
Roeland C. Dietvorst
Willem J.M.I. Verbeke
Richard P. Bagozzi
Carolyn Yoon
Marion Smits
Aad van der Lugt*

Final Version: June 23, 2008
* Author note

Roeland C. Dietvorst is researcher at the Institute for Sales and Account Management, Erasmus University, Burgemeester Oudlaan 50, PA 3062, Rotterdam, The Netherlands, phone: 31 (10) 4081308, fax: 31 (10) 4089169, Dietvorst@few.eur.nl. Willem J. M. I. Verbeke is Distinguished Professor in Sales and Account Management and Director of the Institute for Sales and Account Management, Erasmus University, Burgemeester Oudlaan 50, PA 3062, Rotterdam, The Netherlands, phone: 31 (10) 4082812, fax: 31 (10) 4089169, Verbeke@few.eur.nl. Richard P. Bagozzi is Dwight F. Benton Professor of Marketing, Stephen M. Ross School of Business, and Professor of Social and Administrative Sciences, College of Pharmacy, University of Michigan, 701 Tappan Street, W6732, Ann Arbor, MI. 48104-1234, phone: (734) 647-4904, fax: (734) 936-8715, bagozzi@umich.edu. Carolyn Yoon is Associate Professor of Marketing, Stephen M. Ross School of Business, University of Michigan, 701 Tappan Street, E5615, Ann Arbor, MI. 48104-1234, phone: (734) 764-6355, fax: (734) 936-8716, yoonc@umich.edu. Marion Smits MD, is at the Department of Radiology, Erasmus MC-University Medical Centre Rotterdam, Dr. Molewaterplein 40, 3015 Rotterdam, phone: 31 (10) 4634044, marion.smits@erasusmc.nl. Aad van der Lugt MD, is Associate Professor at the Department of Radiology, Erasmus MC-University Medical Centre Rotterdam, Dr. Molewaterplein 40, 3015 Rotterdam, phone: 31 (10) 4634044, a.vanderlugt@erasusmc.nl. The first four authors contributed equally to the project.

The authors wish to thank Caroline Kailin Schraa-Tam, Jos van der Geest, Mijke Huibers, Dido van Kuijk, Ralph Lagendijk, and Bas Binder for their assistance. They also thank Angela Gutchess and Robert Welsh for input on technical aspects of the fMRI data analyses. This project was sponsored by the Institute for Sales and Account Management, Erasmus University, Rotterdam.
The goal of this paper is to develop a new theory driven scale for measuring salesperson’s interpersonal mentalizing skills: which is the ability of salespeople to “read the minds” of customers in the sense of recognizing customer intentionality and processing subtle interpersonal cues, as well as adjusting one’s volitions accordingly. Based upon research on autism and neuroscience, the authors develop a model of brain functioning that differentiates better, from less, skilled interpersonal mentalizers. The convergent, discriminant, concurrent, predictive, and nomological validities of measures of the scale were established by use of four methods in four separate studies: confirmatory factor analysis, structural equation models, multitrait-multimethod matrix procedures, and functional magnetic resonance imaging. The study is one of the first to test the validity of measures of a scale not only in traditional ways but also by adopting procedures from neuroscience.

Key Words: theory of mind, salesforce behavior, fMRI, personality, construct validity, interpersonal mentalizing
“To predict how economic man will behave, we need to know not only that he is rational, but also how he perceives the world -- what alternatives he sees and what consequences he attaches to them” (Simon 1956, p. 271).

“Ostensive [i.e., intentional] communication opens up a wide, wild, inner world of relationships and meanings, where constant gambles are being taken, and won, and lost. People with autism impervious as they are to such gambles, cannot fully participate in such a world. It may fascinate them, or terrify them, but it will not readily admit them as players. In this world the game is played by self-aware selves that have enough flying power to allow them the necessary bird’s eye view” (Frith 2003, p. 216).

**INTRODUCTION**

In his classic article, Bonoma (2006) cautioned that salespersons should realize that “companies don’t buy, people do”, by which he meant that it is important to attune oneself to the minds of buyers, minds that sometimes change rapidly as a consequence of group dynamics within buying centers (e.g., Dawes, Lee, and Dowling 1998). The imperative for salespeople is to immerse themselves into the nuances of the customer’s organization and pay special attention to subtle cues communicated by customers. In this way, salespeople can put themselves into the shoes of the members of the buying center and mentally simulate what customers indicate or say they want, and why they want to buy. Following recent developments in neuroscience, we term such processes, interpersonal mentalizing (Singer and Fehr 2005). More formally, interpersonal mentalizing refers to the activity of inferring another person’s beliefs, desires, risk preferences, intentions, and other mental states or events, processing subtle cues, and adjusting one’s volitions accordingly (e.g., Frith and Frith 2003, p. 80). Interpersonal mentalizing is an automatic or reflexive process that encompasses specialized regions of the brain.

The ability to engage in interpersonal mentalizing and read the mind of the customer can be linked to the adaptive selling concept, which is a deliberative phenomenon (whereas
interpersonal mentalizing is largely an automatic process) and is defined as “the altering of sales behaviors during a customer interaction or across customer interactions based upon perceived information about the nature of the selling situation which enables salespeople to tailor messages to fit individual needs and preferences” (Franke and Park 2006, p. 693; Szymanski 1988; Spiro and Weitz 1990). In a similar vein, Sujan, Weitz, and Kumar (1994) proposed that adaptive selling is analogous to working smarter, which involves planning so as to better determine the suitability of sales behaviors and activities that will be undertaken in upcoming selling encounters. Sujan (1999, p. 18-19), however, proposes that “We need improved measures of salespeople’s ability to ‘read’ their customers…” and suggests that promising avenues for developing constructs that pertain to the perceptiveness of salespeople’s observations are the ability to identify client’s needs or desires at the underlying, rather than a superficial, motive level, as well as the ability to pick up on non-verbal cues.

Several drivers have been proposed to explain why salespeople interact in adaptive ways or work smarter. One example is salespeople’s incremental learning, which results in the accrual of contextual knowledge concerning selling contexts; that is, adaptation here depends in part on knowledge of how one’s behavior shapes and is shaped by one’s interactions, which requires mental preparation and planning and a certain degree of self-efficacy with one’s ability to alter behavior in sales situations (Sujan et al. 1994). Individual differences in personality traits are another driver; a key individual difference in this regard is self monitoring, which reflects the degree to which individuals regulate their self presentation by altering their actions in accordance with situational cues present in an interaction (Spiro and Weitz 1990).

The functioning of the drivers of adaptation in selling interactions rests on assumptions about, and processes going on in, the mind of salespeople. Yet research to date has utilized methods based only on verbal self-reports. Advances in neuroscience have inspired recent research in such related areas as consumer behavior (Yoon et al. 2006; Shiv et al. 2005) and
economics (Camerer, Loewenstein, and Prelec 2005), and suggest that mental processes can, despite their complexity and relative inaccessibility, be studied more directly.

Our goal of this article is to develop a domain-specific theory of mind scale (hence called a salesperson theory of mind, SToM scale) which gauges salespeople’s ability to engage in interactions with customers based on how well they take into account the intentions and other mental states and events of customers. On the basis of scores of salespeople on the SToM scale (described below), we categorize salespeople according to their theory of mind. Then people scoring relatively high or low on the scale were asked to participate in a laboratory experiment in which their brain activity was monitored during a task which involved listening to stories designed to evoke different opportunities for taking the perspective of both customer and salesperson. Our aim is to pinpoint specific brain areas distinguishing high versus low interpersonal mentalizers and to provide a paper and pencil scale and managerial implications. This study to our knowledge is the first in marketing to test the validity of a new scale using insights from neuroscience, along with traditional methods.

To accomplish this goal, we develop some ideas from neuroscience because these insights provide us with a different view on what might make salespeople successful during selling encounters. Then we describe four studies for investigating the role of mentalizing in personal selling. In the first study, our objective is to identify real situations and tasks that require interpersonal mentalizing by actual salespeople. Here, we do not study interpersonal mentalizing, per se, but rather uncover expressed skills thought to be related to interpersonal mentalizing. In other words, we investigate how interpersonal mentalizing is embodied or enacted within specific selling situations (see Zaltman, 1997, for a similar perspective). Thus, we develop a paper and pencil measure that indirectly operationalizes interpersonal mentalizing concepts in a selling context. We call this the ‘salesperson theory of mind, scale’ to stress the context-specific aspects of our measures and differentiate them from a generalized theory of mind (ToM) scale, which we develop and use to test criterion-related validity. In Study 2, we
replicate the findings of Study 1 and further relate the SToM-scale to performance and other variables related to interpersonal mentalizing. In both Studies 1 and 2, we investigate convergent, discriminant, and criterion-related validity; Study 2 also goes further to examine nomological validity of the measures of the SToM-scale by use of structural equation models. Study 3 then collects data by use of the multitrait-multimethod matrix and uses confirmatory factor analysis to test for the convergent and discriminant validity of measures of interpersonal mentalizing. Then in the fourth study, in order to identify the brain areas involved in interpersonal mentalizing, and validate measures of the scale at the neurological level, we use functional Magnetic Resonance Imaging (fMRI) and experimental treatments to compare salespeople identified as high versus low in interpersonal mentalizing skills, as measured by our scale, and to pinpoint specific differences in neural processing. A brief primer on fMRI methodology, as well as technical details specific to Study 4 is provided in Web Appendix 1.

**ESSENTIALS FROM NEUROSCIENCE**

Salespeople interact with customers for the purpose of understanding customer needs and designing and offering a product or service to meet those needs. The goal is to forge an understanding and a contract that potentially meets the interests of both seller and buyer. From the point of view of the firm, this requires that the salesperson understands the customer’s perspective and skillfully navigates negotiations so as to achieve a signed contract. To be effective, salespeople need to comprehend and interpret the customer’s mental states and processes. Scholars characterize the aspect of mentalizing that is critical for effectiveness of a salesperson as follows: “the ability to generate a “decoupled” representation of the beliefs of the customer about the world, “decoupled” in the sense that they are decoupled from the actual state of the world and that they may or may not correspond to reality” (Singer and Fehr 2005, p. 341). We suggest that the salesperson interprets the interpersonal situation, in general, and in particular mentalizes about the customer, through a process of making inferences and conjectures as to the beliefs, desires, intentions, etc., of the customer. Interpersonal mentalizing
is especially needed in such self-interested exchanges as agency contract negotiations (e.g., Bergen, Dutta, and Walker 1992) and the forming of alliances in order to compete more effectively in certain markets (e.g., Morgan and Hunt 1994).

Neuroscience research reveals that interpersonal mentalizing is an automatic, unconscious and effortless process that involves the activations of a network of hard-wired brain areas or modules\(^2\) (to be described below) as a function of social cues emerging from interactions between people in an encounter. In order to introduce the processes underlying interpersonal mentalizing for further discussion, consider the following admonishment by a participant in a recent experiment wherein the participant had his $10 ultimatum rejected by a player in a game:

“I did not earn any money because all the other players are STUPID! How can you reject a positive amount of money and prefer to get zero? They just did not understand the game! You should have stopped the experiment and explained it to them…” (Camerer et al. 2005, p. 47).

