

# Gender Roles and Empathic Accuracy: The Role of Communion in Reading Minds

Sean M. Laurent · Sara D. Hodges

© Springer Science + Business Media, LLC 2008

**Abstract** Although empathic accuracy is considered a stable skill, few individual difference measures consistently predict performance on Ickes' (e.g., 2001) empathic accuracy measure. Because past work has shown that women are more empathically accurate than men when female gender roles are made salient before an empathic accuracy task, we hypothesized that self-reported communion and related variables might predict empathic accuracy. Participants (194 undergraduates) from a northwestern U.S. university completed an empathic accuracy task and self-report measures of communion and empathy. Communion and empathic concern predicted greater empathic accuracy, but only after controlling for socially desirable responding. The role of communion in empathic inference is discussed, along with the need to include measures of social desirability when examining correlates of empathic accuracy.

**Keywords** Gender roles · Communion · Empathic accuracy · Empathy · Social desirability

## Introduction

Regardless of where a person is from—geographically, politically, or socially—one of the most important skills a human can possess is the ability to correctly know what other people are thinking and feeling. Accurately guessing others' thoughts and feelings facilitates our understanding of other people's intentions, aids us in discerning truth from

untruth, and helps us anticipate the needs or desires of those with whom we interact. Poor guessing, on the other hand, can spark disappointment or disaster, causing misunderstandings or outright conflicts. Given the importance of this ability, it comes as no surprise that researchers want to understand who is good at it, why they are good at it, and under what circumstances they are good at it.

It is important then to uncover elements that might help predict empathic accuracy, a measure of interpersonal accuracy that concerns inferring the thoughts and feelings of a target person. One promising variable that has been previously investigated in relation to interpersonal accuracy more broadly is communion (e.g., Davis and Kraus 1997), an individual difference variable related to gender and gender roles. Similarly, self-reported empathy has also been investigated as a predictor of interpersonal accuracy. Theoretically, it seems as if these variables should also predict empathic accuracy. Therefore, the purpose of the present study is to investigate further whether communion and related constructs such as self-reported empathy can predict empathic accuracy. Below, we describe in more detail the empathic accuracy paradigm and the theoretical rationale for the study. It is important to note that in this study we are investigating empathic accuracy specifically, not interpersonal sensitivity more generally, and our review of the literature reflects this. In addition, our research was conducted in the United States, and thus, generalizing our results to other countries or cultures where gender roles may differ should be done cautiously.

## Empathic Accuracy

The naturalistic act of guessing others' thoughts and feelings is so commonplace that it has been called "everyday mind reading" (Ickes 1997, p. 2). To investigate

---

S. M. Laurent (✉) · S. D. Hodges  
Department of Psychology, University of Oregon,  
1227 University of Oregon,  
Eugene, OR, USA, 97403-1227  
e-mail: slaurent@uoregon.edu

people's skill at this everyday task, i.e. their *empathic accuracy*, Ickes and his colleagues (Ickes 2001; Ickes et al. 1990a; Marangoni et al. 1995) have invented a clever methodology. In Ickes and colleagues' paradigm, a perceiver makes inferences about a target's specific thoughts and feelings which can later be compared to the target's reported actual thoughts and feelings in order to compute an index of the perceiver's accuracy. According to available evidence based on this paradigm, there are reliable individual differences in perceivers' skill levels, indexed by high within-perceiver reliability across multiple targets (Gesn and Ickes 1999; Marangoni et al. 1995). Given this reliability, one might expect empathic accuracy to be predictable by other individual difference variables, ideally getting to the source of why some people are more accurate than others.

#### Individual Differences Predicting Empathic Accuracy

Although several studies (e.g., Buysse and Ickes 1999; Klein and Hodges 2001; Stinson and Ickes 1992) have experimentally demonstrated how situations can moderate perceiver accuracy, few have documented stable individual differences related to empathic accuracy. For example, Ickes et al. (1990b) found that a measure of self-monitoring was positively related to empathic accuracy. To our knowledge, however, this has not been replicated. In the same study, Ickes et al. found that perceivers' grade point averages were positively related to empathic accuracy, although a subsequent study (Ickes, Hancock, Graham, Gesn, & Mortimer, as cited in Ickes et al. 2000a) failed to replicate this effect.

In a different study, Ickes et al. (2000a) found that participant gender interacted with verbal intelligence to predict empathic accuracy scores. For men, there was a positive and marginally significant relation between a measure of verbal intelligence and empathic accuracy, but for women, the same measure was negatively related and non-significant. Ickes et al. concluded that verbal intelligence might be the best candidate predictor of empathic accuracy. The fact that verbal intelligence related to empathic accuracy differently for women and for men in their study suggests that it may be important to test the moderating effect of gender on any candidate predictors of empathic accuracy.

Although few individual difference predictors of empathic accuracy have been found, one variable that has consistently predicted empathic accuracy under certain specified conditions is the gender of the perceiver (e.g., Ickes et al. 2000b). To pursue this connection further, we now consider the issue of gender differences in empathic accuracy, in order to make the case that good candidate predictors of empathic accuracy might be individual difference variables that are also related to gender.

#### Gender Differences in Empathic Accuracy

Gender differences in empathic accuracy have received mixed support, but a convincing argument has been put forth explaining why gender differences sometimes occur and other times do not. In a review of ten studies of empathic accuracy, Graham and Ickes (1997) noted that in seven early studies, no differences between women and men were found, but in three later studies, women were more accurate. They proceeded to search their methods for an explanation. The only difference they found was in the form that participants used to record their inferences about a target person's thoughts and feelings. In the later three studies, the form asked perceivers to rate how accurate they felt their thought/feeling inferences had been. This differed from the earlier form that asked perceivers to judge the emotional valence of the target's thoughts or feelings. Next, Ickes et al. (2000b) conducted a meta-analysis that included the original ten studies and an additional five studies that used one of the two versions of the inference form. They concluded that this very small change was indeed the reason why women were sometimes more accurate than men. To explain the finding, they turned to the discussion by Eisenberg and Lennon (1983) of gender differences in empathy, reasoning that when women were made aware that a component of empathy was being assessed, it activated stereotype-related prescriptions about women's empathy, thus motivating women—but not men—to try harder at the task.

