Physicians' interactions with peers: Empathic accuracy during shift-handovers on intensive-care units

Word count: 9271 words, abstract: 203 words

Number of tables, figures:1 figure, 1 table

None of the authors has any conflict of interest.

Correspondence should be addressed to:
Petra Klumb, Université Fribourg, RM 01 bu. C-1.115, Rue P.A. de Faucigny 2, 1700 Fribourg
petra.klumb@unifr.ch,+41 26 300 7643

This is the peer reviewed version of the following article:

Klumb, P. L., Wicki, C., & Rauers, A. (2019). Physicians' Interactions with Peers: Empathic Accuracy during Shift Handovers on Intensive-Care Units. ***Applied Psychology: Health and Well-Being*****, 11(1), 102-125. doi: 10.1111/aphw.12146,**

which has been published in final form at https://doi.org/10.1111/aphw.12146.

Abstract

Physicians' interactions with peers: Empathic accuracy during shift-handovers on intensive-care units

Background: We investigated how accurately physicians judge colleagues' states during shift handovers on intensive-care units, the role of physician characteristics, and how accuracy is related to handover partners' satisfaction.

Methods: Using mobile phones, we assessed momentary judgments during *N*=272 shift handovers by 36 physicians of five Swiss clinics. Physicians rated their own and their partner's affective states. We calculated the covariation of the perceiver's judgments of the partner's affect with the partner's self-reported affect and the perceiver's own self-reported affect. We then examined the moderation of these covariations by physicians' roles and experience.

Results: Overall, resident physicians were moderately successful in taking their counterparts' perspective: Perceiver's ratings of partner's affect and the latter's self-ratings were significantly related. Associations between perceivers' ratings of their own and their partner's affect were also evident. None of the effects varied as a function of physicians' roles. There was an unexpected effect of job experience; physicians with more experience were more likely to project their own affect into the rating of partner's affect. Physicians' accuracy in judging the partner's tense arousal was related to the partner's satisfaction with the social interaction. This effect may have been mainly driven by instances in which low tension was accurately judged, however.

Key words: Empathic accuracy; bias; intensive care; shift handover; event sampling

Practitioner points:

- Physicians were able to accurately judge internal states of their handover partners, but with increased work experience judgments became more biased.
- Better judgments of tense arousal were associated with partner's interaction satisfaction.

Worktime restriction has increased the number of physicians involved in the care of patients, specifically those in intensive-care units (ICUs). To assure continuity and quality of patient care, and thereby patient safety, handovers (also called handoffs or sign-outs) from shift to shift are necessary (Bangerter, 2002). Handovers are acts of communication that consist in an exchange of information about a patient that accompanies the transfer of control over (or accountability for) the patient from outgoing to incoming health professionals (Cohen & Hilligoss, 2010). Information transmission during handovers uses various modalities: verbal communication including paralinguistic elements, non-verbal behaviors, and also artefacts including paper forms and computer tools (for a review, see Abraham, Kannampallil, & Patel, 2014). Evidence shows that poor handoffs can result in serious problems, such as treatment failures and other preventable adverse events, which in turn can lead to increased length of stay, higher health-care costs, or even malpractice suits (e.g., Arora et al., 2005; Bates & Gawande, 2003; Horwitz, Moin, Krumholz et al., 2008). Therefore, transitions in care have been considered "danger points" in the patient care process (Horwitz, Meredith, Schuur, Shah, Kulkarni, & Jenq, 2009), and handover research has thus far focused on the accuracy and completeness of transmitted information (e.g., Bogenstätter et al., 2009; Pothier et al., 2005).

Completeness and accuracy of transmitted information (e.g., updates on a patient's current condition, any recent or anticipated changes, and required treatments) is considered an indicator of handover effectiveness or quality (Manser et al., 2013). Ratings of handover effectiveness seem to be driven by heterogeneous criteria, however, as Carroll, Williams, and Gallivan (2012) reported. In their research, the quality of the *relationship* between handover partners outweighed the exchange of complete and accurate information in determining handover quality. Differential weighting of criteria such as relationship quality may also explain why the authors found considerable variability in effectiveness ratings across roles and experience.

Currently, little is known about relational aspects of handovers and how these contribute to handover quality (Cohen & Hilligoss, 2010). It may be that skills related to relational communication, such as perspective taking, social support, and trust building play a key role (Carroll et al., 2012). The current research attempts to fill this gap by focusing on physicians' emotions during the handover process, and on their handover partner's ability to read these emotions – a topic neglected in research with health-care providers, particularly in live encounters (Hall, 2011).