Camerer and colleagues note that this particular respondent failed to mentalize effectively about the other party. That is, he failed to realize that many people react to what they perceive as unfair offers by rejecting them, even if by doing so they forego any gain. Such one-sided allocations of attention to cues and formation of dysfunctional categorizations are analogous to reactions to others by people high on Autistic Spectrum Disorders (ASD) (Camerer et al., 2005). People high on ASD tend to respond to social cues during interactions according to rote rules (e.g., by categorizing signals and remembering their meaning according to stereotypes or in literal senses), and as a result they frequently make mistakes in judgment in their interactions (e.g., Eckel and Wilson 2003, p. 247). In other words, in an attempt to read the minds of their interlocuter they use coarse-grained categories (akin to the categories described by Sujan, Sujan, and Bettman 1988). The coarse-grained categories might work for routine situations, but come up short within more complex interpersonal contexts requiring detailed attention to interaction partners, such that flexible, quick, and appropriate reactions can be generated to shape the
conversation eventually to one’s advantage. This seems to be at the heart of Bonoma’s (2006) analysis of the dynamics in buying centers and is consistent with the analysis by Singer and Fehr (2005, p. 343) who argue that mind reading involves the ability to understand the actual motivational state of the interaction partner, motivations that might change rapidly over time and thus require constant reinterpretation.

The way the human brain functions might help to explain why coarse-grained categories dominate judgments by salespeople who seem relatively poor at interpersonal mentalizing (e.g., Camerer et al. 2005). Three functions are of note:

1) **Specialist functions**: people possess specialized brain areas or modules that have evolved through evolution to process different sorts of informational cues such as emotions, intentions, and content related to a specific topic and goals of people with whom they interact (Pinker 1997). When a specialized brain system is triggered by particular cues, processing is rapid, and the task feels relatively effortless to the person engaged. People are generally unaware of the power and sophistication of the processes that enable them to interact with others effectively (Camerer et al. 2005, p. 21).

2) **Parallel modules** that operate in (un)coordinated ways: different regions of the brain operate largely in parallel and at times act in a concerted way, while at other times, they work at odds with each other. The functioning of these regions of the brain can be thought to constitute networks of brain activities. Ramachandran (2004) refers to this as cross-wiring and provides a wide range of examples of such networks. During interpersonal mentalizing, specific brain modules interact in a coordinated way to form a network (discussed more fully below).

3) **Winner takes all**: the brain does not invariably integrate all the signals activated by individual groups of neurons or networks. When two distinct neuronal groups convey different information about the external world, the resulting perceptual judgment often adopts the information from one dominant activated neural group and suppresses or
ignores the information carried by the other weakly activated neural group (Camerer et al. 2005, p. 25).

Consistent with research in neuroscience, we suggest that salespeople low in interpersonal mentalizing will experience weak activation of certain specialized areas in their interpersonal mentalizing brain network (Frith 2003). This involves low integration of the activated information in the brain and utilization of coarse-grained categories when making inferences during social interactions. For those low in interpersonal mentalizing, abstract and coarse-grained categories thus become the key drivers for engaging in conversations (indicative of the “winner takes all” function).

Interpersonal mentalizing is a hardwired brain process that occurs spontaneously and largely unconsciously in social encounters and is centralized in a distinct network of brain regions. Research by neuroscientists shows that the most consistently activated regions with mentalizing tasks are the medial prefrontal cortex (MPFC), located in the middle of the front of the brain, the left and right temporo-parietal junctions (TPJ), located on both sides just above the ears, and the left and right temporal poles (TP), located at the bottom of the temporal lobes (e.g., Frith 2003). In the ideal case, these three areas interact with one another and cooperate as a network to form an overall interpretation of the mental states or events of another person in an interaction (Frith 2003). Table 1 presents a summary of recent findings for studies of mentalizing implicating these three brain regions.

[Table 1 about here]

Based on a growing body of social cognitive neuroscience research, we propose that those who are high (versus low) in interpersonal mentalizing skills will display greater coordinated activation of all three areas implicated in the interpersonal mentalizing network during a mentalizing task. The functioning of this distinct network in the brain provides an
explanation for why some salespeople might be better able than others at taking a bird’s eye view of an interaction and integrating the different pieces of information to their advantage.

**STUDY 1: THE DEVELOPMENT OF THE SToM-SCALE**

To develop the SToM-scale, we first performed a literature search and did a content analysis of research in neuroscience and salesforce behavior to suggest items that could be used in the scale. We then isolated different social situations and interactions where people low in interpersonal mentalizing ability would presumably encounter difficulties (Frith and Frith 2000, p. 337). The literature suggests that people low in interpersonal mentalizing exhibit several characteristics. First, they have difficulties strategically taking the initiative in conversations, which is needed to address needs, cajole, and gauge responses from customers. A second skill lacking in people who are low in interpersonal mentalizing is shaping or providing direction in conversations (Sujan et al. 1994). A third variable differentiating high from low mentalizers is the ability to engage in mutually rewarding interactions. Persons low in mentalizing have difficulties in engaging in tasks that require joint attention and reciprocity; from a salesperson’s perspective, the establishment of joint attention implies that one create a conversational context such that he/she and the customer cognitively elaborate upon the same conversational topics to each others advantage (see Grice’s, 1975, cooperative principle). Finally, persons low in mentalizing have difficulties processing indirect information and hints, as they tend to focus on bare utterances or literal meaning and are less able to grasp and act on the ostensive meanings in communications (Soldow and Thomas 1984).

**Respondents and Procedures**

Sales managers participating in an executive education program were asked to send questionnaires to their salespeople. One hundred and seventy questionnaires were distributed. Respondents were asked to provide a unique code anonymously instead of their name, and then return the completed questionnaire using a self-addressed envelope. As compensation for completing the questionnaire, participants received a gift valued at about $12. For further
motivation, respondents were also informed that their scores on scales would be available to them on the internet site of the Institute for Sales and Account Management, at the university that was sponsoring the project. They were additionally told that following a random selection, they might be invited to participate in an fMRI study of salespeople at the university hospital. Scores on the interpersonal mentalizing-scale were not published on the website prior to the fMRI study, in order to keep the participants unaware of their categorization of being a high or low mentalizer. We received 132 completed questionnaires (for a response rate of 78%). The sample consisted of 90% males and 10% females, their average age was 38.2 years ($SD = 7.4$), and their average sales experience was 9.2 years ($SD = 6.2$). The distribution of gender was representative of the salesforce in the country where the study was conducted.

Results

A total of 33 items was identified by our content analysis. We administered these items to the respondents, along with other measures used to investigate validity (described below). After pruning items due to redundancy and low intercorrelations so as to arrive at a manageable scale, we identified 14 potential items. An exploratory factor analysis using promax rotation and maximum likelihood estimation yielded four factors (explained variance of 48%, KMO = .86). After eliminating one item due to cross loadings, we ended up with 13 items (see Table 2). The four factors are labeled: (1) ability to take initiative in sales and build rapport in conversations ($\alpha = .69$), (2) ability to notice subtle cues during sales encounters ($\alpha = .76$), (3) ability to take a bird’s eye view and supply missing information (i.e., achieve closure) during sales encounters ($\alpha = .66$), and (4) ability to shape/influence interactions with customers in a positive way ($\alpha = .79$).

We correlated the four factors of the SToM-scale with age and sales experience. The findings show that the four factors do not correlate significantly with age ($r = -.087$ to .001) or experience ($r = .016$ to .183). This implies that the dimensions of the SToM-scale reflect more personal dispositions than learned behavior, per se.
Next we scrutinized the validity of the measures of SToM by use of confirmatory factory analysis (CFA) and the partial disaggregate model (Bagozzi and Heatherton 1994; Bagozzi and Edwards 1998). Four analyses were performed: (1) a 4-factor CFA to establish convergent validity of the items for each factor and discriminant validity of items across factors, (2) a second-order CFA to ascertain whether the four factors load satisfactorily on one higher-order factor and thereby constitute more concrete dimensions of an overall abstract construct, (3) a 7-factor CFA to examine criterion-related validity of the measures of the 4-factor SToM-scale with measures of a 3-factor general theory of mind (ToM) scale, and (4) an 8-factor CFA to investigate the discriminant validity of measures of the 4-factor SToM-scale from measures of four factors of important variables studied by contemporary salesforce management researchers (i.e., two dimensions of sales call anxiety, perspective taking ability (an aspect form of empathy), and adaptiveness).

Convergent and discriminant validity of measures of the four dimensions of the SToM-scale. Figure 1, panel a, shows the results for the factor loadings for the CFA model. These loadings are high (.54 to .97), and in conjunction with the satisfactory goodness-of-fit indexes establish that convergent validity has been achieved: goodness-of-fit measures for Study 1 are $\chi^2(14)=17.51, p=.23$, RMSEA=.05, NNFI=.99, CFI=.99, and SRMR=.04. See Web Appendix 2 for definitions of these indexes. Discriminant validity of the measures can be seen in the values of correlations among factors (.43 to .71). These correlations reflect corrections for attenuation due to any unreliability of measures; the raw Pearson product-moment correlations are significantly lower than these correlations. Each of the correlations is significantly less than 1.00 (as indicated by both confidence intervals and $\chi^2$–difference tests) and thus supports the achievement of discriminant validity for the items across the four factors (below we examine discriminant validity of the measures of SToM from measures of other scales).

[Figure 1 about here]
Second-order CFA model of the four dimensions of the SToM-scale. Figure 1, panel b, presents the findings for the second-order CFA of the model. Here we see that the model fits well according to all the goodness-of-fit indexes: $\chi^2(16)=17.85$, $p=.33$, RMSEA=.03, NNFI=.99, CFI=1.00, and SRMR=.04. The second-order and first-order factor loadings are high: second-order loadings range from .61 to .88, first-order loadings from .54 to .97. These results suggest that the four dimensions of the SToM-scale can be organized as distinct, concrete representations of a single, abstract concept of sales theory of mind thinking (i.e., interpersonal mentalizing). Below we show that a certain substructure can be differentiated.

Criterion-related validity of measures of the SToM-scale. To examine the criterion-related validity of the measures of the SToM-scale, we performed a 7-factor CFA of the measures of the SToM-scale and measures of a ToM scale consisting of 3 factors. The measures of ToM comprised 10 items according to criteria proposed by Frith and Frith (2000) and concerning generalized interpersonal mentalizing ability (see Web Appendix 3). For the data in Study 1, we administered these items to the sample of salespeople, factor analyzed them, and found three factors corresponding to three of the four factors for our sales-specific scale (the exploratory factor analyses of the generalized ToM scale are available on request). The three ToM factors capture, respectively, 1) the ability to take initiative in interactions and build rapport (corresponding to our SToM1, building rapport), 2) the ability to process indirect information and hints in conversations (which corresponds to our SToM2 subscale, detecting nonverbal cues), and 3) the ability to cooperate in and coordinate interactions so as to achieve closure (corresponding to our SToM3, taking a bird’s eye view). The literature on interpersonal mentalizing does not address our fourth interpersonal mentalizing-scale factor (shaping the interaction), but we expect all three ToM dimensions to be correlated with SToM4 because such an ability might be expected to be dependent upon the skills summarized by the three ToM dimensions. Overall the CFA model fits well in study 1 according to the goodness-of-fit indexes: $\chi^2(56)=60.91$, $p=.30$, RMSEA=.02, NNFI=.99, CFI=1.00, and SRMR=.05. The relevant
correlations can be found in Table 3 in the entries below the main diagonal and have been highlighted in bold-face type. As hypothesized, ToM1, and SToM1 are highly correlated (.90), ToM2 and SToM2 are highly correlated (.90), and ToM3 correlates moderately highly with STO M3 (.45). Positive correlations between STO M4 and ToM1-ToM3 also occur, as expected (.39, .63, and .18). In sum, the measures of the sales-specific SToM-scale factors achieve criterion-related validity with the measures of the generalized ToM scale factors.

[Table 3 about here]

Discriminant validity of measures of dimensions of STO M-scale, from measures of other scales. We investigated the discriminant validity between measures of the four dimensions of the STO M-scale and measures of three other scales that should be related to the STO M-scale, yet in theory measure different constructs. One of the other scales is a social anxiety scale composed of 12 items and developed by Watson and Friend (1969). We chose social anxiety because it is a common emotion felt by salespeople and should be negatively related to the four dimensions of the STO M-scale. A study by Ramachandran and Oberman (2006) investigating people high on ASD supports our conjecture. Verbeke and Bagozzi (2000) showed the effects of social anxiety in a salesforce but did not examine interpersonal mentalizing as we do herein. The social anxiety scale used herein turned out to have two dimensions which were highly correlated (r = .68).