In a related line of research, Klein and Hodges (2001) experimentally manipulated whether participants were exposed to a task designed to make a stereotyped female gender role salient before they performed an empathic accuracy task. First, participants watched a target video without making any inferences, and then some participants completed a questionnaire asking them how sympathetic they felt toward the target while others completed a filler questionnaire that had nothing to do with empathy. Participants then watched the video again, making inferences about the target's thoughts and feelings. Klein and Hodges found that when perceivers filled out the sympathy questionnaire before completing the empathic accuracy paradigm, women were significantly more accurate than men. When perceivers did not fill out the sympathy questionnaire before the empathic accuracy task, men and women did equally well. In line with Ickes et al. (2000b), Klein and Hodges came to the conclusion that it was not a difference in ability between women and men but a difference in motivation related to gender roles that was driving these results. As a further demonstration that motivation was one key to empathic accuracy, Klein and Hodges showed that even in the presence of the sympathy questionnaire, the gap between men's and women's

empathic accuracy performance closed up when an additional motivator for accuracy—in the form of financial rewards—was provided.

Taken together, these findings suggest that gender roles play a part in increasing empathic accuracy. Although this conclusion has so far been limited to an explanation of why women are sometimes more accurate than men, the possibility exists that self-reported gender roles or variables closely related to these roles, such as self-reported empathy (e.g., Eisenberg and Lennon 1983), might also relate to empathic accuracy.

### Gender Role Adherence and Empathic Accuracy

Gender roles are sets of gender-linked traits and behaviors ascribed to and enacted by women or men to greater or lesser degrees. Although measurement of these roles has taken many forms, two common measures of communion (part of a female gender role) and agency (part of a male gender role) are the Bem Sex Role Inventory (Bem 1974) and the Personal Attributes Questionnaire (Spence et al. 1974). Although some roles are more strongly prescribed for women and others for men (Prentice and Carranza 2002), adherence to these roles by women and men is flexible (e.g., Foushee et al. 1979). In fact, although women report higher levels of femininity and are often stereotyped as being more expressive, communal, empathic, and interdependently oriented (i.e., stereotypically feminine traits) than men, there is considerable overlap between the degree to which individuals, regardless of whether they are men or women, agree that these gender roles describe them (e.g., Schenk and Heinsch 1986). Similarly, although men report higher levels of masculinity and are thought to be more instrumental, agentic, and independently oriented (i.e., stereotypically masculine traits), men and women both report varying degrees of adherence to these roles too.

As discussed above, differential motivation to be empathically accurate has been linked to gender stereotypes. Specifically, women have been found to be more accurate than men when a stereotypically feminine gender role is made salient prior to the completion of an empathic accuracy task. This finding has two important implications. First, if women's motivation to be accurate results from activation of a feminine gender role, then individual differences in self-reported communion—which might be conceptualized as the chronic activation of a feminine gender role—may also be related to empathic accuracy. Second, because of overlap between women and men in the extent to which these roles are reported, their relationship to empathic accuracy should be similar for women and men.

We argue, then, that gender roles such as communion, which shows consistent gender differences but also shows overlap between women and men, might be good candidate

predictors of empathic accuracy. In addition, although empathic accuracy is distinguishable and often independent from the emotional components of empathy (e.g., see Hodges and Biswas-Diener 2007), we predict that empathic concern—the component of empathy characterized by feelings of compassion for another person—will also predict empathic accuracy. We are not the first to test these ideas. For example, traits such as femininity (also called communion; e.g., Helgeson 1993) and self-reported empathy—as measured by the Interpersonal Reactivity Index of Davis (1980), which includes subscales for empathic concern, perspective-taking, personal distress, and fantasy—have also been generally unrelated to empathic accuracy in previous research, or related in unexpected or unreliable ways (cf. Ickes et al. 1990b; Klein and Hodges 2001; Myers and Hodges *in press*).

One explanation offered for these unexpected null results is that a lack of adequate perceiver variance in empathic accuracy scores masks potential relationships between empathic accuracy and individual difference variables (Ickes et al. 2000a). Another explanation suggests that important characteristics of the *target* must be accounted for in order to find predicted relationships (Zaki et al. 2008). Although both of these explanations are reasonable, it is also possible that some unmeasured variable that is related to self-reported communion and empathy is obscuring the relationship between empathic accuracy and empathically relevant traits, and if the influence of this variable were removed, the expected relationships would emerge.

Social desirability is one such variable that is often included in research designs in order to control for systematic response biases that reflect impression management concerns or self-deception (e.g., Paulhus and Reid 1991). To our knowledge, social desirability has not been included in previous attempts to predict empathic accuracy from self-reported communion and empathy. The research described below, however, suggests that it should be.

### The Influence of Social Desirability

Social desirability, often measured with instruments such as the Marlowe–Crowne Social Desirability scale (Crowne and Marlowe 1960; see also Paulhus and Reid 1991), refers to a tendency to respond to questionnaire measures in a way that reflects a positive self-description. In previous research, social desirability has been related to femininity (i.e., communion)—in fact, the impression management scale in the Balanced Inventory of Desirable Responding by Paulhus (1984, 1991) has also been labeled the Communion Management scale (Paulhus 2002), and correlates positively with a measure of communal bias (i.e., the endorsement of values such as obedience; Lönnqvist et al.

2007). Other researchers have also found links between communion and socially desirable responding in both adults (e.g., Lee 1982; Taylor 1981) and children (Hall and Halberstadt 1980).

The importance of including measures of social desirability when investigating the link between gender roles and other variables has also been noted. For example, in research relating gender roles to self-esteem, Marsh et al. (1987, 1989) included measures of social desirability in order to control for possible effects of response bias. They concluded that social desirability was an important factor that should be taken into account, and also found that social desirability was consistently related to communion.

Social desirability has been positively related to Davis's (1980) self-reported empathy scales as well (e.g., Constantine 2000; Miville et al. 1999; Paulhus and Reid 1991), with desirable responding typically positively correlated with greater empathic concern and perspective-taking. Because showing empathy for others is generally considered desirable (e.g., Hodges and Biswas-Diener 2007), including measures of social desirability in studies investigating the link between empathy and other variables makes considerable sense, although only a limited number of previous studies have done so (e.g., when studying empathy in adolescent sex offenders; Curwen 2003).

