study, the participants completed the BSRI and the PAQ, among other measures, on the Internet at least a week after coming to the lab to be tested for empathic accuracy. In this study, the empathic accuracy task consisted of participants watching one of two videos: In both cases, the participants watched a pair of male college students discussing their parents’ divorces, with one of the men designated as the target (i.e., the person whose thoughts were to be inferred). In one pair, the target was quite transparent (his thoughts were relatively easy to infer from what he was saying and the context of the conversation), whereas the target in the other pair was quite hard to read. This difference in readability was independently confirmed by a separate group of research participants who saw the video *and* the targets’ thoughts and who were asked how hard it would be to infer each thought after having seen that segment of the video. Another group of participants also went through the same protocol (individual difference measures separated in time after empathic accuracy), but in place of the divorce targets, these participants saw “GRE woman” (see Laurent & Hodges, 2009b).

Once again, the femininity score from the BSRI significantly predicted the participants' empathic accuracy across all three videos (the transparent male from the parental divorce dyad; the hard to read male from the parental divorce dyad; and "GRE woman").<sup>1</sup> As in Laurent and Hodges (2009a), similar results were also found using the expressiveness scale from the PAQ. The parallel results from Laurent and Hodges (2009a) and Laurent et al. (unpublished data) represent a relatively rare replication of an individual difference measure predicting empathic accuracy when using a standard stimulus target across participants. Furthermore, the communion result was replicated not only for "GRE woman" (Laurent & Hodges, 2009b) but also for the two different "standard stimulus" male targets used in the study by Laurent et al. (unpublished data).

Buoyed by this success, other projects in our lab have also included communion as a possible predictor of individual difference variance in empathic accuracy. In her master's degree project exploring the relationship between empathic accuracy and nonverbal skills, Lewis (2008) gave participants the BSRI and the PAQ along with other personality measures. Lewis used "GRE Woman" as her target for measuring empathic accuracy, which she measured after administering the personality measures and either just before or just after administering the *Interpersonal Perception Task* (IPT-15; Costanzo & Archer, 1989) and the *Diagnostic Analysis of Nonverbal Accuracy* (DANVA; Nowicki & Duke, 1994), which are both measures of nonverbal sensitivity. In a separate undergraduate honors thesis study, Locher (2009) collected PAQ scores, as well as scores on Davis's Interpersonal Reactivity Index (IRI; 1980) as part of a battery of other personality measures given to his participants prior to showing them one of six

---

<sup>1</sup> Laurent et al. (unpublished data) sought to decrease concerns about socially desirable responding which they suspected might have been pronounced in the Laurent and Hodges (2009a) study in part because the self report measures of communion followed soon on the tail of the empathic accuracy task. Although it is not reported in the paper, among other measures, participants also completed the Balanced Inventory of Desirable Responding (BIDR; Paulus, 1984). Not only was it *not* necessary to control for social desirability in order to find the link between communion and empathic accuracy in this study; there was no significant correlation between social desirability and communion—for men or women. We think this is because the social desirability measure and other individual difference measures were collected at least a week later, and in an entirely different context.

empathic accuracy targets. The targets in his study were all first-time mothers of infants talking about their experiences as a new mother, taken from a set of videotaped interviews collected by Hodges et al. (2010).

But alas: The communion measures did not predict empathic accuracy in either of these studies (nor did any of the other individual difference measures, including both nonverbal sensitivity tests). Furthermore, Klein and Hodges (2001) reported in a footnote that BSRI scores did not predict empathic accuracy in their Study 1, and a further re-analysis of their data revealed that BSRI femininity had no effect on empathic accuracy in *any* of the conditions in that study. Furthermore, among the studies showing a null effect for communion, in the Lewis (2008) and the Locher (2009) studies, communion scores were collected as part of a series of individual difference measures that came *before* measuring empathic accuracy, whereas in the Klein and Hodges (2001) study, they were collected after the motivation manipulations and after empathic accuracy assessment, as was the case in the Laurent and Hodges (2009a) study, so placement of the scale does not readily predict when there will be null effects. It must further be pointed out that in the Laurent and Hodges (2009a) study, one of the studies in which we are considering communion to have “worked,” communion was not strictly communion, but a residual of communion after controlling for social desirability. This leaves Laurent et al. (unpublished data) as the only study with a straight communion measure significantly predicting empathic accuracy—and notably, this was the study where all the individual difference measures, including communion, were collected at least a week later in a different setting from the rest of the study.