The second scale we used measured perspective taking (i.e., the ability to put oneself in the place of another), which is one aspect of empathy. We used Davis’ (1980) 6-item scale in this regard and expect the dimensions of the STO M-scale to be positively correlated with perspective taking.

Finally, we used Spiro and Weitz’s (1990) 6-item adaptive selling scale (ADAPTS) and predict that adaptiveness will be positively correlated with the dimensions of the STO M-scale. Spiro and Weitz proposed theoretically that adaptiveness consists of six facets, but the CFA that they ran on data showed that the scale was not unidimensional. They nevertheless treated their scale as a unidimensional scale, which obscures differences amongst facets and violates
psychometric principles of measurement and makes any predictions based on the scale ambiguous. Moreover their 16-item scale contained seven items for facet 6, and only between zero and three items each for facets 1 to 5. As a consequence, we operationalized adaptiveness with 6 of the 7 items for facet 6, which Spiro and Weitz (1990, p.62) defined as “actual use of different approaches in different situations” and measured globally with such general statements as, “I am very flexible in the selling approach I use” (we dropped one item from their facet 6 measures because it was too transparently redundant with one or more of the others). The six adaptiveness measures we used achieved unidimensionality.

Table 4 presents the findings. The model fits well based on the goodness-of-fit indexes: \( \chi^2(76)=95.26, p=.07, \) RMSEA=.04, NNFI=.98, CFI=.99, and SRMR=.05. The four dimensions of SToM correlate negatively with social anxiety (range: -.22 to -.53) and positively with perspective taking (range: .27 to .40) and adaptiveness (range: .46 to .61), as hypothesized. Yet the correlations are significantly less than 1.00 and therefore demonstrate that the measures of SToM are distinct from the measures of social anxiety, perspective taking, and adaptiveness.

[Table 4 about here]

Discussion

The domain-specific SToM scale was shown to consist of four distinct factors, where measures achieved convergent validity within factors and discriminant validity between factors. Further, as our second-order CFA showed, the four SToM factors can be considered as reflecting a single, higher-order abstract representation of SToM with four dimensions. The four SToM dimensions were next shown to achieve criterion-related validity in the sense of systematically correlating with measures of generalized theory of mind skills. Finally, the measures of the four SToM dimensions were shown to be distinct from measures of social anxiety, perspective taking, and adaptiveness. Next we turn to an attempt to replicate these findings in a new sample of salespeople and at the same time demonstrate that the dimensions of SToM relate to performance.
STUDY 2: REPLICATION AND TEST OF PREDICTIVE VALIDITY OF SToM SCALE

The measures of SToM, ToM, social anxiety, perspective taking, and adaptiveness were administered to a new sample of salespeople. In addition, measures of performance were obtained. Finally, in an additional sample of sales managers and their salespersons, the performance measures were validated. Convergent, discriminant, criterion-related, and predictive validity of the measures of SToM were investigated.

Respondents and Procedures

We surveyed a total of 126 salespersons who were students and co-workers of students at an executive education program at a cooperating university. The sample consisted of 91% males, 9% females, average age was 40.0 years (SD = 9.0), and average experience in sales was 12.3 years (SD = 7.8).

The same items used in Study 1 were administered in Study 2. In addition, six items from the Behrman and Perrault (1982) performance scale were used. The six items focus on sales volume, sales quota, selling new products, sales by key accounts, building and maintaining long-term relationships with customers, and profit contributions. Each item asked salespeople to rank themselves on a 10-point scale, where 1 = bottom 10% and 10 = top 10% in sales compared to all salespeople in their company.

Results

Convergent and discriminant validity of measures of the four dimensions of the SToM-scale. Figure 1, panel a, shows the factor loadings for the four SToM factors, where it can be seen that all loadings are high (.70 to .89). The high loadings and satisfactory fit of the CFA model support the achievement of convergent validity: \( \chi^2(14)=17.66, p=.22, \) RMSEA=.04, NNFI=.99, CFI=.99, and SRMR=.02. Discriminant validity is also attained as can be seen in the correlations among factors, which range from .33 to .77 and are all significantly less than 1.00. Again, the discriminant validity of the measures of SToM from measures of other scales will be examined below.
Second-order CFA model of the four dimensions of the SToM-scale. Figure 1, panel b, presents the findings for the second-order CFA model. This model fits well according to all the goodness-of-fit indexes, and the second-order and first-order factor loadings are high: second-order loadings range from .57 to .96, and first-order loadings from .69 to .89: $\chi^2(16)=22.68$, $p=.12$, RMSEA=.056, NNFI=.99, CFI=.99, and SRMR=.04. These results suggest that the four dimensions of the SToM-scale can be organized as distinct, concrete representations of a single, abstract concept of sales theory of mind thinking (i.e., interpersonal mentalizing). Again a particular substructure will be examined below.

Criterion-related validity of measures of the SToM-scale. The findings for the 7-factor CFA of the measures of the SToM-scale and the measures of the ToM scale are shown in Table 2. The model fits well overall: $\chi^2(56)=99.54$, $p=.00$, RMSEA=.066, NNFI=.96, CFI=.98, and SRMR=.05. The entries in the correlations matrix above the diagonal address criterion-related validity. As hypothesized, ToM1, and SToM1 are highly correlated (.97), ToM2 and SToM2 are highly correlated (.87), and ToM3 correlates moderately highly with SToM3 (.48). Positive correlations between SToM4 and ToM1-ToM3 also occur, as predicted (.61, .61 and .42). In sum, the measures of the sales-specific SToM-scale factors achieve criterion-related validity with the measures of the generalized ToM scale factors.

Discriminant validity of measures of dimensions of SToM-scale from measures of other scales. Table 4 presents the results for this test of discriminant validity. The overall fit of the model is good: $\chi^2(56)=75.18$, $p=.05$, RMSEA=.04, NNFI=.98, CFI=.99, and SRMR=.04. The four dimensions of SToM correlate negatively with social anxiety (range: -.23 to -.38), and positively with perspective taking (range: .28 to .33) and adaptiveness (range: .46 to .78), as forecast. Yet the correlations are significantly less than 1.00 and therefore demonstrate that the measures of SToM are distinct from the measures of social anxiety, perspective taking, and adaptiveness.
Predictive validity. Table 3 also presents the correlations between the four SToM factors and anxiety, perspective taking, adaptiveness, and performance factors (see final column in Table). Performance correlated .31, .56, .31 and .48 with the four respective SToM factors, -.35 with anxiety, .25 with perspective taking, and .49 with adaptiveness. This establishes the bivariate predictive validity of the measures of the SToM-scale.

Validation of performance measures. Forty managers at a sales conference were asked to distribute questionnaires to their top and bottom performers. We asked them to give at least two questionnaires each to top and bottom performers and up to 10 if possible. A total of 200 questionnaires were distributed, with 100 to top performers and 100 to bottom performers. We defined top and bottom performers in terms of the ability to achieve high sales, meet quotas, build and maintain relationships with customers, and acquire profitable accounts. The questionnaires contained the same six performance items used in the replication and predictive validity study discussed above and were embedded with many other questions, serving to disguise the purpose of our study. A total of 102 questionnaires was returned: 57 top performers (57% response rate) and 45 bottom performers (45% response rate).

A t-test on the equality of mean performance across top and bottom performers showed that the six performance items do indeed differentiate between high and low performers: $M_{\text{High}} = 7.70$ vs. $M_{\text{Low}} = 6.95$, $t = 4.19$. Thus evidence exists that the Behrman and Perreault (1982) scale items we used are related to actual performance.

Discussion

The SToM measures, which consisted of four distinct dimensions and loaded on one second-order factor, were again found to achieve convergent and discriminant validity in a new sample of salespeople and to achieve criterion-related validity as well. Moreover the measures of the four dimensions of SToM were shown to be distinct from measures of social anxiety, perspective taking, and adaptiveness.
STUDY 3: CONSTRUCT VALIDITY BY THE MULTITRAIT-MULTIMETHOD (MTMM) MATRIX APPROACH AND NOMOLOGICAL VALIDITY

Studies 1 and 2 examined aspects of validity for the SToM-scale but did so by use of a single method. In Study 3, we perform a true construct validity assessment by use of CFA applied to data gathered by two methods: a “does not describe me/describes me completely” scale and a “disagree/agree” scale, both measured with 7-point items. The sample included 132 salespersons obtained using methods similar to that employed in Study 2: average age=38.3 (SD=8.9), and 80% males, 20% females.

The resulting MTMM matrix that we formed consists of two indicators by each method for each SToM factor (“traits”). This yields a 16x16 matrix of correlations. Application of a 6-factor CFA model (4 SToM traits and 2 method factors) showed that the 2 method factors were highly correlated, after correction for attenuation ($\phi=.96, SE=.04$). Therefore we ran a 5-factor CFA model (4 SToM traits and one method factor). This model gave a very good fit to the data: $\chi^2(82)=169.55, p=.00$, RMSEA=.08, NNFI=.97, CFI=.98, and SRMR=.04. Trait variance ranged from .46 to .85 (average=.66), and indeed of the 16 measures, only one (SToM3a) yielded less than 50% trait variance and then only slightly below the .50 standard. Random error variance ranged from .00 to .49 (average=.25), which is low. Method variance ranged from .00 to .44 (average=.09), which is very low; indeed only one of 16 method factor loadings was significant. Overall, the construct validity of the measures of the SToM scale, in terms of convergent validity, is excellent.

Discriminant validity was also achieved for the measures of the SToM scale. The respective correlations of SToM1 with SToM2-SToM4 were .40($SE=.08$), .49($SE=.08$), and .55($SE=.07$). SToM2 correlated .79($SE=.04$) and .76($SE=.04$) with SToM3 and SToM4, respectively, and SToM3 and SToM4 correlated .78($SE=.78$). These correlations, which are corrected for attenuation and therefore are higher than the raw Pearson product-moment
correlations, fall far, and significantly, below 1.00, thereby demonstrating achievement of
discriminant validity for the measures of SToM.

We also investigated predictive validity in a multivariate sense (sometimes also called,
nomological validity) by examining a structural equation model wherein two SToM factors (one
first-order factor for rapport building; and a second-order factor for detecting nonverbal cues,
taking a bird’s eye view, and shaping the interaction) predicted adaptiveness, perspective taking,
social anxiety, and performance, and in turn, adaptiveness, perspective taking, and social anxiety
also predicted performance (see Figure 2). The two SToM factors represent intangible relational
and instrumental aspects of sales theory of mind, respectively. Notice in Tables 3 and 4 that the
correlations among the four SToM factors are consistent with such an interpretation in that
SToM2 - SToM4 correlated highly and uniformly with each other, while SToM1 correlated
moderately with SToM2 - SToM4.