### The Present Study

Because gender differences in empathic accuracy have been linked to gender roles, we propose that, in part, women are more empathically accurate than men when a feminine gender role is made salient simply because women more often than men identify with and adhere more strongly to a communal orientation. We therefore believe that greater communion (and related constructs) will predict greater empathic accuracy, even when a feminine gender role has not explicitly been made salient, and that these variables should predict empathic accuracy equally well for women and men. We also believe that this relation will be clearly evident only when controlling for the influence of socially desirable responding. Furthermore, because gender role adherence is in large part based on the different social roles that men and women are expected to fill, it is likely that the relationship of social desirability to self-reported gender roles will differ by participant gender.

In summary, we hypothesize that 1) greater communion as measured by the BSRI and the PAQ will predict greater empathic accuracy, that this relationship will only emerge when taking into account the influence of social desirability, and that this relationship will not differ for women and men, and 2) that greater self-reported empathic concern will also predict greater empathic accuracy for women and men, but again, only after controlling for the influence of social

desirability. In addition to our primary hypotheses, we will also explore whether agency (an aspect of a masculine gender role that is measured by the BSRI and PAQ agency scales), perspective-taking, personal distress, and fantasy (i.e., elements of self-reported empathy measured by the IRI other than empathic concern) will be related to empathic accuracy. Multiple regression analyses will be used to test all hypotheses relating to the prediction of empathic accuracy—this will allow for explicit tests of whether coefficients differ for women and men. Furthermore, if no interactions with gender emerge, as we hypothesize, then multiple regression analyses using women's and men's combined data will provide the greatest statistical power to test our primary hypotheses.

## Method

### Participants

Participants were 194 students (125 women) from the University of Oregon who received partial credit for a course requirement in exchange for their participation. One subject's data were dropped because she did not complete the questionnaires. Participants' ages ranged from 18–54 years ( $M=19.8$ ). The sample reflected the general racial and ethnic composition of students at the university: 75% percent of the sample self-identified as White, Caucasian or European–American, 11.9% as Asian or Asian–American, 3.6% as Latino/a, and 2.1% as Black or African–American. Two participants (1%) listed their ethnicity as “other” or “mixed,” and 6.7% did not respond to the question about ethnic identity.

### Video Stimulus

The video stimulus in the current study was used previously by Klein and Hodges (2001). Following the procedure outlined in Klein and Hodges, we used a recording of a female university student who was videotaped while she discussed a recent academic setback (after receiving her scores from the Graduate Record Exam, she found that her math score was below a minimum needed for entry into a graduate program to which she wanted to apply). After being interviewed and videotaped, the target woman in the video watched the video herself and stopped it at points where she remembered having had a specific thought or feeling, and wrote down the content of each thought or feeling she remembered. There were four instances where the target recalled a specific thought or feeling during the video, which lasted approximately 5 min. For the current study, the original video was transferred to digital video disk.

## Procedure and Measures

Participants completed the empathic accuracy task either alone or in groups of two or three. The experimenter (the same male in all cases) instructed participants that they would be watching a video of a woman discussing a recent academic setback. (In addition, participants were given 4 min at the beginning of this study to think about and write down as many things about themselves that they considered feminine or masculine, or were asked to write about a control topic. This writing task did not affect empathic accuracy or interact with any other variables to predict empathic accuracy, and regression models that controlled for condition were nearly identical to those presented here with one exception that is mentioned in the results below.) Next, the experimenter played the video stimulus all the way through. Using Ickes' empathic accuracy paradigm (e.g., Ickes 2001), when the video finished, the experimenter informed participants that they would watch the same video again, but this time it would be stopped at certain points. Participants were told that they should try to guess what the person in the video was thinking or feeling at those points. After making sure that participants understood the task, the experimenter then played the video again and paused it at each point where the target had reported a thought or feeling. Participants were verbally instructed to write down their inferences of what the target was thinking or feeling at these points on a response form provided to them.

Following the empathic accuracy assessment, the experimenter collected each participant's response form, seated participants away from each other, and gave them a packet of paper-and-pencil measures. The packet included a standard demographics form, the Personal Attributes Questionnaire (PAQ; Spence et al. 1974), the Bem Sex Role Inventory (BSRI; Bem 1974), the Interpersonal Reactivity Index (Davis 1980), and the Marlowe–Crowne Social Desirability Scale (Crowne and Marlowe 1960).

The Personal Attributes Questionnaire (PAQ) contains two eight-item subscales that measure communion and agency, socially desirable gender-linked personality traits related to femininity and masculinity, respectively (see Spence and Helmreich 1978 and Spence et al. 1974 for the full scales, including information on reliability and validity; see also Spence 1984, 1993, and Helgeson 1994, for information on the validity of the PAQ as a measure of communion and agency). Items from both subscales are measured on a five-point scale that is anchored by a pair of contradictory characteristics (e.g., *very rough* vs. *very gentle* or *very passive* vs. *very active*). These subscales were originally labeled femininity and masculinity, but Spence (1991, 1993; Spence and Buckner 2000) and others (e.g., Helgeson 1993) have concluded that although the

subscales of the PAQ (and the BSRI) do measure gender-related traits, they do not measure *masculinity* and *femininity* in any global sense. We agree with this conclusion, and will refer to the subscales as PAQ communion and PAQ agency to reflect this.

The Bem Sex Role Inventory (BSRI) has two twenty-item subscales for measuring femininity (F scale) and masculinity (M scale). For the full scales and information on reliability and validity, see Bem (1974; see also Spence 1991 and Helgeson 1994, for information on the validity of the BSRI as a measure of communion and agency). Although these scales are labeled femininity and masculinity, there is debate as to whether the BSRI actually measures these constructs, as discussed above. Instead, it has been demonstrated that the M and F subscales measure essentially the same dimensions of communion and agency that are measured by the PAQ (e.g., Spence 1991). For purposes of the current research, we will refer to the subscales as BSRI communion and BSRI agency. Items are measured on a seven-point scale from 1 (*never or almost never true*) to 7 (*almost always true*). Sample items are “warm” (communion scale) and “independent” (agency scale).

The Interpersonal Reactivity Index (IRI) is a self-report empathy scale (see Davis 1980, for the full scale; see also Davis 1983, for information on the reliability and validity of the IRI as a measure of individual differences in empathy). It consists of four components labeled Perspective-Taking (e.g., “I try to look at everybody's side of a disagreement before I make a decision”), Fantasy (e.g., “I daydream and fantasize, with some regularity, about things that might happen to me”), Empathic Concern (e.g., “I often have tender, concerned feelings for people less fortunate than me”), and Personal Distress (e.g., “I tend to lose control during emergencies”). All items are measured on a 5-point scale from “does not describe me very well” to “describes me very well.”