The valence of emotional communication can provide important information in the handover process. For example, negative emotional information (such as how tense or tired a partner is) can point to potential problems or the possibility of a bad outcome (Baumeister et al., 2001), whereas a partner's positive affective expression may signal that no deviations from the normal routine are to be expected (e.g., Carver, 2003). When correctly interpreted, conclusions from both positive and negative emotional information may support the communicative process. Understanding others' emotions can be used to adapt one's own communication behavior (e.g., Hall, 2011; Parkinson, 1996; Schmidt Mast & Hall, in press), thus making it more beneficial for one's partner. An example of such a benefit is reduced uncertainty for instance in a patient regarding diagnosis and care in doctor-patient communication (see e.g., Gemmiti et al., 2017), or in a colleague regarding decisions on patient care in the course of the upcoming shift (see Arora et al., 2005).

Emotional information can be conveyed both verbally and nonverbally. In the area of doctor-patient communication, nonverbal information has been shown to serve important functions (e.g., Hall, 2011; Roter, Frankel, Hall, & Sluyter, 2006). Regarding handovers, nonverbal behavior may help physicians (a) enhance common ground as well as joint attention,

(b) monitor the degree to which a handover partner processes and reacts to transmitted information (Frankel et al, 2012), and (c) adapt their communication to the unique needs of a specific handover (Manser et al., 2013). For instance, if the physician handing over information notices that the incoming colleague is distracted the former has to redirect the latter colleague's attention to important details of the message. Better understanding of emotional cues thus contributes to a partner's reception of information and effective delivery of care in the upcoming shift, but very little is known about the concrete interactions taking place during a handover and physicians' ability to successfully interpret emotional cues.

Therefore, our study addresses the issue of communication competence in the handover process – a competence required by the Accreditation Council for Graduate Medical Education (ACGME, 2017) - and one that is known to be related to both objective and subjective performance outcomes in other occupational contexts (see review by Elfenbein & Ambady, 2002). Shifting attention from the technical content of the handover to its communicative basis, we focus on the assumptions individuals make about other people's internal states (*empathic inferences*) as well as their ability to make these assumptions correctly (*empathic accuracy;* Ickes, Bissonnette, Garcia, & Stinson, 1990).

Empathic accuracy is a function of the perceiver's skill to make accurate empathic inferences and the other person's tendency to express (or conceal) emotions (Zaki et al., 2008). Since both can vary considerably from person to person, it is important to observe multiple couplings of different partners when assessing empathic accuracy (Hall, 2011). In the current study, our goals were to (a) investigate differences in empathic accuracy across multiple handover partners, (b) test the influence of perceivers' characteristics (role and experience) on empathic accuracy, and (c) relate accuracy to outcomes.

Physicians' Empathic Accuracy

A partner's state is important for the regulation of communication processes in general (Schmid Mast & Hall, in press) and may be even more relevant in the high-stakes environment of intensive-care units (Frankel et al., 2012). Despite the challenges involved in correctly inferring another person's state in noisy and complex social situations, we therefore assumed that physicians still pay attention to their partner's state because of its information value (as described above). Rules or norms in the hospital culture can prescribe the suppression of some emotions (e.g., in order to maintain a professional appearance; cf. Kelly & Barsade, 2001; Morgan & Krone, 2001). For instance, anger or anxiety may be expressed less overtly and frequently, therefore physicians may be primarily attuned to variations in positive expressions. In the present study we therefore examine positive and negative emotions separately; we chose pleasant mood and tense arousal as specifications of these two dimensions that may be expressed differentially during shift handovers. The latter is a sign of colleagues' distress to which physicians may attend (similarly to patient distress, see Hall, 2011).

We predicted that physicians would be capable of achieving a significant degree of empathic accuracy for both positive and negative emotions based on their previous experience in the specific communication situation of the handover and based on research suggesting the general ease of making empathic judgments (e.g., Hall, Andrzejewski, Murphy, Schmid Mast, & Feinstein, 2008; but see Hall, Stein, Roter, & Rieser, 1999, who found physicians to have low overall levels of accuracy in judging their patients' emotional state).