[Figure 2 about here]

Because two methods were used to measure SToM, adaptiveness, perspective taking, and
social anxiety (method 1: “does not describe me at all” to “describes me completely”; method 2:
“strongly disagree” to “strongly agree”), we ran the structural equation model shown in Figure 2
twice, once for each method. For method 1, the overall model fit well: $\chi^2(86)=142.20$, $p=.00$,
RMSEA=.07, NNFI=.97, CFI=.98, and SRMR=.05. It can be seen in Figure 1 that rapport
building influences performance through social anxiety: the greater the rapport building, the
lower the social anxiety, and the greater the performance. The only other effect on performance
is a direct effect from SToM2* (the second-order factor with first-order SToM2- SToM4 factors
loading on it), where the greater the SToM2*, the greater the performance (here the effect only
approaches significance: $\beta=.35$, $t=1.76$). The other results of note include the dependence of
adaptiveness and perspective taking on SToM2*, and the dependence of social anxiety on
SToM1. For method 2, the overall model fit well: $\chi^2(86)=135.66$, $p=.00$, RMSEA=.06,
NNFI=.98, CFI=.98, and SRMR=.04. As shown in Figure 2, rapport building again influences performance indirectly through social anxiety. SToM2* now has a strong direct effect on performance (β=.61, t=2.97). Similar to method 1 findings, we again see that adaptiveness and perspective taking are dependent on SToM2*, and social anxiety is dependent on rapport building.

In sum, SToM processes are the primary drivers of performance. SToM1 (i.e., rapport building) indirectly (through social anxiety) and SToM2* (the second-order factor upon which SToM2- SToM4 load) directly influence performance. Greater rapport building apparently reduces social anxiety, and the less the social anxiety, the greater the performance. The instrumentality of SToM2* functions to affect performance straight on. Adaptiveness (which is measured with overall or summary measures) and perspective taking are dependent on SToM processes but have no effects on performance, over and above the more basic SToM processes.

To both validate the SToM-scale and understand better the bases for interpersonal mentalizing, we turn now to our study of brain processes of salespersons.

**STUDY 4: DO DIFFERENT PATTERNS OF BRAIN ACTIVITY OCCUR BETWEEN HIGH AND LOW INTERPERSONAL MENTALIZING SALESPEOPLE DURING INTERPERSONAL MENTALIZING TASKS?**

To the extent that the SToM-scale measures salespeople’s ability to interpersonally mentalize, we would expect to see different patterns of brain activity between salespersons scoring high and low on interpersonal mentalizing tasks. More specifically, based on the recent research of neuroscientists with autistic and normal persons, we would expect that high as opposed to low scorers on the SToM-scale will display greater activation in the MPFC, TPJ, and TP regions (e.g., Frith and Frith, 2003, Amodio and Frith 2006; and Castelli et al. 2002). Thus, we propose:
Hypothesis: A comparison between the brain activity of salespeople high and low on interpersonal mentalizing ability, during performance of a mentalizing task (relative to performance on a non-mentalizing task), will show greater activations of the MPFC, TPJ, and TP.

Participants

From the sample of 132 salespeople in Study 1, 20 right-hand males were recruited for the fMRI study. Table 5 presents the means, standard deviations, and t-tests for comparison of high versus low scorers on the SToM-scale across various criteria. High versus low participants differed on all four dimensions of SToM. High and low scorers on the SToM-scale did not differ in age or experience, but they did differ as expected on the other scales. The high interpersonal mentalizing (IM) group scored higher on adaptive selling and perspective taking, and lower on social anxiety than the low IM group.

[Table 5 about here]

Method and Materials

The purpose and design protocol for the experiment were approved by the appropriate institutional review board, and all participants gave written informed consent. The stories serving as stimuli were presented auditorily, consistent with the method used by Nieminen-von Wendt et al. (2003).

The fMRI protocol consisted of three experimental conditions: interpersonal mentalizing, process, and unlinked sentences. Participants listened to five stories of each type presented in one of two different counterbalanced orders. Interpersonal mentalizing is the critical condition in which the cognitive task involves the use of “theory of mind” in order to understand why and how the characters in the story interact. The process condition serves as a closely matching control condition, in which the cognitive task involves nearly the same cognitive processes as in the interpersonal mentalizing condition with the exception that the stories do not explicitly require the use of “theory of mind” in order to understand why and how the characters operate.
or interact. Finally, in the unlinked sentences condition, participants listened to a series of sentences that did not form a coherent story. The unlinked sentences condition serves as baseline control condition, in which the cognitive task involves the use of language, and memory. Under each experimental condition, every story was followed by a question that the respondent were asked to answer silently to oneself. The number of words and types of words in the stories were distributed as evenly as possible over the different conditions. The stimuli were presented in the participant’s mother tongue; and an English translation is presented in Web Appendix 4.

Durations of the stories, including the questions, were between 33 and 36 seconds long, and were on average equivalent in terms of time length across the three experimental conditions. Each participant was then given about 6 seconds to think about an answer for each question following the presentation of a story.

A separate group of twenty-five individuals informed about the purpose of the study were asked to evaluate the 15 scenarios. After being given definitions of the stimuli, the 25 people identified each of the 15 scenarios as being either interpersonal mentalizing, process, or unlinked sentence scenarios. They were also asked to describe the scenarios and were counted as giving correct responses if their descriptions were sensible and could be interpreted. Finally they rated on 10-point scales their own confidence in the classification and how clear they thought the scenarios were. The three respective scenarios were correctly classified with 96.8%, 99.2%, and 99.2% accuracy. Answers to the stories were correct for 92.0%, 95.6%, and 100% of interpretations, respectively. The respective average confidence ratings were 8.26 (SD = .94), 8.22 (SD = 1.16), and 9.54 (SD = .72). The average clarity ratings were 8.16 (SD = 1.12) for the interpersonal mentalizing and 7.86 (SD = 1.15) for the process scenarios. Clarity ratings for unlinked sentence scenario were not meaningful given their nature.

Functional Image Analysis

Imaging was conducted using a full-body 3.0 T GE scanner (General Electric, Milwaukee, WI) fitted with an 8-channel receive-only head coil. For the structural imaging, a
high resolution image of the brain was acquired with a 3D T1-weighted inversion recovery fast spoiled gradient recalled echo sequence (echo time (TE)/ repetition time (TR)/inversion time = 2.1/10.4/300 ms, flip angle = 18°, matrix = 416x 256, field of view (FOV) = 25 cm, slice thickness 1.6 mm with 50% overlap). For the functional imaging, a time series of 210 volumes, with 39 Slices in the transverse plane, was obtained using single shot gradient-echo planar imaging (TR = 3000 ms, TE = 30 ms, flip angle = 75°, resolution = 3.5 mm x 3.44 mm x 2.3 mm, and FOV = 22 cm).

During the functional run, a new story was presented every 42 seconds and volume acquisitions were made during the entire 42 second periods. This resulted in 14 whole-brain fMRI volume acquisitions per story, of which the first 13 were used for analysis (the last volume was excluded from analyses, because during this period participants heard 3 beeps, which signaled an inter stimulus interval).

Functional image data were preprocessed and analyzed using Statistical Parametric Mapping (SPM2, Wellcome Department of Cognitive Neurology, London, UK). Linear image realignment, co-registration, non-linear normalization to stereotactic anatomical space (MNI), and spatial smoothing 3-dimensional Gaussian kernel, 8mm full-width at half maximum (FWHM) were performed for each participant using standard statistical parametric mapping methods. A high-pass (cutoff period, 250 sec) frequency filter was applied to the time series.

Based on our hypothesis, greater activations were predicted for high versus low mentalizers in the regions implicated in mentalizing—specifically, the MPFC, TPJ, and TP. We first tested the hypothesis in a conservative way with a random effects group analysis at coordinates defined by previous studies, and second in an explorative way by searching for groups of voxels in which the activity across subjects correlates with the individual SToM measures. Because the predictions were limited to specific anatomical regions, we adopted a region-of-interest (ROI) approach.3 Such an approach tests the contrasts only in those specific regions rather than across the entire brain and, by reducing the degree of correction needed for
multiple comparisons, allows greater sensitivity in detecting effects. Thus small volume
 corrections (SVC; Worsley et al., 1996) were applied to the three a priori regions of interest. The
 MPFC region was defined using MARINA software (Walter et al. 2003) which has predefined
 anatomical regions that can be used as masks. The MPFC mask consisted of the MARINA “left
 and right superior frontal gyrus, medial” regions. For the TPJ and TP regions, a sphere with a
 10 mm radius was used according to the coordinates of previous studies. We used Saxe and
 Wexler (2005) for TPJ (centered at x = 54, y = -54, z = 14; and x = -48, y = -60, z = 21; for right
 and left, respectively) and Fletcher et al. (1995) for TP (centered at x = 44, y = 18, z = -16; and x
 = -44, y = 20, z = -16; for right and left, respectively). Before using SVC, we transformed
 coordinates given by these studies from Talairach space to MNI space (www.mrc-
cbu.cam.ac.uk). The contrasts of interest were then tested in these regions in a second level
 random-effects group analysis.

For the correlational analysis we extracted the mean percent signal change associated
with interpersonal mentalizing compared to the process task, and compared to the unlinked
sentences task, and then examined their correlations with participants’ SToM scores. The sizes
of the ROI’s are larger for the correlational analysis and were created with WFU Pickatlas
software toolbox, by selecting the left and right temporal lobes and the medial prefrontal cortex.
Unless otherwise specified, all results were threshold at $p = .005$ (uncorrected) with a cluster
size greater than $k = 10$. This cluster size was chosen to ensure that all activations were at least 2
contiguous voxels in acquired space.

Based on our hypothesis, we expected that the areas implicated in mentalizing (i.e.,
MPFC, TPJ, and TP) would be more strongly activated in high versus low interpersonal
mentalizers. As a test of our hypothesis, a comparison between high IM and low IM groups was
conducted for the interpersonal mentalizing versus process task and the interpersonal
mentalizing versus unlinked sentences task.
As predicted, the test of interpersonal minus process revealed more activation of the MPFC and the TPJ (see Table 6a for significant interaction effects). However, no difference obtained for the TP. First, in the interpersonal mentalizing minus process task condition, three clusters in the MPFC were significantly more activated in high IM versus low IM group. The high IM compared to the low IM group also displayed greater levels of activation in the MPFC when performing the interpersonal mentalizing minus the unlinked sentences task (see Table 6b). In addition, the high IM group showed greater activation in the right TPJ than the low IM group in the test of interpersonal mentalizing minus process (see Table 6a), as well as the test of interpersonal mentalizing minus unlinked sentences (see Table 6b).

Comparison between high IM and low IM groups, for both the contrast of interpersonal mentalizing versus process, and interpersonal mentalizing versus unlinked sentences, did not yield significant effects in the TP region. Furthermore, in comparisons of the low IM minus the high IM group, none of the areas associated with mentalizing was more active in the low group.

[Table 6 about here]

As further test of our hypothesis, a correlational analysis was performed between the individual SToM scores and the activity during the interpersonal mentalizing versus process task, and the interpersonal mentalizing versus unlinked sentences task. Results revealed three areas in which the activity showed significant positive correlations with SToM scores for the interpersonal mentalizing versus process task: right MPFC ([8 58 20], \(r = .69, p < .005\)), right TPJ ([54 -68 -2], \(r = .69, p < .005\)), and left TPJ ([-66 -28 -4], \(r = .61, p < .005\)) (see Figure 3). Two clusters in the left and right TP showed a similar but nonsignificant trend in terms of correlations with SToM scores for the interpersonal mentalizing versus process task: left TP ([-38, 10, -30], \(r = .52, p < .05\)), right TP ([48, 2, -8], \(r = .45, p < .05\)). Significant positive correlations were also found with SToM scores for the interpersonal mentalizing versus unlinked sentences task in the following regions: left TPJ ([-64 -28 -4], \(r = .67, p < .005\)), left TPJ/STS ([-60 -12 4], \(r = .63, p < .005\)), and right TPJ ([64 -42 6], \(r = .60, p < .01\)). Two small
clusters in the MPFC showed a similar trend in terms of correlations with SToM scores for the interpersonal mentalizing versus unlinked sentences task, but the cluster sizes were smaller than 10 voxels. Furthermore, for both contrasts interpersonal mentalizing versus process, and interpersonal mentalizing versus unlinked sentences none of the regions showed a negative correlation with SToM measures.