In sum, at this point, communion appears to have much in common with other individual difference measures when it comes to empathic accuracy: It is unreliable as a predictor of

empathic accuracy, and/or it may interact in complex and not immediately comprehensible ways with other aspects of the data collection, such as motivation manipulations and the order in which the participants complete various measures. To further take the wind out of communion's sails, when significant effects of communion have been found, they are very small (an  $r^2$  of .03 for the communion residual in the study by Laurent and Hodges (2009a); an  $r^2$  of .02 across the videos used in Laurent et al., unpublished data).

### **Women's (diminutive and limited) intuition?**

Thus far, we have been able to at least tentatively link superior empathic accuracy to women by means of both a situational motivator and a dispositional motivator, both of which are uniquely associated with women in some way: Women can be motivated to be more empathically accurate by appealing to—or perhaps more correctly, challenging—their ability to behave in ways consistent with the female gender role, and people who score high on communion (who tend to be women), which is a central component of the female gender role, are sometimes more empathically accurate. However, that's not the whole story, due to a small nagging loose end: Even without challenging their gender role and even when controlling for communion scores, women are sometimes *still* significantly more empathically accurate than men—and they are often *non-significantly* more empathically accurate than men, a point we turn to now.

Ickes, Gesn, and Graham (2000) were able to neatly categorize all of the then-available studies at the time as either “Advantage: Women,” all of which included the question about “How well did you think you did?”; or “No Advantage,” none of which had the “how well did you do?” question and which also included a few studies whose results non-significantly favored

men. However, since then, a number of studies have favored women to varying degrees, even in the absence of the “How well did you do?” question. Notably, the studies producing a significant advantage for women have all used the same “standard stimulus” target person (“GRE Woman”)—but use of this particular target is no guarantee that women will be more accurate, and, furthermore, it is not easy to explain when the female advantage will emerge.

Using “GRE Woman,” Laurent and Hodges (2009a) found a significant sex effect for empathic accuracy, such that women performed better than men, even after accounting for the fact that women scored, on average, higher on communion. Klein and Hodges (2001) also used this same video as their target stimulus when they found a significant effect for sex. More importantly, even in the conditions in Klein and Hodges designed to produce minimal empathic accuracy differences between men and women (i.e., collecting the empathic concern data after assessing empathic accuracy or telling the participants that empathic accuracy was a cognitive skill), the effect sizes for sex were similar to that found for sex in Laurent and Hodges (2009a). Finally, in her master’s thesis that used “GRE Woman” as the target person, Lewis (2008) also found that women significantly outperformed men in their empathic accuracy ( $d = .31$ ).

However, once again, any hope of a simple, consistent pattern of results is dashed by the results of additional studies: Pickett, Gardner, and Knowles (2004) used “GRE woman” in a study examining how social exclusion affects interpersonal sensitivity and reported no significant main effects for sex on empathic accuracy. Their results for sex are summarized briefly in a footnote, and understandably, no effect size is given (only that there was no significant effect). It is possible that the manipulations to create Pickett et al.’s different experimental conditions may have reduced sex differences in some cells, or perhaps the fact that all participants completed a “Need to Belong Scale” prior to empathic accuracy had some impact.