On the other hand, there is evidence that transferring physicians do not have sufficient understanding of the perspective of their receiving counterparts on the handoff to be able to infer their specific needs (Manser et al., 2013). For example, outgoing physicians tend to overestimate

the similarity between their own and their handover partners' perspective, and consequently to overrate the effectiveness of their communication (Chang et al., 2010). In complex and demanding real-life settings like an intensive-care unit (ICU), the simplified strategy of projection may be employed more frequently than estimated in laboratory studies. Using one's own internal states as a reference point for judging the other person may be adaptive as this simplified strategy saves cognitive resources (Epley et al., 2004; Rameson 2012). Hence, we assumed that physicians are able to judge the affective state of their partners but will also rely on their own affective states when judging their handover partners thus biasing their judgments.

Hypothesis 1a: Perceivers are able to read others' affective states with a significant degree of correctness: Physicians' affect judgments are related to their partners' self-reported affect.

Hypothesis 1b: Perceivers use their own affective states as anchors when judging others' affective states: Physicians' affect judgments are related to their own affect.

Figure 1 about here

Individual Characteristics: Communication Role

The roles of transferring and receiving patient information that are assumed by outgoing and incoming physicians, respectively, are characterized by diverging demands, states, needs, and expectations (Carroll et al., 2012; Manser et al., 2013). This may lead to differences in empathic accuracy because perceivers' depth of information processing has been shown to vary with the goals they bring into the respective situation (Ickes, 2011; Klein & Hodges, 2001). For instance, physicians may aim to receive as much information as possible or aim to keep the handover short. In the latter case, physicians' motivation to be attentive to their counterpart's state may be limited. Since incoming physicians depend strongly on the information and assessments received

during the handover to inform their upcoming work, they should be more motivated to closely attend to their outgoing colleagues' affective communications.

Furthermore, outgoing physicians may also display less empathic accuracy, because correct empathic inferences rely, at least in part, on effortful processing (Zaki et al., 2009). We argue that the cognitive resources necessary for effortful processing are available to a higher extent to incoming physicians as compared to outgoing physicians for two reasons. First, incoming physicians are mainly listening and can concentrate on their handover partners rather than having to think about how to select, order, and present information as is the case for their outgoing colleagues (i.e., their task in the handover imposes fewer demands on cognitive resources; e.g., Rameson et al., 2012; see also Sweller, 1988, for a theoretical account). In addition, incoming physicians typically have more samples of information (e.g., tone of voice, speech velocity, posture, mimics, gestures, etc.) available on which to base their judgments that may make the task less demanding for them. Second, incoming physicians are likely to be in a more rested state in contrast to the fatigued state of many outgoing physicians (Frankel et al., 2012), and mental fatigue is associated with simplified information processing strategies (Bodenhausen, 1990). Hence, we assume role-related variability of accuracy and bias.

Hypothesis 2: Accuracy and bias vary as a function of work roles: Incoming physicians will show higher accuracy and lower bias than their outgoing colleagues.

Individual Characteristics: Experience

Individual differences in experience interacting with the handover partner and in the handover context (i.e., the ICU) could influence judgments. There is some evidence linking familiarity with an individual or a culturally defined group to accuracy in emotion recognition (Elfenbein & Ambady, 2003; Colvin et al., 1997). With experience, perceivers know about a

target's typical emotional expressivity, allowing them to perceive variations from this average as effective cues to the target's current affect in a given situation. Likewise, physicians' experience with the handover situation at the particular ICU should increase the validity of their empathic inferences. With more ICU experience, physicians should be better able to structure the handover by separating important from less relevant information (Carroll et al., 2012; Chang et al., 2010) and to more efficiently process handover information freeing up processing resources (Keller et al., 2016) for empathic inferences. Hence, we assume both types of experience will be positively associated with accuracy.

Hypothesis 3: Accuracy and bias vary as a function of work experience: More, as opposed to less, experience in the ICU and with the handover colleague will be associated with higher empathic accuracy.

Outcomes: Handover Effectiveness and Interaction Satisfaction

Evidence shows empathic accuracy to be associated with various positive outcomes, for example, workplace effectiveness, relationship quality, and social satisfaction (Elfenbein & Ambady, 2002; Blanke et al., 2016; for a review also see Hall et al., 2009). Being able to correctly judge the handover partner's inner states may allow physicians to better adapt their communication to the needs of the specific handover (Manser et al., 2013), and to respond more adequately to the social affordances arising during this task. This should contribute to handover effectiveness and thereby to colleague's satisfaction (as well as patient safety). We therefore hypothesized that empathic accuracy would allow for more successful communication, as indicated by ratings of handover effectiveness and interaction satisfaction.