The Marlowe–Crowne Social Desirability scale (MCSDS) is a widely used (Paulhus and Reid 1991), 33-item measure of socially desirable responding (the full scale can be found in Crowne and Marlowe 1960, and information on the reliability and validity of this scale as a combined measure of impression management and self-deception can be found in Paulhus 1984, 2002). All items are rated as true or false, and a point is assigned every time a “true” item (e.g., “I have never intensely disliked anyone.”) or “false” item (e.g., “I can remember ‘playing sick’ to get out of something.”) is endorsed that indicates an unrealistic but socially desirable response. Higher scores indicate higher desirable responding.

For all scales (i.e., for the MCSDS and all subscales from the PAQ, BSRI, and the IRI), scores were created by taking the mean of participants' responses to the individual

items. Information on the reliability (Cronbach's  $\alpha$ ) of all scales can be found in Table 1. With the exception of the IRI *Personal Distress* scale ( $\alpha=.69$ ), the reliability of all scales was adequate and comparable to reliabilities reported in the literature (e.g., BSRI, Bem 1974; IRI, Davis 1980; MCSDS, Paulhus 1984; PAQ, Spence and Helmreich 1978). Furthermore, as low reliability tends to attenuate correlations among variables, any correlation or regression coefficient based on scales with low reliability would tend to *underestimate* rather than overestimate the reported relationship (John and Benet-Martínez 2000).

## Results

### Empathic Accuracy Coding

Using the coding scheme devised by Ickes et al. (1990b), participants' thought and feeling inferences were coded for accuracy by six independent raters on a three-alternative scale that ranged from zero (essentially different content from the target's reported thoughts or feelings) to 1 (somewhat similar content) to 2 (essentially the same content). Interrater reliability for the coders' ratings (Cronbach's  $\alpha$ ) was .86. Ratings were then averaged across coders and across inferences for each participant, giving a single index of accuracy between 0 and 2, which was then divided by 2 to create an index of accuracy ranging from .00 (no accuracy) to 1.00 (maximum accuracy).

### Data Analysis Method

Multiple regression was used to test all hypotheses, and separate models were tested for each scale. First, a model

using gender, BSRI communion, and BSRI agency as simultaneous predictors of empathic accuracy was computed. Following this, a model using gender, PAQ communion and PAQ agency was computed, and last, a model using gender and the subscales from the IRI (i.e., empathic concern, perspective-taking, fantasy, personal distress) was computed. Tolerances for all variables in each model were also examined as a way of checking for possible multicollinearity. These analyses revealed that multicollinearity was not an issue (i.e., for all variables, tolerances were above .81, and for most variables, tolerance was above .95). Gender was dummy coded (women coded as "0" and men coded as "1"). Two-way interaction terms were then computed in order to test whether our predictor coefficients differed by gender (see Aiken and West 1991). Because none of these interactions (e.g., communion $\times$ gender) were significant (including higher order interactions), they are not discussed further.

Simultaneously analyzing the data from women and men with multiple regression analyses offers several significant advantages over other types of analyses. For example, unlike conducting separate analyses for women and men, this parsimonious and recommended approach (cf. Aiken and West 1991; Jaccard and Turrisi 2003; Pedhazur 1997) allows for explicit significance tests of whether regression coefficients differ for women and men, which is a question of key interest here. Furthermore, this approach conserves statistical power by including the data for women and men together in each analysis.

### Preliminary and Exploratory Analyses

Before conducting our primary analyses, we ran a MANOVA to explore whether our candidate predictor variables

**Table 1** Descriptive statistics of scales predicting empathic accuracy.

	Women mean (SD)	Men mean (SD)	$\alpha$	Correlations with canonical variate
Empathic accuracy	.46 (.15)*	.40 (.16)*	.86	-.42
BSRI communion	5.08 (.55)*	4.70 (.67)*	.81	-.74
BSRI agency	4.84 (.65)*	5.06 (.70)*	.85	.35
PAQ communion	4.12 (.47)*	3.87 (.54)*	.75	-.55
PAQ agency	3.50 (.55)**	3.66 (.60)**	.70	.28
IRI perspective-taking	3.53 (.72)	3.48 (.62)	.80	-.03
IRI fantasy	3.41 (.74)*	3.12 (.72)*	.78	-.46
IRI empathic concern	3.95 (.55)*	3.64 (.55)*	.75	-.64
IRI personal distress	2.59 (.56)*	2.38 (.56)*	.69	-.40
Social desirability	15.20 (.14)	15.49 (.15)	.71	.09

BSRI Bem Sex Role Inventory, PAQ Personal Attributes Questionnaire, IRI Interpersonal Reactivity Index

The empathic accuracy measure is has endpoints of 0 (no accuracy) and 1 (maximum accuracy). The BSRI measures have endpoints of 1 and 7. The PAQ and IRI measures have endpoints at 1 and 5. The Social Desirability measure has endpoints of 0 and 33. Correlations with the canonical variate are from a MANOVA analysis that used gender as a single independent variable and all measures as a combined dependent variable. Means within the same row marked with \* differ at  $p<.05$  and means marked with \*\* differ at  $p<.10$  (using univariate  $F$  tests).

and empathic accuracy combined would differentiate women from men. Results indicated that women and men differed on the single canonical variate, Wilks'  $\Lambda = .84$ ,  $F(10, 175) = 3.20$ ,  $p = .001$ . Further examination of the correlations of each variable with the canonical variate showed strong loadings for all variables except for social desirability and perspective-taking, and follow-up univariate analyses ( $F$  tests) confirmed that women differed from men on all variables except for these two. Table 1 gives raw means, standard deviations, and reliability coefficients of all predictor variables and empathic accuracy, and also includes correlations of each variable with the single canonical variate that emerged in the MANOVA analysis. As expected, these variables—which have shown gender differences in the past—differentiated men and women well, with women's higher scores on communion and empathic concern (as well as two other empathy subscales from the IRI: personal distress and fantasy) contributing most strongly to the canonical variate, and men's higher scores on agency also contributing (see Table 1). Somewhat surprisingly however, women's slightly higher average empathic accuracy also contributed somewhat to the combined dependent variable, and a univariate follow-up analysis confirmed that women were more empathically accurate than men overall. Still, this is not entirely inconsistent with prior research that sometimes finds gender differences in empathic accuracy—albeit usually only when the link between empathy and the female gender role is primed.