However, harder to explain away is why Laurent and Hodges (2009b) found no sex effects using this same particular target: the women's advantage was far from significant, with empathic accuracy scores of .32 versus .30,  $d=.10$ . In Laurent and Hodges' (2009b) study, the empathic accuracy task was the first thing participants did upon their arrival in the laboratory, so there were no prior measures that could have reduced the sex effect. Furthermore, the sample who saw "GRE Woman" contained 49 women and 27 men, so the sample size was in the ballpark range of other studies that have shown significant sex effects on empathic accuracy.

Any number of factors may contribute to when sex differences are found for this target; we can only speculate about them at this point. Perhaps something as haphazard as the time of year or academic status affects participants' reactions to "GRE Woman": Depending on their own scholastic stresses or plans for the future, "GRE Woman's" tale of woe may have been more or less compelling to certain college student participants. Still, even with the two non-significant results (Pickett et al., 2004; Laurent & Hodges, 2009b), the odds are in favor of "GRE Woman" producing at least a small—and sometimes larger—empathic accuracy advantage for women.

What *is* it about this video? Perhaps there is something about her or her plight that particularly speaks to women. There is nothing explicitly stereotypically female in the content of what she says (e.g., her inability to get into the graduate program of her choice is not attributed to sexism, nor does she talk about stereotypical problems faced by future female professionals such as tension between career and family). However, she does talk about her low quantitative score on the GRE, which could influence women who are probably well aware of unflattering stereotypes about their math skills (Spencer, Steele, & Quinn, 1999). "GRE Woman's" appearance is quite feminine, and she is physically attractive in a traditional way. Ironically, however, past research by Ickes et al. (1990) suggests that, if anything, this attractiveness should

boost *men*'s empathic accuracy, given their finding that opposite sex pairs did a better of job inferring each other's thoughts when their partner was attractive.

Our best guess is that there is something about the pathos of this one target that speaks loudly enough to female subjects that, at some level, even in the absence of other cues, they come to believe that the focus of the study might be related to what this particular target is telling them. Indeed, anecdotally, researchers and experimenters (especially female ones) who have shown this video to participants often report that "GRE woman" really "gets" to them (the effect may occur in part because academic failure looms as such a threatening concept for this specific segment of society!). Another possibility is that "GRE Woman" may send particularly strong nonverbal cues about what she is thinking, and consistent with women's nonverbal decoding advantage, female perceivers pick up on these cues to a greater extent than male perceivers. Even though perceivers seem to use verbal information more than nonverbal information when it comes to empathic accuracy (Gesn & Ickes, 1999; Hall & Schmid Mast, 2007), really explicit nonverbal cues might still help, and we think that "GRE Woman" may be particularly nonverbally expressive.

The special "magic" of "GRE Woman" is consistent with work by several interpersonal perception researchers who in recent years (Funder, 1995; 1999; Thomas & Maio, 2008; Zaki, Bolger, & Ochsner, 2008) have highlighted that accuracy is a product of aspects of the perceiver *and* aspects of the target. For example, Zaki et al. found that participants who reported greater trait empathy were only more accurate at identifying how positive or negative targets felt when the targets were highly expressive (this task was a somewhat different than inferring the targets' thoughts). Similarly, Thomas and Maio's (2008) gender-tailored motivations only improved empathic accuracy for easy-to-read targets. The combination of female perceivers and "GRE

Woman” may be another specific illustration of such an interaction—one that has come to light in part simply because it has been tested in multiple studies.