Hypothesis 4: Empathic accuracy is related to handover outcomes: Smaller discrepancies between actor's partner ratings and partner's self-ratings will be associated with higher ratings of handover effectiveness and interaction satisfaction.

Method

The study took place at the ICUs of five different Swiss hospitals that treat patients undergoing extensive surgical interventions and in need of special care after the surgery. Physician residents worked in three shift cycles and information was typically handed over for about 45 minutes between 7 and 8 a.m., 3 and 4 p.m., and 11 and 12 p.m. At each site, we obtained approval of the research plan by the local ethics committee as well as participating physicians' informed consents.

Participants

Participants were 36 (N_1) physician residents working at 5 different ICUs. Eighteen (50%) of them were male and 16 (44%) female, and they were aged 28-44 years (mean age = 33.59 years, *SD* = 3.33). Two participants did not specify their age and sex. The ICU is one of the rotational positions within the physician resident's medical specialization training. The average duration of their stay on the ICU is about 6 months.

We sampled N_2 = 272 shift handovers between different dyads of physicians, with a mean of 18 handovers per physician (*SD* = 11.6). Note that the dyads were not always the same. Physicians interacted with different partners over the course of repeated handovers. They could change their roles (from outgoing to incoming physician and vice versa) and specific dyads of physicians could interact repeatedly, with varying or non-varying roles. We elaborate on this when presenting the analysis strategy.

Procedure

Via an online survey platform, participants provided demographic information and answers to questionnaires in an initial assessment. During the event sampling phase, we measured momentary affective states and characteristics of the handover at the beginning and the end of each handover. This questionnaire was administered via I-Phones to ensure it could be filled out anywhere immediately before and after each shift handover. At the end of the event sampling phase, physician residents filled in a questionnaire similar to the initial assessment. This study only draws on demographic information and data from the questionnaires filled out at the end of each shift handover.

Measures

Physicians' affective state. To characterize physician's momentary state, we focused on two affective dimensions: pleasant mood and tense arousal (factor adopted from the non-specific psychological distress scale; Almeida & Kessler, 1998). Participants rated their own affective state and that of their handover partners on five-point scales from 1 (*not at all*) to 5 (*very much*) for simple adjectives such as "good." Ratings for "good", "content", "calm", and "relaxed" (Multidimensional Mood Questionnaire*, MDBF*; Steyer et al., 1997; item intercorrelations all $r >$ .82, $p < .0001$) were aggregated into the broad scale of pleasant mood (self-ratings: range = 1-5, $M = 3.76$, $SD = 0.95$; judgments: range = 1-5, $M = 3.95$, $SD = 0.91$). Ratings for "nervous" and "unconfident" (*MDBF* and Profile of mood states, *POMS*; Biehl et al., 1986; item intercorrelation: $r = .71$, $p < .0001$) were aggregated into the scale of tense arousal (self-ratings: range: 1-4, $M = 1.64$, $SD = 0.83$; judgments: range: 1-4, $M = 1.63$, $SD = 0.79$). Self-reports for tense arousal and pleasant mood were negatively correlated within individuals ($r = -.52$, $p <$ .001). Handover partners' self-reports were uncorrelated for both tense arousal ($r = -.04$, $p = .45$)

and pleasant mood (r = .03, p = .51). After aggregating, we transformed the scales to approach normality (following recommendations by Tabachnick & Fidell, 2013). Eventually, this yielded four different scores for each partner (each dimension rated once for the participant and once for the partner).

Modeling empathic accuracy (hypotheses 1-3). Our study was informed by two methodological approaches for measuring empathic accuracy: the naturalistic assessment paradigm of Ickes' group (e.g., Marangoni et al., 1995) and the truth and bias model by West and Kenny (2011). A partner's self-reported affect is regarded here as the "truth" which should be uncovered by the perceiver. Hence, judgments of a partner's affective state can be modeled as a function of two forces: the partner's self-reported affect ("accuracy" after Ickes et al., 1990; "truth" after West & Kenny, 2011), and the perceiver's own self-reported affect ("projection" or "bias"). Bias is conceptualized here as the degree to which the judgment is similar to the judge's own current affective state. When dyad partners are in similar states (e.g., because background noise during the handover may increase tense arousal in both partners), the strategy of using one's own experience to infer one's partner's experience can contribute to accuracy. When partners are not in similar states, projecting own affect onto the partner (i.e., relying on egocentric anchoring or bias without sufficient adjustment) is not conducive to accurate judgments.