We also examined correlations among our predictor variables separately for women and men (Table 2). As expected, social desirability was correlated with all of our candidate predictor variables (i.e., BSRI and PAQ communion, and empathic concern). Although social desirability did not contribute much to the canonical variate (i.e., the correlation of social desirability with the gender differentiating variate was .09), the correlations of social desirability with the communion variables and empathic concern

differed by gender (using  $r$  to  $Z$  transformations), with larger correlations for men than women in all cases, all  $ps < .05$ . To confirm that socially desirable responding was more closely associated with communion and empathic concern for men than women, we followed up with separate exploratory factor analyses for women and men, using scale scores (BSRI and PAQ communion and agency, empathic concern, and social desirability) rather than items as the units of analysis. For women and men both, two factors emerged, but the rotated factor loadings differed by gender. For women, the two communion scales and empathic concern loaded strongly on the first factor (all loadings above .81), while the two agency scales loaded strongly on a second factor (all loadings above .90). Social desirability, however, loaded weakly across both factors (.35 and .22, respectively). For men, the first factor showed strong loadings from the communion scales (both loadings above .86), and empathic concern (.72), but also from social desirability (.73). The second factor again had high loadings from both agency scales (above .94), but unlike for women, had a very low loading for social desirability (−.02).

These analyses suggest that, particularly for men, a large part (i.e., greater than 25% of shared variance) of self-reported communion reflected a tendency to respond in a socially desirable way. To remove this strong influence of social desirability, we regressed our candidate predictors on social desirability (separately for women and men) and saved the residuals from these analyses. These residual scores represent measures of communion and empathic concern that are not biased by a tendency to respond in a socially desirable manner (i.e., they are uncorrelated with social desirability), and were the measures we used in subsequent analyses to predict empathic accuracy. Similarly, because of significant correlations of social desirability with perspective-taking and personal distress, we regressed these variables on social desirability and saved the residuals, using

**Table 2** Correlations of predictor variables.

		1.	2.	3.	4.	5.	6.	7.	8.	9.
1.	BSRI Communion	–	.04	.74*	.00	.33*	.12	.56*	.00	.52*
2.	BSRI Agency	−.03	–	.08	.79*	.04	−.01	.05	−.42*	−.02
3.	PAQ Communion	.69*	.04	–	.23	.39*	−.01	.46*	−.04	.51*
4.	PAQ Agency	−.10	.71*	.03	–	.01	−.14	.12	−.46*	.09
5.	IRI Perspective Taking	.14	.01	.24*	−.06	–	.19	.36*	−.18	.42*
6.	IRI Fantasy	.18*	.03	.18*	−.01	−.11	–	.28*	−.10	−.07
7.	IRI Empathic Concern	.53*	−.10	.60*	−.13	.32*	.29*	–	−.15	.34*
8.	IRI Personal Distress	.21*	−.37*	.07	−.35*	−.14	.14	.14	–	−.15
9.	Social Desirability	.09	.02	.22*	.15	.26*	−.08	.19*	−.28*	–

Correlations for women are reported below the diagonal. Correlations for men are reported above the diagonal.

BSRI Bem Sex Role Inventory, PAQ Personal Attributes Questionnaire, IRI Interpersonal Reactivity Index

\* $p \leq .05$

them as predictors in our subsequent analyses. On the other hand, residual scores were not used for those variables that were uncorrelated with social desirability.

An alternative approach was to include our measure of social desirability in our primary regression analyses in order to control for its influence. These analyses predicted empathic accuracy from gender, communion, and agency, while controlling for social desirability and the interaction of social desirability with gender (in effect, controlling for the differential correlations for women and men of our predictors with social desirability). The results using this alternate strategy were entirely consistent with those reported here (i.e., the coefficients for all hypothesized predictor variables were exactly the same as those reported here), but we think the results we report below are more straightforward and easier to interpret.

Supporting our hypothesis concerning the necessity of controlling for social desirability, in all regression analyses that did not control for social desirability, the coefficients for our candidate predictor variables were typically smaller and were not significant. Accordingly, in all analyses reported below, we use measures that are free of the influence of social desirability (i.e., the residuals) to predict empathic accuracy. For the sake of simplicity, we refer to these residuals by the names of the scales they were derived from (e.g., BSRI communion).

#### Primary Regression Analyses—Empathic Accuracy on Communion and Agency

Our primary interest (Hypothesis 1) was to test whether communion would predict empathic accuracy, and to see whether this effect would replicate across different scales measuring this construct. Specifically, we hypothesized that the BSRI and PAQ communion scales would predict empathic accuracy, that this relationship would not differ by gender, and that the relationship would only emerge when controlling for social desirability. Our first analysis examined the relation of the BSRI communion and agency subscales to empathic accuracy. We were also interested in whether the gender difference we found would remain even after controlling for the influence of communion.

Our results showed that, as predicted, BSRI communion significantly predicted empathic accuracy,  $\beta=.18$ ,  $p=.01$ , whereas BSRI agency was unrelated to empathic accuracy. We also found that even when communion was accounted for, women were significantly more empathically accurate than men,  $\beta=-.17$ ,  $p=.02$ . This finding suggests that beyond the greater communion that women typically report—which predicts empathic accuracy—women showed an advantage over men in inferring a target's thoughts and feelings, even when a female gender role had not been made explicitly salient. Table 3 gives standardized

and unstandardized regression coefficients along with standard errors and significance tests for the BSRI analyses (as well as those for the PAQ and the IRI; see below). As noted above, none of the two-way interactions with gender were significant, showing that communion predicted empathic accuracy similarly for women and men. Also noted above, when social desirability was not accounted for in the analyses, the coefficient for communion was smaller and not significant.

Next, in a separate model, we regressed empathic accuracy on PAQ communion and agency, essentially replicating our first analysis with a different scale. Again, we found that communion ( $\beta=.13$ ,  $p=.068$ ) but not agency predicted empathic accuracy—albeit marginally (see Table 3). Although the initial writing activity described above in the method section had no main effect on empathic accuracy, when our analysis controlled for this manipulation, the coefficient for PAQ communion increased ( $\beta=.16$ ) and was significant,  $p=.05$ . In this analysis, we also found again that coefficients did not differ for women and men, that the prediction was not significant without controlling for social desirability, and that women were more accurate than men,  $\beta=-.17$ ,  $p=.02$ .