“GRE Woman” is the only target for whom a (mostly) consistent female empathic accuracy advantage has been found in the absence of additional motivations, but it must be noted that, in preparing this chapter, we looked beyond non-significant  $p$  values to the effect sizes for sex in many of our lab’s recently collected datasets that used other targets. Although there were no other videos that produced a *significant* female advantage, in *every* case, the non-significant difference favored women. These results included two studies using “standard stimuli”: accuracy in reading the two male targets discussing parental divorce (both the transparent and hard-to-read one) in Laurent et al. (unpublished data) with sex effect sizes of  $d = .39$  and  $d = .28$ , respectively; and average accuracy across the six new mother targets in Locher’s (2009) honors thesis with an effect size  $d = .18$ . The other three studies involved members of dyads inferring their partner’s thoughts, such that different thoughts and feelings were being inferred in each dyad—what Ickes (2001) refers to as the unstructured dyadic interaction paradigm. One of these studies (Myers, 2009) consisted of pairs of same sex college students getting to know each other and produced a sex effect size of  $d = .06$  for empathic accuracy. The other two (Hodges, 2004; 2009) involved same sex pairs of college students discussing parental divorce (sex effect sizes of  $d = .30$  and  $d = .32$ , respectively).

These last three studies all used same sex dyads, which introduces a further confound—men’s empathic accuracy was thus always measured using male targets, and female’s empathic accuracy was always measured using female targets. Although we know of no evidence that people are more empathically accurate with same sex targets (see Klein & Hodges, 2001), research showing that women are more nonverbally expressive than men (DePaulo, 1992; Hall,

1984) would suggest that female perceivers in these last three studies might have had an advantage due to targets they saw. However, this speculation is tempered somewhat by Gesn and Ickes' (1999) finding that empathic accuracy relies predominantly on verbal information (see also Hall & Schmid Mast, 2007), a finding which suggests that nonverbal expressiveness contributes only minimally to perceivers' empathic accuracy. Furthermore, the Myers (2009) study shows the smallest effect size for sex, so it seems clear that merely having female targets to read is not enough to make women more accurate. Beyond that, the sex effect size for the "easy to read" *male* target in Laurent et al. (unpublished data) is one of the largest.

Granted, some of the effect sizes for the studies reviewed here are effectively zero—and none constitute more than a "small" effect size (Cohen, 1988). However, all of them are on the same side of zero. If we were forced to bet on who would score higher on a test of empathic accuracy, a man or a woman, we would bet on the woman (but we would keep the wager small!). Thus, the dichotomous variable of perceiver sex seems to recall the findings we reported earlier using continuous variables associated with gender roles (or at least with the feminine gender role): Although differences are sometimes found that support a female advantage, these differences are small, they are not consistently found and there are not easy answers to the question of *when* they will be found.

### **Conclusions and future directions**

In many ways, studies trying to link empathic accuracy with feminine gender roles and being female mirror the study of empathic accuracy generally, where individual difference and personality variables have been disappointing, inconsistent, and generally unreliable predictors of empathic accuracy (Ickes et al., 1990; Ickes, Buysse, et al., 2000). If anything, the factors of

being feminine and female have fared slightly better than other variables: At least there is more than one set of significant results supporting each, and there are no results showing significant (or even suggestive) effects in the *opposite* direction. However, what *is* there is far from a compelling portrait of “women’s intuition” in this particular domain.

Effects that take the form of motivation by sex interactions paint a clearer picture. Three studies we have reviewed in this chapter suggest that challenging women’s ability to be empathically accurate improves their performance. The same challenges, even when they are “tailored” for men (e.g., the condition in Thomas and Maio’s 2008 Study 2 that consisted of a message that men are not very good at interpersonal sensitivity) do not appear to have parallel effects on men.

The intriguing motivation by sex interactions that have been found thus far open up interesting directions for future research. Researchers might start by trying to answer the question *why* challenging women’s empathic accuracy abilities appears to make them better. It may be that asking women to evaluate their performance not only reminds them of the stereotype that women are better at empathy-related tasks, but may also remind them of the stereotype that men are bad at these tasks. If this is the case, then women's superior performance may be explained by “stereotype lift,” or the performance boost observed following awareness that an outgroup is negatively stereotyped (Walton & Cohen, 2003). Because stereotype lift seems to have its largest effect on individuals whose performance is most closely linked to their social identities (Marx & Stapel, 2006), this may be especially true of more communal women. In contrast, perhaps it is the women who have the greatest concerns about how well they are meeting female gender roles who are most sensitive to the challenge, which would predict a different and perhaps even opposite pattern of results, with un-feminine women or women who

experience uncertainty or anxiety about how well they fit their gender role responding more to challenges.