In each handover dyad, incoming and outgoing physicians rated themselves and their handover partner on the same affect dimensions. Using both partners' own, self-reported affective states as simultaneous predictors we modeled judgment of the partner's affect. The covariation of the perceiver's judgment with the partner's affect reflects "truth" (i.e., empathic accuracy) and the covariation of the perceiver's own affect with the rating given for the partner's affect reflects "bias" (i.e., how much the rating of the other person's state is influenced by one's

own current affective state). Following recommendations by West and Kenny (2011), we centered the outcome and both physicians' affect at the mean of the partner's affect. That is, we subtracted the mean of the partner's own, self-reported affect from each individual score.

We implemented the truth and bias model in a multilevel framework to account for statistical interdependencies arising from the design of the present study, which constituted an incomplete round-robin design with varying roles. During each shift handover, physicians judged their partner's affect and were judged by their partner. The individual affect judgments (N = 530) served as the outcome variable (level 1).

These individual judgments cannot be treated as independent observations, however. The data were derived from 272 handovers observed among 36 physician residents. Over time, each physician resident repeatedly judged various physicians' affect, while also repeatedly being judged by various other physicians. It should be noted that the observations were not neatly nested but crossed (as individuals changed their dyad membership). We therefore used crossed-random effects multilevel models (Baayen et al., 2008; Locker, Hoffman, & Bovaird, 2007) and included variance components (i.e., random intercepts; e.g., Bolger & Laurenceau, 2013) for the perceiver and for the target at level 2 to account for statistical interdependencies in the data, with each component significantly improving the model fit. These variance components were included to account for statistical interdependencies arising from individual differences pertaining to how physicians perceived their partners on average, and to how physicians were being perceived by others on average. In contrast, the fixed effects (i.e., the predictor variables) were included for hypothesis testing. Analyses were implemented in SAS 9.2 for Windows, using the mixed procedure (SAS PROC MIXED, REML).

Communication role. For each handover assessment, physicians indicated whether they were beginning or ending their shift. Communication role was coded as 0 for incoming and as 1 for outgoing physicians.

Work experience on the ICU. During the initial online assessment, participants reported the number of months of their work experience on the ICU (Tesluk & Jacobs, 1998). On average, they had worked there for 3.75 months (range: 0-23, *SD* = 5.07).

Experience with a specific colleague. For each interacting dyad, we additionally calculated the dyadic minimum and maximum of work experience on the ICU (i.e., the shortest and longest amount of time either partner had worked on the ICU). The dyadic minimum score served as an indicator for the experience specific dyad partners had in working together at the ICU (i.e., the minimum score denotes the first time point at which both partners were working at the ICU, *M* = 2.23 months, range: 0-23, *SD* = 3.48). The dyadic maximum score (*M* = 6.64, range = 0-23, *SD* = 6.71) was computed for use as a control variable, as we will elaborate later.

Predicting handover outcomes with empathic accuracy (hypothesis 4). Using five-point scales from 1 "not at all" to 5 "very much," we assessed two outcome measures: satisfaction with the social interaction ("I am satisfied with the handover interaction," range: 1-5 *M* = 4.18, *SD* = 0.88; correlation between the partners' ratings: $r = .20$, $p < .001$) and perceived handover effectiveness ("All the questions were resolved in the handover," range: 1-5 *M* = 4.15, *SD* = 0.86; correlation between the partners' ratings: $r = .10$, $p < .05$). The latter was adopted from the teamwork dimension of the handover-quality rating tool used by Manser et al. (2013). To test the hypothesis that empathic accuracy is related to these outcomes, we utilized a different indicator of empathic accuracy than for hypotheses 1-3: Here, we calculated the *absolute difference* between the perceiver's partner-rating and the partner's self-rating (tense arousal:

range = 0-3.5, $M = 0.69$, $SD = 0.72$; correlation between the partners' scores: r = .39, $p < .001$; pleasant mood: range = 0-3, $M = 0.95$, $SD = 0.74$; correlation between the partners' scores: r = .46, $p > .001$). Both partners' difference scores (i.e., both partners' empathic accuracy scores) were simultaneously entered as predictors to predict the partners' respective outcome in Actor-Partner-Interdependence models (APIM; Kenny, Kashy, & Cook, 2006). This allowed us to test the association of both partners' empathic accuracy with both partners' appraisal of the handover, while accounting for the statistical interdependence of the partners' scores.