[Figure 3 about here]

In sum, we generally find support for our hypothesis: that is, when the neural responses in the interpersonal mentalizing condition were compared with those in the process and unlinked sentences conditions, the MPFC and right TPJ regions were differentially activated in the high as opposed to the low IM group. In addition to the MPFC and right TPJ, a correlational analysis revealed that the left TPJ was also significantly correlated with SToM measures. This effect was however weaker in the TP region for the contrast of interpersonal mentalizing versus process, and the TP was equally activated in high and low for the contrast of interpersonal mentalizing versus unlinked sentences.

Finally, the tests for the contrast interpersonal mentalizing versus process and interpersonal mentalizing versus unlinked sentences yielded somewhat different results, this is mainly due to the noisy nature of the experiment, and because of the different cognitive tasks involved in the process task and unlinked sentences task.

**DISCUSSION**

In the present study, we presented a new theory-based SToM-scale inspired from recent ideas on neuroscientific research on autism. Both psychometric methods and fMRI based research were used to validate this scale. Our research follows a call made by Sujan (1999, p. 18-19) that we need improved measures of salespeople’s ability to “read” their customers. Such scales should seek to tap into salespeople’s ability to identify their client’s needs or desires at the underlying, rather than superficial, motive level. A core conclusion from neuroscience is that our brain consists of modules, modules which are activated by different cues in the environment
which, depending on individual differences, become activated in different intensities. But because salespersons during sales encounters both evoke and process these cues, such activations are coordinated in the brain to form a coherent interpretation (sense-making as it is called) of what occurs during a sales conversation (see Camerer et al., 2005, for an overview). Therefore we developed a brain model that explains sales people’s ability to engage in interpersonal mentalizing.

The research consisted of four studies. In the first study, a paper and pencil measure (the SToM-scale) was developed so as to measure verbal expressions of the degree of interpersonal mentalizing that salespeople exhibit. The results showed that salespeople exhibit different degrees of interpersonal mentalizing which can be represented in four distinct, but related, dimensions, and further, the measures of SToM achieve convergent, discriminant, and criterion-related validity. Moreover, high versus low scorers on the SToM-scale are relatively more adaptive in selling situations, are better able to take the perspective of customers, and show less fear of being evaluated negatively in selling situations. The second study replicated findings of Study 1 and also showed that the four dimensions of SToM relate significantly with performance. The performance measures were then validated on a new sample of high and low performers. Our third study examined the construct validity of measures of SToM by use of the MTMM matrix and CFA and also tested nomological validity. The measures showed high trait variance, low error variance, and very low method variance. Performance was found to be driven largely by SToM: rapport building influenced performance indirectly through social anxiety and the other three dimensions of SToM influenced performance directly. The fourth study was conducted in order to discover whether different functioning of brain regions provide evidence for individual differences in the ability to mentalize interpersonally, and in addition provide evidence that the four dimensions of SToM discriminate between high and low mentalizers. We hypothesized that the high (vs. low) interpersonal mentalizing (IM) group displays relatively greater activations of specific regions of the brain (i.e., the MPFC, TPJ, TP)
that have been consistently reported in the literature to be associated with mentalizing tasks. This hypothesis was largely confirmed: the high (vs. low) IM group showed more activity during the mentalizing task in the MPFC and TPJ regions of the brain, but this effect was much weaker in the TP regions, and was nonexistent when we compared the interpersonal mentalizing task with the unlinked sentences task.

A closer inspection of the data shows that the TP regions were in fact activated highly in both the high and low IM groups for our salesperson participants. To the extent that such activation relates to the formation and use of mental scripts (e.g., Frith and Frith 2003, p. 465), we speculate that both high and low IM salespersons equally use script-based thinking, while differing in the ways described above. It thus appears that only for salespersons high in IM is the entire network consisting of the MPFC, TPJ, and TP fully activated, whereas for persons low in IM, only part of the network, the TP, is fully activated. This interpretation is consistent with our conjectures, made earlier in the article, that people low in IM overly rely on script-based (categorical) thinking, whereas people high in IM integrate such thinking with the use of ostensive cues and interpersonal sensitivity (see Table 1). However as one reviewer pointed out, it is also possible that the high IM group paid more attention to the task. But this also might imply that they are more intrigued by the content of the interpersonal stories as manifest in more thoroughly activated brain processing.

As Camerer et al. (2005) note, the more we know about functional specialization in the brain and how these regions collaborate in performing different tasks, the more these come to substitute for time honored distinctions between categories used to study human behavior; such implications are likely to occur as well for how we study salesforces in the future. In our research, the findings suggest that the capability to interpersonally mentalize reflects the ability to grasp subtle cues intuitively and effortlessly, as well as to go beyond information given in an interaction in order to take a holistic point of view (a “bird’s eye view”, so to speak). This latter ability involves generating coherent, yet conjectural, stories about one’s interaction, which are
revised as the conversation continues. A second important implication is that people differ in their utilization of their mentalizing networks, and these differences have a number of behavioral correlates. A possible explanation of the differences in the pattern of brain activity between the high and low IM groups could be that this reflects a difference in cognitive strategy in computing information about mental states of others. The high IM group displayed an activity pattern in which the MPFC and TPJ play a major role during interpersonal mentalizing; and this might reflect salespersons abilities to be more dynamic, flexible, and adaptive interaction partners. The pattern of brain activity during interpersonal mentalizing suggests that the MPFC and TPJ regions of the brain are significantly less activated in low than high IM salespersons. Because only the TP is fully activated for persons low in IM, whereas the MPFC, TPJ, and TP are activated for persons high in IM, it appears that the pattern of responses for the person low in IM is consistent with the winner takes all metaphor we discussed earlier in the article. Here we suggested that persons low in IM act primarily in rigid ways and/or according to previously learned scripts. Either the person low in IM fails to process social stimuli fully in interpersonal interactions (because the MPFC and TPJ are less active) or else the TP dominates the person’s responses in the sense of overwhelming whatever activity exists in the MPFC and TPJ. The latter is consistent with a winner takes all perspective.

Interpersonal mentalizing also seems to be related to research on mindfulness in the organization science literature. However, researchers in this tradition to date have limited inquiry to the analysis of verbal reports by qualitative methods (e.g., Weick and Sutcliffe 2006). Whereas the information processing perspective emphasizes a two-step process consisting of the categorization of customers followed by implementation of canned policies contingent on the categorization, the mindfulness research has focused on disciplined observation of communication in a holistic sense and interpretation (“sense making”) of communication in the light of the situation-person interface. Weick and Putnam (2006, p. 283) perceptively point out the limitations of the contingency approach as follows: “When people engage in distinction-
making, they begin to realize just how quickly we put our experiences into tidy and unexamined conceptual boxes (Kabat-Zinn, 2002, p. 69), how reluctantly we are to examine those conceptual boxes, and how much is discovered when we examine these boxes”. People low in interpersonal mentalizing seem to be especially prone to categorical thinking in the rigid way that Weick and co-authors characterize it, and at the same time people low in interpersonal mentalizing appear to be relatively insensitive to ostensive cues and nuances in everyday human interaction. People high in interpersonal mentalizing, by contrast, actively engage in on-going sense-making as an interaction ebbs and flows. This occurs apparently in their interpersonal mentalizing brain network which becomes highly activated. Sense-making is manifest in a dynamic, back-and-forth interpretation between (a) the content of what is said and what is not said, including nonverbal communication and inference making of the desires and intentions of the interaction partner and (b) a decoupled, “bird’s eye view”, perspective of how the on-going interaction relates to motivations and expectations of the institutions and people connected to the interaction. Needless to say, people high in interpersonal mentalizing have an advantage which those low in interpersonal mentalizing lack. Our study suggests that the difference occurs in specific brain regions that vary across high and low interpersonal mentalizers and a paper and pencil scale can capture aspects of interpersonal mentalizing in this sense.

Can interactive mindfulness be learned? This is a difficult question to answer at this stage of our knowledge of mentalizing and what is required to cultivate mindfulness. But we believe that through observational learning, role playing, and practice, salespeople can be trained to become better in the practice of mindfulness and perhaps even enhance their mentalizing abilities to a certain extent. The first step in such training is to make people aware that the anxiety they might experience during sales conversations might be a consequence of undeveloped skills in interpersonal mentalizing (see Ramachandran and Oberman 2006) and that anxiety can be reduced to the extent that they develop the discipline to now and then assume a posture of a detached, abstract observer of their own interactions as they occur, so as to provide
the opportunity to interpret the flow of ostensive cues at multiple, specific occasions across an on-going interactions. This might involve, for example, the sub-vocal, posing of questions at different points in time (e.g., “Did the customer’s hint to the effect that she wished we could bundle our offerings mean that her company would order more in the long-run to achieve this short term benefit?”) (Richardsen and Piper 1986).

Moreover, role-playing might stimulate interpersonal mindfulness. Brief simulated interactions could be video-taped, and a skilled, sensitive trainer could analyze the tapes with the salesperson, pointing out what to watch for in ostensive cues and how to respond effectively (e.g., Soldow and Thomas 1984). For example, a customer might show signs of discomfort that could be traced to a mechanical or overly assertive style by the salesperson. Such role-playing could take place, if appropriate, in the presence of other salespeople of the firm, because salespeople will differ in their styles and abilities to mentalize, and shared learning could be facilitated. Considerable development and trial and error may be needed to institute effective role-playing exercises of this sort. Note, too, that the diagnosis, training, and coaching of mindfulness may be best conducted by people identified as particularly skilled in interpersonal mentalizing and practiced in mindfulness. To the extent that mindfulness can be trained, this will have neurological implications as well. In his regard, many researchers (e.g., Hariri and Forbes 2007) have proposed that through life experiences circuits in the brain get wired and rewired – this has come to be called neuro-plasticity. This speculation points to areas for future research.

**SToM and ADAPTS**

How do SToM and ADAPTS differ? SToM is a multidimensional scale designed to measure specific implications of interpersonal mentalizing, whereas ADAPTS has been treated by Sprio and Weitz (1990) as a unidimensional scale to measure a general or overall tendency to practice adaptive selling. Further, SToM refers to largely automatic processes in which salespersons read the minds of customers so to speak and in turn co-create the nature and course of the interaction with customers, whereas adaptive selling is a largely deliberative process in
which salespersons identify customers and selling situations so as to respond according to coarse-grained a priori learned categories, and is thus more of a one-way pattern of communication, albeit informed by earning in an adaptive sense. The first dimension of SToM, rapport building, captures a quality of the relationship between salesperson and customer, whereas the three other dimensions SToM (detecting nonverbal cues, taking a bird’s eye view, and energizing the sales interaction) reflect particular practices or things salespeople do to influence the sales outcome. Although general adaptiveness as measured by ADAPTS might predict sales performance, the items do not identify the specific reasons for, or mechanisms behind, their effects. The SToM scale measures things that might be considered to be the bases for general adaptiveness and thus constitute managerial policy variables for which salespeople can be selected, trained, and coached in order to improve adaptiveness (and influence perspective taking, coping with social anxiety, and performance). Our tests of hypotheses support the effects of the dimensions of SToM on performance, through social anxiety, and directly. Indeed SToM may be considered not only more psychologically fundamental and managerially useful than ADAPTS, but it may even supersede ADAPTS in its effects on performance, as our findings suggest.
Footnotes

1. Other researchers have called such mental activities the application of one’s “Theory of Mind” (ToM) to interpret what is going on in the mind of an interaction partner (Baron-Cohen 1995; Singer and Fehr 2005).

2. We use the terms, brain modules, brain areas, brain regions, brain systems, and neural substrates, interchangeably.

3. Results of the analyses for the TP activations can be found in Web Appendix 5.


Castelli, Fulvia, Chris D. Frith, Francesca Happé, and Utah Frith (2002), “Autism, Asperger Syndrome, and Brain Mechanisms for the Attribution of Mental States to Animated Shapes”, Brain, 125, 1839-49.