Another important question to pursue is whether challenges to women's empathic accuracy ability work to improve accuracy via motivating them to try harder, as we have provisionally hypothesized in this chapter. Motivation has been presumed to be behind better empathic accuracy in several studies, but we know of only one study showing a direct link between self-reported motivation and empathic accuracy (Laurent et al., unpublished data) and confess to knowing of a few less successful attempts languishing in file drawers. Of course, self report may not be the best measure of motivation, and future studies would benefit from getting different shots of motivation from other angles (see Hall, this volume, for the challenges of measuring motivation in this particular domain). Other ways of measuring or manipulating motivation would well serve the study of empathic accuracy, as motivation has emerged as one of the most fruitful predictors of empathic accuracy generally, not just for women whose abilities have been challenged. The next step would be then to show *what* specifically is increased by greater motivation—whether motivation works by increasing people's attributions of cognitive complexity as demonstrated by Thomas and Maio (2008), or whether it works via additional mechanisms as well.

As for further research examining the small and slippery main effects of sex and communion, it will require fortitude. Theoretical models that could compellingly explain complex interactions need to catch up with the motley collection of unhypothesized sets of results that have accumulated. A tolerance for very small amounts of variance explained may help, as will the use of more sensitive statistical models such as Hierarchical Linear Modeling (HLM). The latter is not only more powerful because it increases the degrees of freedom in

analyses, but also because it allows a precise allocation of variance in empathic accuracy, partitioning it into the variance attributable due to between-perceiver characteristics, as well as the variance attributable to particular targets or to the particular thoughts to be inferred. Earlier studies have already suggested that how easy a target is to read accounts for the lion's share of the variance in empathic accuracy (Gesn & Ickes, 1999; Ickes, Buysse, et al., 2000; Marangoni, Garcia, Ickes & Teng, 1995; Thomas & Maio, 2008); HLM allows this variance to be divvied up among thoughts *within* a target.

Taking a wider view, it may be important to reconsider how the studies we have reviewed in this chapter fit into our larger understanding of empathic accuracy—not just what predicts it, but how it plays a role in the success of social interactions (Gleason, Jensen-Campbell & Ickes, 2009; Hodges et al., 2010; Myers & Hodges, 2009). Several recent studies (e.g., Hodges, 2005; Hodges et al., 2010; Myers, 2009) have somewhat surprisingly shown that empathic accuracy is neither predicted by independent variables that are theoretically and empirically linked to positive social interactions (such as similar life experiences or reciprocal disclosure), nor is empathic accuracy correlated with outcomes associated with positive social interactions, such as rapport and feeling understood. One working idea is that empathic accuracy is indeed a skill within the human arsenal of tools used in interpersonal interactions, but one that plays a backup or reserve role to more basic (and less effortful!) strategies such as the use of schemas and stereotypes in inferring others' thoughts (Gesn & Ickes, 1999; Locher, 2009; Myers & Hodges, 2009) or that is used mostly when these other strategies encounter conflicting or inconsistent input. If women (and other groups considered to have superior social skills) were more successful at using these other strategies, or if perceptions of other people are more important than actual accuracy in determining the success of social interactions, it could simultaneously

explain why women are credited with special intuition *and* why any advantage women hold over men when it comes specifically to inferring others' thoughts is so precarious.