#### Primary Regression Analyses: Self-Reported Empathy

Our next analysis focused on the relation, if any, of empathic accuracy to self-reported empathy. Hypothesis 2 was that empathic concern—a communal facet of empathy (e.g., Skoe

**Table 3** Summary of regression analyses predicting empathic accuracy from BSRI and PAQ communion and agency and IRI self-report empathy.

	<i>B</i>	<i>SE<sub>B</sub></i>	$\beta$	<i>p</i> ≤
BSRI				
Gender	-.06	.02	-.17	.02
Communion	.05	.02	.18	.01
Agency	-.00	.02	-.02	.75
PAQ				
Gender	-.05	.02	-.17	.02
Communion	.04	.02	.13	.07
Agency	-.02	.02	-.07	.36
IRI				
Gender	-.05	.02	-.16	.03
Perspective-taking	-.04	.02	-.16	.04
Fantasy	.01	.02	.07	.38
Empathic concern	.05	.02	.17	.03
Personal distress	.01	.02	.03	.71

BSRI communion, PAQ communion, and IRI Perspective Taking, Empathic Concern, and Personal Distress are residuals that are uncorrelated with social desirability

BSRI Bem Sex Role Inventory, PAQ Personal Attributes Questionnaire, IRI Interpersonal Reactivity Index



et al. 2002)—would predict empathic accuracy similarly for women and men, but only when the influence of social desirability was controlled. Conceptually, this was another internal replication with a measure that represented a related but different facet of communion. We also thought we might reproduce an unusual finding from previous research, i.e. that perspective-taking would be negatively related to empathic accuracy (Hodges and Kiel unpublished).

As predicted, we found a significant positive relationship between empathic accuracy and empathic concern,  $\beta=.17$ ,  $p=.03$ . We also found, as in previous research, that perspective-taking negatively predicted empathic accuracy,  $\beta=-.16$ ,  $p=.04$ . Again, gender also predicted empathic accuracy, with women more accurate than men,  $\beta=-.16$ ,  $p=.03$ .

## Discussion

Prior research has shown that making a female gender role salient before an empathic accuracy task increases women's empathic accuracy relative to men's. Based on this finding, we hypothesized that gender role adherence in the form of self-reported communion and empathic concern (which has also been related to communion; e.g., Skoe et al. 2002) would predict empathic accuracy. We also hypothesized that this association might only emerge after controlling for the influence of socially desirable responding. Consistent with this hypothesis, we found that, first, social desirability was correlated with both of our included measures of communion (BSRI; Bem 1974; PAQ 1974) and with empathic concern (Davis 1980), and that this correlation was especially strong for men. This is particularly interesting because, as in past research, women rated themselves as more communal (e.g., Spence and Buckner 2000) and empathic (e.g., Eisenberg and Lennon 1983) than men, but did not differ in social desirability. Speculatively, this higher correlation for men might have reflected a desire for some men to appear more communal than they are, particularly since they had just finished a task that involved trying to get in touch with a target's thoughts and feelings. It is also possible, however, that because communion is a feminine trait, some men higher in communion tried to respond in a less communal fashion, in order reassert their masculinity after doing a feminine task. Either of these possibilities may have contributed to a higher correlation between communion and social desirability for men, and further research might help clarify this.

Second, we found that only after statistically removing the influence of socially desirable responding from communion and empathic concern did these variables significantly predict greater empathic accuracy. The same analyses, not controlling for social desirability, had smaller,

non-significant coefficients. This is an important distinction, because earlier research that has examined the relationship of communion and empathic concern to empathic accuracy (Ickes et al. 1990b; Klein and Hodges 2001) found no effect, but also did not report controlling for social desirability. The present data suggest that in order to find the relationship between gender-related personality traits and empathic accuracy—which was the primary focus of this research—it may be important to control for relationships of these variables with social desirability. In other words, simply reporting oneself as communal in order to manage one's impression (as some men appear to do) or because of self-deception does not help one to be more empathically accurate. But for women and men both, the extent to which one's self-reported communion reflects actual differences in communal orientation and does not reflect a bias in responding, does matter for empathic accuracy.

Another interesting finding that came out of this research was the higher empathic accuracy for women compared to men. Previous research (Ickes et al. 2000b; Klein and Hodges 2001) has shown that typically, women are only more empathically accurate after a female gender role has been made salient. Still, even when this is not the case, women are sometimes more accurate than men. For example, Klein and Hodges (2001) found that women often had a slight, although non-significant, advantage over men even when a feminine gender role had not been made explicitly salient, and even when the empathic accuracy task was framed as a "male" task. The magnitude of the non-significant differences reported in Klein and Hodges (e.g.,  $d=.39$  in the cognitive instruction condition, p. 724) is comparable to the significant gender difference found here ( $d=.38$ ), and the significance of the effect in the current study is probably attributable to the greater statistical power in the current research. Also, in other research that did not attempt to prime a feminine gender role and used a measure of mind-reading accuracy similar to Ickes' (2001) paradigm, women were found to be more accurate than men in the domain of intimate relationships (Thomas and Fletcher 2003).

In line with these findings, it is also prudent to acknowledge that gender differences are not always found on all measures of interpersonal sensitivity, even when a feminine gender role *has* been made explicitly salient. For example, Koenig and Eagly (2005) showed that when participants were primed to believe that a test of interpersonal sensitivity was a feminine (versus neutral) task, women's accuracy was not affected. Similarly, Horgan and Smith (2006) found that women's interpersonal sensitivity was not affected by a manipulation that primed a feminine (versus a neutral or masculine) role. Of course, neither of these studies used the Ickes' paradigm, which underscores the possibility that different measures of

interpersonal sensitivity are capturing different aspects of what it means to be interpersonally accurate (e.g., Lewis and Hodges unpublished), and highlights the fact that in the present research we are focusing only on empathic accuracy using the Ickes paradigm, and are not trying to generalize our results to other measures of interpersonal sensitivity.