TABLE 1. LOCATION AND FUNCTION OF BRAIN REGIONS ASSOCIATED WITH INTERPERSONAL MENTALIZING

<table>
<thead>
<tr>
<th>Regions</th>
<th>Summary of Findings</th>
<th>Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPFC: Medial prefrontal</td>
<td>The MPFC is involved when people reflect on ostensive cues which might signal faking by another person; the MPFC is especially active during interpersonal mentalizing tasks. People in game theory settings take an intentional stand and interpret and predict their opponent’s behavior; this involves MPFC activation.</td>
<td>Grèzes, Frith, and Passingham (2004a; 2004b) as reviewed in Amodio and Frith (2006); Fletcher et al. (1995)</td>
</tr>
<tr>
<td>cortex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TPJ: Temporo-parietal</td>
<td>The TPJ is the most consistently activated area with mentalizing tasks. The right TPJ especially displays selective sensitivity for the onsets of cues about mental states of others and is a key driver in constructing a coherent model of the protagonist’s mind.</td>
<td>Frith and Frith (2001)</td>
</tr>
<tr>
<td>junction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TP: Temporal poles</td>
<td>Left and right TP converge for all sensory modalities. Lesion studies show this region to be particularly associated with social knowledge in the form of scripts.</td>
<td>Frith and Frith (2003)</td>
</tr>
</tbody>
</table>
### TABLE 2. *THE SALES THEORY OF MIND (SToM) SCALE*

**Factor 1: Rapport building**

1. When I am with a customer (e.g., in the elevator before a sales meeting) I can easily kindle a small conversation.
2. I find it difficult to talk to a customer about topics that are not business-related. (R)
3. When at a business meeting or a reception, I can easily start off a conversation on a general topic such as the weather.

**Factor 2: Detecting nonverbal cues**

4. I find it difficult to discern the nonverbal cues of customers during a conversation. (R)
5. At times I realize that I do not pick up the hints in sales conversations; after the meeting colleagues explain to me what happened during the conversations. Only then do I realize what happened during the conversation. (R)
6. During a sales conversation, if customers hint of something, I do take that into consideration as we are speaking together.

**Factor 3: Taking a bird’s eye view**

7. When I realize that someone does not possess the right amount of knowledge in or during a sales conversation, I can easily add some information to bring focus to the conversation, thus making it easier for people to understand what I want to say.
8. When I realize that people do not understand what I’m saying, I put what I want to say in a broader perspective in order to explain what I mean.
9. I always try to understand the industry context in which a customer operates; and by using examples from that context, I add any missing information.

10. Sometimes I summarize for customers what has been said up to that point in the meeting, this make for a smoother conversation!

Factor 4: shaping the interaction

11. I make sure that I positively influence the atmosphere in a sales conversation.

12. I can easily act in ways that gives a sales conversation a positive twist.

13. I can easily make people feel more comfortable during a sales conversation.

R = reverse coded
### TABLE 3. SUMMARY OF FINDINGS FOR STUDIES 1 AND 2: CRITERION-RELATED VALIDITY, THEORY OF MIND (ToM) AND SALES THEORY OF MIND (SToM)

<table>
<thead>
<tr>
<th>Parameter estimates for factor inter-correlation matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Study 1 below, Study 2 above diagonal)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. ToM₁</td>
<td>1.00</td>
<td>.41</td>
<td>.35</td>
<td>.97</td>
<td>.52</td>
<td>.57</td>
<td>.61</td>
</tr>
<tr>
<td>2. ToM₂</td>
<td>.37</td>
<td>1.00</td>
<td>.68</td>
<td>.16</td>
<td>.87</td>
<td>.68</td>
<td>.61</td>
</tr>
<tr>
<td>3. ToM₃</td>
<td>.13</td>
<td>.45</td>
<td>1.00</td>
<td>.22</td>
<td>.62</td>
<td>.48</td>
<td>.42</td>
</tr>
<tr>
<td>4. SToM₁: Rapport building</td>
<td><strong>.90</strong></td>
<td>.52</td>
<td>.08</td>
<td>1.00</td>
<td>.33</td>
<td>.44</td>
<td>.57</td>
</tr>
<tr>
<td>5. SToM₂: Shaping the interaction</td>
<td>.40</td>
<td><strong>.90</strong></td>
<td>.24</td>
<td>.54</td>
<td>1.00</td>
<td>.66</td>
<td>.73</td>
</tr>
<tr>
<td>6. SToM₃: Detecting nonverbal cues</td>
<td>.33</td>
<td>.43</td>
<td><strong>.45</strong></td>
<td>.44</td>
<td>.61</td>
<td>1.00</td>
<td>.75</td>
</tr>
<tr>
<td>7. SToM₄: Taking a bird’s eye view</td>
<td><strong>.39</strong></td>
<td>63</td>
<td>.18</td>
<td>.56</td>
<td>.69</td>
<td>.63</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Key: ToM₁=ability to process information and hints in conversations, ToM₂=ability to take initiative in interactions and build rapport, ToM₃=ability to cooperate in and coordinate interactions so as to achieve closure.
TABLE 4. SUMMARY OF FINDINGS FOR STUDIES 1 AND 2: DISCRIMINANT VALIDITY FOR SALES THEORY OF MIND (SToM), ANXIETY, PERSPECTIVE TAKING, AND ADAPTIVENESS (Study 1 below, Study 2 above diagonal)

<table>
<thead>
<tr>
<th>Parameter estimates for factor inter-correlation matrix</th>
</tr>
</thead>
<tbody>
<tr>
<td>1  2  3  4  5  6  7  8  9</td>
</tr>
<tr>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td>1. Social anxiety1a</td>
</tr>
<tr>
<td>1.00</td>
</tr>
<tr>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td>2. Social anxiety2</td>
</tr>
<tr>
<td>.68</td>
</tr>
<tr>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td>3. Perspective taking</td>
</tr>
<tr>
<td>- .40</td>
</tr>
<tr>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td>4. Adaptiveness</td>
</tr>
<tr>
<td>- .33</td>
</tr>
<tr>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td>5. SToM1: Rapport building</td>
</tr>
<tr>
<td>- .32</td>
</tr>
<tr>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td>6. SToM2: Shaping the interaction</td>
</tr>
<tr>
<td>- .43</td>
</tr>
<tr>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td>7. SToM3: Detecting nonverbal cues</td>
</tr>
<tr>
<td>- .53</td>
</tr>
<tr>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td>8. SToM4: Taking a bird’s eye view</td>
</tr>
<tr>
<td>- .31</td>
</tr>
<tr>
<td>--------------------------------------------------------</td>
</tr>
<tr>
<td>9. Performance</td>
</tr>
<tr>
<td>--</td>
</tr>
</tbody>
</table>

In Study 2, all anxiety items loaded on one factor.
TABLE 5. DESCRIPTIVE STATISTICS FOR STUDY 4 PARTICIPANTS, BY SToM-SCALE SCORES

<table>
<thead>
<tr>
<th></th>
<th>High SToM scorers</th>
<th>Low SToM scorers</th>
<th>t-statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td></td>
</tr>
<tr>
<td>Age (years)</td>
<td>34.20 (7.52)</td>
<td>40.10 (10.05)</td>
<td>-1.49</td>
</tr>
<tr>
<td>Experience in sales (years)</td>
<td>8.30 (4.55)</td>
<td>9.90 (7.50)</td>
<td>.22</td>
</tr>
<tr>
<td>SToM</td>
<td>6.45 (.33)</td>
<td>5.18 (.40)</td>
<td>7.79***</td>
</tr>
<tr>
<td>SToM1: Rapport building</td>
<td>6.53 (.57)</td>
<td>4.67 (.85)</td>
<td>5.78***</td>
</tr>
<tr>
<td>SToM2: Detecting nonverbal cues</td>
<td>6.43 (.39)</td>
<td>5.37 (.82)</td>
<td>3.71**</td>
</tr>
<tr>
<td>SToM3: Taking a bird’s eye view</td>
<td>6.45 (.40)</td>
<td>5.47 (.70)</td>
<td>3.81**</td>
</tr>
<tr>
<td>SToM4: Shaping the interaction</td>
<td>6.37 (.40)</td>
<td>5.10 (.86)</td>
<td>4.22**</td>
</tr>
<tr>
<td>Adaptive selling</td>
<td>6.32 (.44)</td>
<td>5.14 (.64)</td>
<td>4.85***</td>
</tr>
<tr>
<td>Social anxiety</td>
<td>2.08 (.66)</td>
<td>3.63 (.71)</td>
<td>-5.07***</td>
</tr>
<tr>
<td>Perspective taking</td>
<td>5.43 (.66)</td>
<td>4.45 (.85)</td>
<td>2.99*</td>
</tr>
</tbody>
</table>

Note: All subjects are right-handed males. *p<.05, **p<.005, ***p<.001
TABLE 6. FOCI OF INCREASED ACTIVATION FOR CONTRASTS

a) Activations related to Interpersonal Mentalizing minus Process task contrasts

<table>
<thead>
<tr>
<th>Anatomical Region</th>
<th>MNI coordinates</th>
<th>Statistical effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L/R  x  y  z</td>
<td>Z-value  Cluster size(k)</td>
</tr>
<tr>
<td>Medial prefrontal cortex</td>
<td>R 10 58 20 3.86 64</td>
<td>*  *</td>
</tr>
<tr>
<td>Medial prefrontal cortex</td>
<td>R 2 48 42 3.60 30</td>
<td>*  †</td>
</tr>
<tr>
<td>Medial prefrontal cortex</td>
<td>L -14 48 36 3.71 60</td>
<td>*  *  *</td>
</tr>
<tr>
<td>Temporo-parietal junction</td>
<td>R 62 -46 4 3.35 18</td>
<td>*  *  *</td>
</tr>
</tbody>
</table>

b) Activations related to Interpersonal Mentalizing minus Unlinked Sentences task contrasts

<table>
<thead>
<tr>
<th>Anatomical Region</th>
<th>MNI coordinates</th>
<th>Statistical effects</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>L/R  x  y  z</td>
<td>Z-value  Cluster size(k)</td>
</tr>
<tr>
<td>Medial prefrontal cortex</td>
<td>L -14 52 34 3.77 51</td>
<td>*  *</td>
</tr>
<tr>
<td>Temporo-parietal junction/ Superior temporal sulcus</td>
<td>R 64 -44 6 4.20 46</td>
<td>*  *  *</td>
</tr>
</tbody>
</table>

I = interpersonal mentalizing group (high vs. low)

T = task

L = left, R = right

*p < .05, † p < .10
Figure 1.
Confirmatory Factor Analysis Models and Results for Sales Theory of Mind (SToM) Scale
(Study 1 findings not in parentheses, Study 2 in parentheses)

a. First-order 4-factor model

b. Second-order model
Note: SToM2* refers to a second-order sales theory of mind factor for which the three first-order factors shown load on this factor. All ellipses designate first-order factors, except for SToM2*, which is a second-order factor. All measures, factor loadings, and error variances are omitted from figure for simplicity. Coefficients in figure are standardized regression parameters.
FIGURE 3

STUDY 4: SIGNIFICANT CORRELATIONS BETWEEN SToM SCORES AND NEURAL ACTIVITY FOR INTERPERSONAL MENTALIZING MINUS PROCESS TASK CONDITION
Web Appendix 1

PRIMER ON fMRI AND SPECIFIC PROCEDURES USED

Functional magnetic resonance imaging (fMRI) is a tool for measuring brain activity over time. It can be used to produce activation maps showing which parts of the brain are involved in particular mental processes. The technique is non-invasive and the data have relatively good spatial and temporal resolution. Whereas conventional MRI provides images of structure (e.g., bone vs. muscle vs. fat), functional MRI provides images that estimate function (brain activity). During the past decade, fMRI has become an important research technique for studying normal brain functions in humans. The primer is necessarily brief; the reader is referred to standard references on fMRI methodology (Buxton 2002; Huettel, Song and McCarthy 2004) for more detailed discussions on key concepts presented in this section.