## References

- Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42*, 155-162.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2<sup>nd</sup> ed.). New York: Academic Press.
- Costanzo, M., & Archer, D. (1989). Interpreting the expressive behavior of others: The Interpersonal Perception Task. *Journal of Nonverbal Behavior, 13*, 225-245.
- Crowne, D. P. & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349-354.
- Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *Catalog of Selected Documents in Psychology, 10*, 85.
- Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multi-dimensional approach. *Journal of Personality and Social Psychology, 44*, 113-126.
- DePaulo, B. M. (1992). Nonverbal behavior and self presentation. *Psychological Bulletin, 111*, 203-243.
- Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin, 94*, 100-131.
- Fernandez-Duque, D., Hodges, S. D., Baird, J. A., & Black, S. E. (in press). Empathy in the behavioral variant of frontotemporal dementia. *Journal of Clinical and Experimental Neuropsychology*.
- Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review, 102*, 652-670.

- Funder, D. C. (1999). *Personality judgment: A realistic approach to person perception*. New York: Academic Press.
- 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.
- Gleason, K. A., Jensen-Campbell, L. A., & Ickes, W. (2009). The role of empathic accuracy in adolescents' peer relations and adjustment. *Personality and Social Psychology Bulletin*, 35, 997-1011.
- Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin*, 85, 845-857.
- Hall, J. A. (1984). *Nonverbal sex differences: Communication accuracy and expressive style*. Baltimore: The Johns Hopkins University Press.
- Hall, J.A., & Schmid Mast, M. (2007). Sources of accuracy in the empathic accuracy paradigm. *Emotion*, 7, 438-446.
- Hall, J. A., & Schmid Mast, M. (2008). Are women always more interpersonally sensitive than men? Impact of content domain and motivation. *Personality and Social Psychology Bulletin*, 34, 144-155.
- Helgeson, V. S. (1994). Relation of agency and communion to wellbeing: Evidence and potential explanations. *Psychological Bulletin*, 116, 412-428.
- Hodges, S. D. (2004). You just can't understand: Shared experience and parental divorce. In J. A. Simpson (Chair), The perils of accuracy and inaccuracy in relationships and social interactions. Symposium conducted at the meeting of the Society of Personality and Social Psychology, Austin, TX.

- Hodges, S.D. (2005). Is how much you understand me in your head or mine? In B.F. Malle & S.D. Hodges (Eds), *Other minds: How humans bridge the divide between self and others*. (pp. 298-309). New York: Guilford Press.
- Hodges, S. D. (2009). Components of empathic understanding: The experience of being understood. In J. Zaki (Chair), *Sharing and understanding the emotions of others: Relationships between different facets of empathy*. Symposium conducted at the 2009 meeting of the Society of Personality and Social Psychology, Tampa, FL.
- Hodges, S. D., Kiel, K. J., Kramer, A. D. I., Veach, D., & Villanueva, B. R. (2010). Giving birth to empathy: The effects of similar experience on empathic accuracy, empathic concern, and perceived empathy. *Personality and Social Psychology Bulletin*, *36*, 398-409.
- Hodges, S. D., & Myers, M. W. (2007). Empathy. In R. F. Baumeister and K. D. Vohs (Eds.), *Encyclopedia of social psychology* (pp. 296-298). Thousand Oaks, CA: Sage.
- Hurd, K., & Noller, P. (1988). Decoding deception: A look at the process. *Journal of Nonverbal Behavior*, *12*, 217-233.
- Ickes, W. (1993). Empathic accuracy. *Journal of Personality*, *61*, 587-610.
- Ickes, W. (2001). **Measuring empathic accuracy**. In J. A. Hall & F. J. Bernieri, (Eds.) *Interpersonal sensitivity: Theory and measurement*. (pp. 219-241). Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Ickes, W. (2003). *Everyday mind reading: Understanding what other people think and feel*. Amherst, NY: Prometheus Books.
- Ickes, W., Buysse, A., Pham, H., Rivers, K., Erickson, J.R., Hancock, M., Kelleher, J., & Gesn, P.R. (2000). On the difficulty of distinguishing “good” and “poor” perceivers: A social relations analysis of empathic accuracy data. *Personal Relationships*, *7*, 219-234.