Still, with several studies showing that women are not typically more empathically accurate than men unless a feminine gender role has been made salient, it becomes important to speculate why the gender difference occurred here. One possible explanation may lie in the nature of the video stimulus used, which was the same in Klein and Hodges (2001) and in the present research. As noted by other researchers, success in the empathic accuracy paradigm relies more heavily on verbal rather than nonverbal information, although nonverbal information still plays a role in accuracy (Gesn and Ickes 1999; Hall and Schmid Mast 2007). Other research in interpersonal sensitivity, however, shows that women have a distinct advantage in decoding nonverbal information (e.g., Hall 1978; McClure 2000). It is possible, then, that the female target in the current research was particularly expressive as she discussed her academic setback. If this particular video contained more nonverbal cues than other videos, it could explain why women's accuracy was greater. Further research using this same target video may help to clarify this.

One last finding of note is that similar to other research (Hodges and Kiel unpublished), perspective-taking was *negatively* related to empathic accuracy for women and men both. This counter-intuitive finding is particularly interesting when one considers that perspective-taking is positively correlated with empathic concern and communion (both BSRI and PAQ), which were all related to *greater* empathic accuracy. Similar to communion and empathic concern, perspective-taking was also positively correlated with social desirability. One possibility is that those people who are higher in perspective-taking are actually projecting their own states onto the target (e.g., Nickerson 1999), and imagining how they might think or feel in the target's situation rather than imagining how the target feels. To the extent that they are dissimilar to the target, this might explain the negative correlation. Regardless of the cause, this pattern of intercorrelations suggests the possible presence of some unknown suppressor variable that would explain the negative relationship of perspective-taking to empathic accuracy, and future investigations might fruitfully explore this possibility.

## Conclusion

As noted earlier, the motivation to succeed at empathic accuracy tasks has been linked to gender roles that typically differ for women and men. However, the distributions of women's and men's adherence to these roles overlap in the

extent to which women and men describe themselves as connected to or concerned for others. Using this as a starting point, we sought to show that higher communion might be conceptualized as the chronic activation of at least one part of a feminine gender role, and that women and men who adhere to this communal orientation might have an advantage in an empathic accuracy task. We further hypothesized that this relationship might be evident only when a tendency to respond in a socially desirable way is controlled for.

Our hypotheses were generally confirmed—and internally replicated. Across two scales measuring communion, as well as a scale measuring empathic concern, we showed that these variables do predict empathic accuracy—in effect, replicating our finding within the same study. Of importance, though, is that these variables only predicted empathic accuracy when socially desirable responding was controlled for. This finding provides some of the first evidence that it may not simply be gender itself that is predictive of empathic accuracy when a female gender role is made salient: The communal role that women generally adhere to more than men also plays some part in success at this task.

Understanding other minds is not easy. Like working on a jigsaw puzzle where the photograph is blurry and some of the pieces may be missing, empathic inference is a challenging and cognitively difficult task (Hodges and Wegner 1997). As such, it may turn out that understanding what consistently predicts empathic accuracy is also a challenging task. The current study provides one piece of the puzzle of how humans understand other minds, by highlighting the positive relation between communion and empathic accuracy.

**Acknowledgements** The authors would like to thank Jennifer Freyd, Adam Kramer, Karyn Lewis, John Myers, Mike Myers, Rebecca Neel, and Gerard Saucier for their insightful comments on an earlier draft of this paper. We would also like to thank Tim Mathews for serving as the experimenter and Carlie Ashcraft, Mary D'Anna, John Meyers, Rebecca Neel, Carissa Sharp, and Liana Vega for their help in coding.

## References

- Aiken, L. S., & West, S. G. (1991). *Multiple regression: testing and interpreting interactions*. Newbury Park, CA: Sage.
- Bem, S. L. (1974). The measurement of psychological androgyny. *Journal of Consulting and Clinical Psychology, 42*, 155–162.
- Buyse, A., & Ickes, W. (1999). Topic-relevant cognition and empathic accuracy in laboratory discussions of safer sex. *Psychology & Health, 14*, 351–366.
- Constantine, M. G. (2000). Social desirability attitudes, sex, and affective and cognitive empathy as predictors of self-reported multicultural counseling competence. *The Counseling Psychologist, 28*, 857–872.
- Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349–354.
- Curwen, T. (2003). The importance of offense characteristics, victimization history, hostility, and social desirability in assessing