Data Acquisition

Both conventional MRI and fMRI work by sending out perfectly safe radiofrequency (RF) pulses and then listening for echoes. The RF pulses excite hydrogen protons (found in the water molecules throughout the body) into a higher energy state. When these protons relax into their lower energy state they emit a signal (or echo) that is detected by the MRI machine. The magnetic environment surrounding each proton influences how long it takes it to relax from the high-energy state to the low-energy state. Because bone, muscle, fat, and other types of tissues provide slightly different magnetic environments, the relaxation times of protons in these tissues are different. MRI can detect these differences and can therefore distinguish these different types of tissue.

Functional MRI works according to the same principles. It turns out that blood that is carrying oxygen provides a different magnetic environment than does blood that is not carrying
oxygen. MRI machines can be tuned to be particularly sensitive to this difference. And because oxygenated blood tends to be sent to parts of the brain that are active, fMRI can be used to estimate neural activity. It is important to remember that fMRI does not measure neural activity directly, but rather a Blood Oxygen Level Dependent (BOLD) signal that is strongly correlated with neural activity. This BOLD signal tends to lag behind the associated neural activity and to be more spread out in time. Modeling the relationship between neural activity and the BOLD signal (the so-called hemodynamic response function) is therefore critical in analyzing fMRI data.

During a typical fMRI experiment, a subject participant is asked to lie still on his or her back in an MRI machine for up to 90 minutes. An experimental session usually consists of 4 or 5 anatomical/structural scans of the brain taken during either the first or last 6 to 15 minutes. The participant is simply asked to lie still during this period and is not performing any tasks. These anatomical scans serve two purposes. First, they are used as guides in specifying exactly where the functional data should be collected (e.g., throughout the brain, only in the frontal lobe). Second, they provide a high-resolution image of the brain anatomy upon which the functional data can be overlaid; without the anatomical landmarks in a high-resolution structural image, it would be very difficult to determine where any observed brain activity actually occurred.

Functional data are collected in a series of “runs” (usually 5-10), each of which lasts 3 to 10 minutes. During each run, the participant performs whatever tasks the experimenter has designed. Often visual stimuli are projected on a monitor in front of the participant (or onto goggles) and the participant can make responses by making finger presses on buttons. Auditory and tactile stimuli (and even tastes and smells) are sometimes used as well. Vocal responses have occasionally been used but doing so is problematic because it introduces head movement. While the task is being performed, the MRI scanner is recording the BOLD signal throughout the brain every couple of seconds. These images are then analyzed to localize different mental
processes to different parts of the brain by identifying areas that are significantly more active during some conditions/tasks than others.

The most common and powerful way of testing the effect of the independent variable on the dependent variable is via a blocked design. As in any blocked-design experiment, a block in an fMRI study is composed of trials that are grouped together in time to represent a level of an independent variable. Thus experimental conditions are separated into distinct blocks with each condition presented for an extended period of time. Blocks can be as short as several seconds and as long as a minute or two, although block length is typically kept constant across conditions. Transitions between blocks represent changes in the level of an independent variable. Although many research questions are amenable to use of block designs, some questions may not be appropriate for blocked designs because the nature of the experimental task preclude the separation of different types of trials into distinct blocks.

Greater experimental flexibility is offered by event-related designs; they allow for detection of neural activity associated with discrete events that are short in duration and whose timing and order need to be randomized. Although event-related designs have reduced detection power relative to blocked designs, they tend to have good estimation power. The design allows for characterization of precise timing and waveform of the hemodynamic response associated with a discrete event. (For a detailed explanation of the relative strengths and weaknesses of the two designs as well as mixed designs, see Liu, Frank, Wong and Buxton, 2001.) Further, impressive gains in estimation power can be realized in event-related designs through the use of “jitter”—i.e., randomization of the intervals between successive presentations of events over some relatively long time period (Ollinger, Shulman and Corbetta 2001).

During the course of an fMRI experiment, functional (BOLD) images of the entire brain are recorded every 1 to 3 seconds. Depending on the length of the experiment, there may be 500 to 1500 of these functional brain images. Each of these images is divided up into a large number of small cubes called voxels (the three dimensional analog of pixels). The size of the voxels is
usually on the order of 3 to 5 mm cubic and it typically requires 25,000 to 50,000 voxels to cover the entire brain. Over the course of the entire experiment, the data from a single voxel constitute a time series of BOLD signals from the 500 to 1500 time points. Each of the tens of thousands of voxel time series is analyzed relatively independently in an attempt to identify voxels whose time series are significantly correlated with the experimental manipulations.

Preprocessing

Before the individual voxel timeseries are analyzed, however, a few preprocessing steps are typically performed on the data. Many researchers filter the data to exclude voxels that are outside the brain (in order to reduce analysis time). Some researchers attempt to correct for the fact that different slices within a single brain image are actually collected sequentially rather than at exactly the same point in time. For example, if a functional brain image is collected every two seconds, the first slice in that image is collected nearly two seconds before the last slice in that image. Many researchers will therefore shift the time series in time (via interpolation with earlier and later time points) in order to ensure that the voxels from different slices are in sync with each other.

The most important preprocessing step is probably correcting for motion. As previously mentioned, the timeseries from individual voxels are analyzed in an attempt to identify brain areas whose activity correlates with experimental manipulations. The underlying assumption is that the data from a voxel corresponds to the same brain area throughout the entire experiment. If a participant moves during an fMRI experiment, however, then the brain area to which a specific voxel corresponds will change. Therefore, virtually all researchers perform some kind of motion correction before analyzing the data. The standard approach is to perform a rigid-body transformation that includes six parameters (pitch, yaw, roll, and translation in x, y, and z) on the brain image from each time point until it best fits the brain image from the first time point. This process is called realignment.
The brains of different individuals obviously differ in size and shape. If results from different participants are going to be combined, it is therefore necessary to transform the data into some standard, template brain. This process is called normalization and it involves two steps. First, a set of parameters for a best-fitting, non-linear normalization transformation is computed. This transformation is usually the one that does the best job of mapping the structural brain image into the template brain, because the structural image has a higher resolution than the functional images. Second, this transformation is applied to the functional images to map them into the same space as the template brain (the functional images are usually in the same space as the structural image; if they are not, they must be coregistered into the same space first).

Another common preprocessing step is spatial smoothing. Essentially, functional brain images are blurred a little bit by convolving them with a Gaussian kernel (replacing the value at each voxel with a weighted average of its value and the values of surrounding voxels). There are three motivations for smoothing the data. First, realignment is not perfect and so the brain area to which a given voxel corresponds changes slightly over the course of the experiment. By smoothing the data, differences between nearby voxels due to motion are minimized. Second, normalization is imperfect and so the same voxel in different participants is unlikely to correspond to exactly the same brain area. Again, smoothing the data minimizes the effect of these small differences. Third, when a known smoothing kernel has been applied to the data, it makes it possible to apply a more sensitive correction for multiple comparisons. This issue will be discussed in more depth in the next section.

Model Fitting

Once the data have been preprocessed, each voxel is analyzed individually in an attempt to find voxels whose timeseries are significantly correlated with the experimental manipulations. As previously discussed, there are tens of thousands of voxels to be analyzed, so this approach
corresponds to doing many, many univariate analyses. The standard approach is to fit a general linear model against each voxel’s time series. The model would include covariates corresponding to the different conditions in the experiment. So, for example, if the participant repeatedly alternated between 10 seconds of visual stimulation and 10 seconds of rest, then the model might include a covariate that had the value 1 for each timepoint corresponding to visual stimulation and the value 0 for each timepoint corresponding to rest. It would probably also include an intercept term (value 1 at all time points) to model the baseline level of fMRI signal in the timeseries (which is typically far from 0). Fitting this model against a voxel’s timeseries would then correspond to finding the weighted sum of these covariates that best fits the actual time series. The weights or coefficients associated with each covariate in this best fit are called the beta values and they are used to compute statistical values (e.g., t-values) associated with each voxel for a given contrast of covariates. For example, the t-value corresponding to a single covariate’s effect is simply its beta divided by the standard error of the mean. Similarly, the t-value for a contrast between covariates is the difference between their betas divided by the standard error of the difference of the means. The statistics of interest are computed for every voxel to evaluate the probability that the voxel is consistent with the null hypothesis. The statistical tests from all voxels in the brain are then combined and displayed together in a statistical parametric map (or SPM) which is simply a brain image in which the value at each voxel is its corresponding statistic. These SPMs are in turn thresholded and overlayed on structural images in order to graphically display which areas of the brain exhibit activity that passes the desired threshold of statistical significance. Often different color schemes are used to aid in visualization (e.g., red for t-values above 3.5, yellow for t-values above 5.0, etc.).

In constructing a statistical model for fMRI data, it is important to keep in mind that the data reflect blood-oxygen levels, not direct neural activity. In particular, because blood-oxygen levels are delayed and extended in time relative to the underlying neural activity, the model covariates must also be delayed and extended in time. The standard approach is to create
covariates based on experimental conditions and then to convolve those covariates with a model of the hemodynamic response function.

Another important issue to keep in mind is that there is substantial temporal autocorrelation in fMRI data. That is, the data from timepoint X is not statistically independent of the data from timepoint X+1. As a result, the actual number of degrees of freedom is much smaller than it would be if the data from different time points were truly independent. The number of degrees of freedom has a substantial impact on the statistical values and so most analysis packages provide a way of estimating the effective degrees of freedom.

When this kind of analysis is done on voxels over the entire brain, a very substantial multiple comparisons problem arises. After all, when tens of thousands of voxel timeseries are being analyzed, it is quite likely that some of them would exhibit large statistical values by chance alone. The simplest way to address this problem is to apply a Bonferroni correction. Rather than using an alpha level of $p=0.05$ as is customary, one could use an alpha level of $p=0.05/n$ (where $n =$ number of voxels). The more standard approach is to look for clusters of contiguous voxels above some threshold where the cluster size is significant. If one knows how spatially smooth the data being analyzed are, then it is possible to estimate how likely it is to observe a cluster of N contiguous voxels all of which have a statistical value above a given threshold (most analysis packages provide this functionality). Given this approach requires knowing how smooth the data are, many researchers smooth their data during preprocessing as a means of imposing a known amount of spatial smoothness. Another approach for correction of the multiple comparisons problem is to evaluate statistical tests on a small predetermined region-of-interest (or ROI) and to exclude voxels outside the ROI from the analysis altogether.

ROI analysis involves prespecifying a set of anatomical regions of interest, and then to perform statistics across these regions (see Poldrack, 2007, for a discussion of ROI analysis). Because it is generally the case that regions specified in this approach are relatively large (e.g., the entire superior frontal gyrus), even if the region is significantly active, this activation may
occur in a small proportion of voxels in the ROI. This would mean that simply averaging across the entire region could swamp the signal from this small number of voxels with noise from the remaining non-activated voxels. This would be problematic because you may well have an *a priori* hypothesis as to an area of expected activation in a statistical parametric map based on prior findings in the literature. In such a case, to correct for multiple comparisons across the whole image would be too conservative, as you are restricting your interest to a subset of the comparisons being made. More recently, researchers have used a small volume correction approach developed by Worsley et al. (1996) in order to address this problem. This involves restricting the voxel-wise analysis to a ROI and then controlling for multiple comparisons only in those voxels.