- Ickes, W., Gesn, P.R., & Graham, T. (2000). Gender differences in empathic accuracy: Differential ability of 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. J. K., & 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. (2009a). Gender and empathic accuracy: The role of communion in reading minds. *Sex Roles*, 60, 387-398.
- Laurent, S. M., & Hodges, S. D. (2009b). Trying hard to guess your mind: Motivation and empathic accuracy. Poster presented at the 2009 meeting of the Society of Personality and Social Psychology, Tampa, FL.
- Laurent, S. M., Hodges, S. D., & Lewis, K. L. (unpublished data). *Modeling mind-reading: Using multi-level modeling to investigate target readability, motivation, and personality in empathic accuracy*. University of Oregon.
- Lewis, K. L. (2008). *Empathic accuracy and nonverbal decoding: Related or distinct constructs?* Unpublished master's thesis, University of Oregon.
- Locher, B. (2009). *Empathic accuracy and the use of stereotypes in inferring the thoughts and feelings of others*. Unpublished Honors College thesis, University of Oregon.
- 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.
- Marx, D. M., & Stapel, D. A. (2006). Understanding stereotype lift: On the role of the social self. *Social Cognition*, 24, 776-791.

- McClure, E. B. (2000). A meta-analytic review of sex differences in facial expression processing and their development in infants, children, and adolescents. *Psychological Bulletin*, *125*, 424-453.
- Myers, J. E. (2009). *Predicting rapport in dyads: Mattering over mind-reading*. Unpublished master's thesis, University of Oregon.
- Myers, M. W., & Hodges, S. D. (2009). Making it up and making do: Simulation, imagination and empathic accuracy. In K. Markman, W. Klein, & J. Suhr (Eds.), *The handbook of imagination and mental simulation* (pp. 281-294). New York: Psychology Press.
- Nowicki, S., & Duke, M. P. (1994). Individual differences in the nonverbal communication of affect: The Diagnostic Analysis of Nonverbal Accuracy Scale. *Journal of Nonverbal Behavior*, *18*, 9-35.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, *40*, 598-609.
- Pickett, C. L., Gardner, W. L., & Knowles, M. (2004). Getting a cue: The need to belong and enhanced sensitivity to social cues. *Personality and Social Psychology Bulletin*, *30*, 1095-1107.
- Rosenthal, R., & DePaulo, B. M. (1979). Sex differences in eavesdropping on nonverbal cues. *Journal of Personality and Social Psychology*, *37*, 273-285.
- Spence, J. T., & Helmreich, R. L. (1978). *Masculinity and femininity: Their psychological dimensions, correlates, and antecedents*. Austin: University of Texas Press.
- 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.

- Spencer, S. J., Steele, C. M., & Quinn, D. M. (1999). Stereotype threat and women's math performance. *Journal of Experimental Social Psychology, 35*, 4-28.
- Thomas, G., & Maio, G. R. (2008). Man, I feel like a woman: When and how gender-role motivation helps mind-reading. *Journal of Personality and Social Psychology, 95*, 1165-1179.
- Wagner, H. L., MacDonald, C. J., & Manstead, A. S. (1986). Communication of individual emotions by spontaneous facial expressions. *Journal of Personality and Social Psychology, 50*, 737-743.
- Walton, G. M., & Cohen, G. L. (2003). Stereotype lift. *Journal of Experimental Social Psychology, 39*, 456-467.
- Zaki, J., Bolger, N., & Ochsner, K. (2008). It takes two: The interpersonal nature of empathic accuracy. *Psychological Science, 19*, 399-404.

### Acknowledgements

The authors would like to thank Alison Lenton for helpful discussions that influenced this chapter; Andreas Gloeckner, Stephan Dickert, and the “Intuitive Experts” Research Unit of the Max Planck Institute for Research on Collective Goods for providing insight about some of the findings contained in this chapter; Sara Loitz for her timely and wise comments on an earlier draft of the chapter; and the Department of Psychology at the University of Edinburgh for providing a sabbatical refuge to the first author while the chapter was being written.