- empathy of male adolescent sex offenders. *Sexual Abuse*, 15, 347–364.
- 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 multidimensional approach. *Journal of Personality and Social Psychology*, 44, 113–126.
- Davis, M. H., & Kraus, L. A. (1997). Personality and empathic accuracy. In W. Ickes (Ed.), *Empathic accuracy* (pp. 144–168). New York: Guilford.
- Eisenberg, N., & Lennon, R. (1983). Sex differences in empathy and related capacities. *Psychological Bulletin*, 94, 100–131.
- Foushee, H. C., Helmreich, R. L., & Spence, J. T. (1979). Implicit theories of masculinity and femininity: dualistic or bipolar? *Psychology of Women Quarterly*, 3, 259–269.
- 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.
- Graham, T., & Ickes, W. (1997). When women's intuition isn't greater than men's. In W. Ickes (Ed.), *Empathic accuracy* (pp. 117–143). New York: Guilford.
- Hall, J. A. (1978). Gender effects in decoding nonverbal cues. *Psychological Bulletin*, 85, 845–857.
- Hall, J. A., & Halberstadt, A. G. (1980). Masculinity and femininity in children: development of the children's Personal Attributes Questionnaire. *Developmental Psychology*, 16, 270–280.
- Hall, J. A., & Schmid Mast, M. (2007). Sources of accuracy in the empathic accuracy paradigm. *Emotion (Washington, D.C.)*, 7, 438–446.
- Helgeson, V. S. (1993). Implications of agency and communion for patient and spouse adjustment to a first coronary event. *Added. Journal of Personality and Social Psychology*, 64, 807–816.
- Helgeson, V. S. (1994). Relation of agency and communion to well-being: evidence and potential explanations. *Psychological Bulletin*, 116, 412–428.
- Hodges, S. D., & Biswas-Diener, R. (2007). Balancing the empathy expense account: strategies for regulating empathic response. In T. Farrow, & P. Woodruff (Eds.), *Empathy in mental illness* (pp. 389–407). New York, NY, US: Cambridge University Press.
- Hodges, S. D., & Wegner, D. M. (1997). Automatic and controlled empathy. In W. Ickes (Ed.), *Empathic Accuracy* (pp. 311–339). New York: Guilford.
- Horgan, T. G., & Smith, J. L. (2006). Interpersonal reasons for interpersonal perceptions: gender-incongruent purpose goals and nonverbal judgment accuracy. *Journal of Nonverbal Behavior*, 30, 127–140.
- Ickes, W. (1997). Introduction. In W. Ickes (Ed.), *Empathic Accuracy* (pp. 1–16). New York: Guilford Press.
- Ickes, W. (2001). Measuring empathic accuracy. In J. Hall, & F. Bernieri (Eds.), *Interpersonal sensitivity: theory, measurement and applications* (pp. 219–241). Hillsdale, NJ: Erlbaum.
- Ickes, W., Bissonnette, V., Garcia, S., & Stinson, L. L. (1990a). Implementing and using the dyadic interaction paradigm. In C. Hendrick, & M. S. Clark (Eds.), *Review of personality and social psychology: vol. 11. Research methods in personality and social psychology* (pp. 16–44). Newbury Park, CA: Sage.
- Ickes, W., Stinson, L., Bissonnette, V., & Garcia, S. (1990b). Naturalistic social cognition: empathic accuracy in mixed-sex dyads. *Journal of Personality and Social Psychology*, 59, 730–742.
- Ickes, W., Buysse, A., Pham, H., Rivers, K., Erickson, J. R., Hancock, M., et al. (2000a). 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. (2000b). Gender differences in empathic accuracy: differential ability or differential motivation? *Personal Relationships*, 7, 219–234.
- Jaccard, J., & Turrisi, R. (2003). *Interaction effects in multiple regression (2nd ed.)*. Sage University Papers Series on Quantitative Applications in the Social Sciences, 07-072. Thousand Oaks, CA: Sage.
- John, O. P., & Benet-Martinez, V. (2000). Measurement: reliability, construct validation, and scale construction. In H. T. Reis, & C. M. Judd (Eds.), *Handbook of research methods in personality and social psychology* (pp. 339–369). New York: Cambridge.
- 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.
- Koenig, A. M., & Eagly, A. H. (2005). Stereotype threat in men on a test of social sensitivity. *Sex Roles*, 52, 489–496.
- Lee, A. G. (1982). Psychological androgyny and social desirability. *Journal of Personality Assessment*, 46, 147–152.
- Lönnqvist, J. E., Verkasalo, M., & Bezmenova, I. (2007). Agentic and communal bias in socially desirable responding. *European Journal of Personality*, 21, 853–868.
- 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.
- Marsh, H. W., Antill, J. K., & Cunningham, J. D. (1987). Masculinity, femininity, and androgyny: relations to self-esteem and social desirability. *Journal of Personality*, 55, 661–685.
- Marsh, H. W., Antill, J. K., & Cunningham, J. D. (1989). Masculinity and femininity: a bipolar construct and independent constructs. *Journal of Personality*, 57, 625–663.
- 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*, 126, 424–453.
- Miville, M. L., Gelso, C. J., Pannu, R., Liu, W., Touradji, P., Holloway, P., et al. (1999). Appreciating similarities and valuing differences: the Miville–Guzman universality-diversity scale. *Journal of Counseling Psychology*, 46, 291–307.
- Myers, M. W., & Hodges, S. D. (in press). 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*. New York: Psychology.
- Nickerson, R. S. (1999). How we know—and sometimes misjudge—what others know: imputing one's own knowledge to others. *Psychological Bulletin*, 125, 737–759.
- Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology*, 46, 598–609.
- Paulhus, D. L. (1991). Measurement and control of response bias. In J. P. Robinson, P. Shaver, & L. S. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (pp. 17–59). San Diego, CA: Academic.
- Paulhus, D. L. (2002). Socially desirable responding: the evolution of a construct. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49–69). Mahwah, NJ, US: Erlbaum.
- Paulhus, D. L., & Reid, D. H. (1991). Enhancement and denial in socially desirable responding. *Journal of Personality and Social Psychology*, 60, 307–317.
- Pedhazur, E. J. (1997). *Multiple regression in behavioral research: explanation and prediction, 3<sup>rd</sup> Ed.* London: Wadsworth.
- Prentice, D. A., & Carranza, E. (2002). What women should be, shouldn't be, are allowed to be, and don't have to be: the contents of prescriptive gender stereotypes. *Psychology of Women Quarterly*, 26, 269–281.
- Schenk, J., & Heinsch, R. (1986). Self-descriptions by means of sex-role scales and personality scales: a critical evaluation of recent

- masculinity and femininity scales. *Personality and Individual Differences*, 7, 161–168.
- Skoe, E. E., Cumberland, A., Eisenberg, N., Hansen, K., & Perry, J. (2002). The influences of sex and gender-role identity on moral cognition and prosocial personality traits. *Sex Roles*, 46, 295–309.
- Spence, J. T. (1984). Masculinity, femininity, and gender-related traits: a conceptual analysis and critique of current research. In B. A. Maher, & W. Maher (Eds.), *Progress in experimental research* (vol. Vol. 13, (pp. 2–97)). San Diego, CA: Academic.
- Spence, J. T. (1991). Do the BSRI and PAQ measure the same or different concepts? *Psychology of Women Quarterly*, 15, 141–165.
- Spence, J. T. (1993). Gender-related traits and gender ideology: evidence for a multifactorial theory. *Journal of Personality and Social Psychology*, 64, 624–635.
- Spence, J. T., & Buckner, C. E. (2000). Instrumental and expressive traits, trait stereotypes, and sexist attitudes: what do they signify? *Psychology of Women Quarterly*, 24, 44–62.
- 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 (Ms. No. 617).
- Stinson, L., & Ickes, W. (1992). Empathic accuracy in the interaction of male friends versus male strangers. *Journal of Personality and Social Psychology*, 62, 787–797.
- Taylor, D. (1981). Social desirability and the Bem Sex Role Inventory. *Psychological Reports*, 48, 503–506.
- Thomas, G., & Fletcher, G. J. O. (2003). Mind-reading accuracy in intimate relationships: assessing the relationship of the relationship, the target, and the judge. *Journal of Personality and Social Psychology*, 85, 1079–1094.
- Zaki, J., Bolger, N., & Ochsner, K. (2008). It takes two: the interpersonal nature of empathic accuracy. *Psychological Science*, 19, 399–404.