**Apparatus**

Imaging was conducted using a full-body 3.0 T GE scanner (General Electric, Milwaukee, WI) fitted with an 8-channel receive-only head coil.

**Imaging Procedures**

For the structural imaging, a high resolution image of the brain was acquired with a 3D T1-weighted inversion recovery fast spoiled gradient recalled echo sequence (echo time (TE)/repetition time (TR)/inversion time = 2.1/10.4/300 ms, flip angle = 18°, matrix = 416x 256, field of view (FOV) = 25 cm, slice thickness 1.6 mm with 50% overlap). For the functional imaging, a time series of 210 volumes, with 39 Slices in the transverse plane, was obtained using single shot gradient-echo planar imaging (TR = 3000 ms, TE = 30 ms, flip angle = 75°, resolution = 3.5 mm x 3.44 mm x 2.3 mm, and FOV = 22 cm).

Functional image data were preprocessed and analyzed using Statistical Parametric Mapping (SPM2, Wellcome Department of Cognitive Neurology, London, UK). Linear image realignment, co-registration, non-linear normalization to stereotactic anatomical space (MNI),
and spatial smoothing 3-dimensional Gaussian kernel, 8mm full-width at half maximum (FWHM) were performed for each participant using standard statistical parametric mapping methods. A high-pass (cutoff period, 250 sec) frequency filter was applied to the time series.

References


DEFINITION OF INDEXES USED TO INTERPRET THE GOODNESSES-OF FIT OF
CONFIRMATORY FACTOR ANALYSIS AND STRUCTURAL EQUATION MODELS
IN STUDIES 1-3

1. Root Mean Square Error or Approximation (RMSEA).

The RMSEA is a population-based index of fit that is defined as

\[
\sqrt{\frac{\chi^2 - df}{N - 1}} / df
\]

where \(\chi^2\) = chi-square for a model of interest, \(df\) = degrees of freedom, and \(N\) = sample size. One set of guidelines maintains that a “close-fit” or “good fit” occurs when \(RMSEA < .05\), a “reasonable fit” or “acceptable fit” happens for values greater than .05 but less than or equal to .08, a “mediocre fit” occurs for values greater than .08 but less than or equal to .10, and a “poor fit” results for values greater than .10. The RMSEA is relatively insensitive to sample size, but of course findings under very low (\(N<100\)) and very large (\(N>1000\), say) sample sized, as well as deviations from normality, should be regarded with caution. The RMSEA tends to penalize complex models and favors parsimonious models.

2. Nonnormed Fit Index (NNFI).

Also known as the Tucker and Lewis index, the NNFI is defined as

\[
\frac{\chi^2_n / df_n - \chi^2_J / df_J}{\chi^2_n / df_n - 1}
\]

where \(\chi^2_n\) = chi-square for the null model of modified independence (i.e., the model where only error variances are estimated), \(\chi^2_J\) = chi-square of a focal model to be tested, and \(df\) = degrees of freedom. Depending on the author, values of the > .90 or > .95 are considered “good fits”. The NNFI takes into account model complexity, but again
caution should be applied for testing very small or very large samples and nonnormal data.

3. Comparative Fit Index (CFI)

The comparative fit index (also called the relative noncentrality index) is defined as

\[
\frac{(\chi^2_n - df_n) - (\chi^2_f - df_f)}{\chi^2_n / df_n}
\]

Depending on the author, values of the CFI > .90 or CFI > .95 are considered “good fits”. Although the CFI is relatively insensitive to sample size (at least for N not too small or too large), it does not compensate for more complexity.

4. Standardized Root Mean Square Residual (SRMR).

The SRMR is a measure of the average of residuals in a model and is defined as the square root of the mean squared differences between elements of the predicted and observed variance-covariance matrix. Depending on the author, values of the RMSEA < .08 or < .07 are considered satisfactory.
References


Web Appendix 3

THE GENERAL THEORY OF MIND (ToM) SCALE

1. I find it easy to understand non-verbal signals of other people.

2. I immediately notice when people do not smile sincerely.

3. I notice more quickly than others when people seem to possess a hidden agenda.

4. I find it easy to keep a conversation going about everyday topics or topics that do not have any urgency.

5. When I’m in an elevator with others I can easily start small talk.

6. When I’m sitting on a terrace I tend to elaborate on what motivates or drives people passing by.

7. I enjoy watching movies that provoke me to imagine the experiences of the characters.

8. I often think about deeper motivations of other people.

9. I enjoy speculating on what other people are thinking.

10. I tend to explain people’s behavior at a more sophisticated level than others.
All original versions of the following scenarios were presented in Dutch. In this appendix, they have been translated from the original language version into English, and therefore do not always reflect the same time length as the original language version.

Interpersonal mentalizing task

Scenario 1:
Sjaak is a salesperson who has just explained to Renée his own perspective about future trends in their market. Renée is the buyer in a customer’s firm and tries to sell Sjaak’s perspective on the market to his colleagues. Suddenly Sjaak realizes that he has provided Renée with the wrong information, and he immediately calls Renée. Renée is irritated and responds, “Do you know that you may have hurt my reputation? Sjaak apologizes and says, “I want to explain my mistakes to your colleagues personally.”

Why is it that Sjaak wants to explain his mistakes in person?

Scenario 2:
Before visiting a customer, Jacqueline always browses that customer’s website. While browsing one of these websites she notices that the director, whom she has known for a long time, still works for the firm in question; but she also notices that many new people have joined the firm. Jacqueline is especially curious about what these new people think of her firm. However, Jacqueline first decides to talk with the director, the person she has known a long time; therefore she calls him to suggest having dinner together.

Why did Jacqueline ask the director to have dinner with her?

Scenario 3:
Wouter is a street-smart salesperson and always tries to consider the personal interests of his customers. He mentions a customer’s personal interests to his secretary so that she can look for a gift that fits the customer’s needs exactly. He knows that when he surprises his customers, they invite him for dinner. Before sending a surprise present, Wouter calls the customer and says, “Hey, pal, take note: now I am not sending you a bill!”

Why does Wouter call the customer and make this statement?

Scenario 4:
Henk talks to a buyer, Janine. As the conversation evolves, Henk realizes that Janine shies away from sensitive issues. He starts to realize that Janine’s influence in the firm might be far less than he had assumed. Consequently, Henk considers how he can get around Janine without hurting her pride. He tells Janine, “During our next meeting perhaps it would be convenient to have a colleague from our technical staff join us, so would you also invite a colleague of yours?”

Why does Henk suggest that Janine invite other people to join the conversation?

Scenario 5:
Ralph, who is a buyer, talks to Pieter and to Pieter’s secretary. Ralph notices that Pieter is unfairly skeptical about his story while Pieter’s secretary is more receptive to his arguments. Ralph then adds something to the conversation. He tells Pieter a funny anecdote about how his own secretary once provided him with an insight which allowed him to avoid a grave mistake.

Why does Ralph mention this anecdote about his own secretary?

Process task

Scenario 1:
In a steel company the buying process occurs via a well-defined method: the buyers first study how earlier firms supplied goods; and, in collaboration with the technical staff, they make up a request for a proposal. This RFP is then sent by e-mail to salespersons from different firms, who then indicate by e-mail whether they can match the request for proposal. Subsequently, using economical arguments, the buyers determine which salesperson will deliver the goods.

On what bases do buyers make decisions about which salesperson will deliver goods?

Scenario 2:
An account manager visits his customers every year. According to a well-defined protocol he has to visit all the factory plants; and, in order to plan these visits, he uses a call-plan system. This planning system determines how different plants can be visited in the shortest amount of time. The account manager studies the planning results and notices that the plant in Amsterdam is the last one he has to visit.

Why does the account manager visit the Amsterdam plant last?

Scenario 3:
Long before the Christmas season, Mr. Versteeg, a salesperson, looks at the rules his company has devised for determining how much to spend on presents to be sent to his customers. Next he chooses two presents that match the set price. Another department then determines which present best fits the company policy rules; this evaluation process lasts a few weeks. Finally, presents are bought and are sent by mail to the customers.

Why does Mister Versteeg begin deciding so early what presents to buy for his customers?

Scenario 4:
For the customer, the buying process occurs via well-defined protocols: the buying customer asks for a meeting with the company’s technical staff via e-mail. During the meeting, alternatives from different suppliers are discussed in order to determine which supplier best meets the company strategy. The resulting information is then sent to a manager, who instructs others to design a checklist for the buying parties.

How does Miss Maartens, a customer, know that her buying follows the company policy?

Scenario 5:
An account manager of a bio-logistics company visits the customer in order to solve a logistics problem. The problem is that two of the customer’s three locations are being supplied by goods beyond the keeping abilities date. He explains to his customer that bio-logistics currently delivers the product in only one plant and that the other two plants are having their goods
delivered internally. The account manager suggests that it would be best to have the goods delivered to all the plants.

Why will a customer make more profit with the expansion of this service?

Unlinked sentences task

Scenario 1:
The company alignment has four plants spread over the Benelux. It is now already the second time that Mister Jansen has been invited to give a presentation. Frank has been account manager for 14 years, and he trains new buyers in his firm. Because of the intense competition from the Internet the future looks different. Peter’s office is on the third floor. The problems with traffic jams have risen quickly in the Randstad.

On which floor is Peter’s office?

Scenario 2:
On Main Street there is a large parking lot from which one can reach the train station. The construction of a network causes delay in information services. Miss Versteeg is an accountant and a mother of three children. The bicycle repairman just repaired a tube. The vacation time planned for this year is a bit unlucky because it falls at the time of an ad campaign. When the train arrives in the station at 4 o’clock we have four more hours before the theater performance starts.

Who repaired the tube?

Scenario 3:
This year the weather warmed so quickly that the skating rink closed one month earlier. The buyer today is not present; he is at the new plant. At the courtroom they say that they will come up with a verdict within 6 weeks. The e-mail did not arrive because many people are working with the server. There is a strike in the public transportation system.

Why did the e-mail not arrive?

Scenario 4:
The new broadcast about the nuclear experiments will be repeated at twelve o’clock. Gerard read enough and now has fallen asleep. Education takes on average five years, but it also can be finished in four years. We now live in an information age. New bridges are always built higher and longer, but where does all this end? It is time to move because this house is past its prime. The shops close at 9 p.m.

Why is it time to move?

Scenario 5:
People are working hard on the new block, and they expect it to be ready at the end of next year. People are starting to ask when they will come with the new folder? One can ask if our vision about the future will catch on in the marketplace. The number of customers is rising according to a pattern. The housing market at this time is a bit unstable because the future of the tax deduction for rent is unclear. Around the Christmas season, the days are always short.

Why is the housing market unstable?
TEMPORAL POLE ACTIVATIONS FOR HIGH AND LOW INTERPERSONAL MENTALIZING (IM) GROUPS

Main effect of Interpersonal Mentalizing Group (High vs. low): $F < 1$

Interaction of IM and Task: $F < 1$

High IM Group:

ToM – Process: $F(1, 9) = 21.00, p < .001$

ToM – Unlinked Sentences: $F(1, 9) = 23.02, p < .001$

Low IM Group:

ToM – Process: $F(1, 9) = 10.10, p < .01$

ToM – Unlinked Sentences: $F(1, 9) = 38.37, p < .0001$
Main effect of Interpersonal Mentalizing Group (High vs. Low): $F < 1$

Interaction of IM and Task: $F (2, 18) = 1.04, p = .36$

High IM Group:

ToM – Process: $F (1, 9) = 10.16, p < .01$

ToM – Unlinked Sentences: $F (1, 9) = 17.00, p < .001$

Low IM Group:

ToM – Process: $F (1, 9) = 14.26, p < .01$

ToM – Unlinked Sentences: $F (1, 9) = 32.30, p < .0001$