Results

Physicians' Accuracy

Hypothesis 1 posited that physicians are able to judge their counterparts' affect but that they also rely on their own states for their judgements. To test this hypothesis, we predicted perceivers' judgments with the other partner's self-rated affect (truth) and the perceiver's own, self-reported affect (bias). Table 1 displays the coefficients for physicians' accuracy and bias. Accuracy and bias differed significantly from zero, indicating that perceivers' judgments are more accurate than chance, but are also subject to a systematic bias: Physicians rely on their own current affect to judge their partners. They seem to rely somewhat more on their own state (estimate for tense arousal = 0.37, $SE = 0.05$, $p < .001$; estimate for pleasant mood = 0.37, $SE = 0.04$, $p < .001$) than that of their counterparts (estimate for tense arousal = 0.24, $SE = 0.04$, $p < .001$; estimate for pleasant mood = 0.21, $SE = 0.04$, $p < .001$).

Whereas this indirect path (i.e., via one's bias; West & Kenny, 2011) can contribute to accuracy when perceiver and partner are actually similar, here, it is not adaptive because handover partners' similarities regarding their affect self-ratings approach zero ($r = .03$, $p = .51$ for pleasant mood and $r = -.04$, $p = .45$ for tense arousal). In this case, relying on one's own

affect to judge the partner is not conducive to empathic accuracy; it may even distort the judgment. Following a procedure introduced by West and Kenny (2011), we determined the indirect contribution of physicians' bias to their empathic accuracy. This potentially effective strategy increases physicians' overall accuracy by 5% for pleasant mood, resulting in an overall estimate for total accuracy of 0.22, whereas for tense arousal, it decreases total accuracy by 4%, resulting in an estimate of 0.23. There still remains some degree of accuracy in the tense arousal judgment, but it results from direct empathic inferences rather than bias.[1]

Influence of Communication Role

According to hypothesis 2, accuracy levels differ as a function of physicians' roles. To test this hypothesis, we again predicted the perceiver's judgment with the partner's self-reported affect (truth) and the perceiver's own, self-reported affect, separately for tense arousal and for pleasant mood. We included the perceiver's current role during a given handover, the interaction of truth and the perceiver's role, and the interaction of bias and the perceiver's role. In the resulting model, none of these interactions reached significance (all *p*s >.13). That is, neither accuracy of judgment nor bias was moderated significantly by the physicians' roles.

Influence of Experience

Hypothesis 3 referred to the effect of work experience on the ICU and specific experience within dyad constellations. To test these hypotheses, we again built on our basic model in which we predicted the perceiver's judgment with both partners' self-rated affect (i.e., with truth and

[1] West and Kenny (2011) distinguish between assumed similarity, or bias (the model estimate for the judge's own affect when predicting the judge's judgment) and real similarity (the implied correlation between the partners' self-reported affects). They furthermore distinguish between direct accuracy (estimate for the target's own affect when predicting the judge's judgment) and indirect accuracy (the product of assumed similarity / bias and real similarity). Total accuracy is the sum of indirect and direct accuracy.

bias). We used two different models for tense arousal and pleasant mood and repeated the analyses for two different indicators of experience (work experience on the ICU and experience with a specific colleague). To test for associations of experience with empathic accuracy, we included experience (grand-mean centered) as a predictor, as well as two interaction terms: first, the interaction of the partner's self-rated affect ("truth") and indicators of experience and second, the interaction of the judge's own, self-rated affect ("bias") with experience. For work experience on the ICU, we found significant interactions of bias and experience for both tense arousal (estimate = 0.43, *SE* = 0.14, $p < .01$) and pleasant mood (estimate = 0.29, *SE* = 0.12, $p < .05$) indicating that more work experience was associated with higher bias in the judgments. There was no interaction of experience with the partner's self-rated affect, indicating that direct accuracy was unrelated to experience (tense arousal: estimate = -0.03, *SE* = 0.10, $p = .76$; pleasant mood: estimate = -0.01, *SE* = 0.08, $p = .92$). A comparable pattern emerged for specific experience in working with the particular colleague. Increased experience with a colleague coincided with more bias in judgments of tense arousal (estimate = -0.25, *SE* = 0.11, $p < .05$) as well as those of pleasant mood (estimate = -0.19, *SE* = 0.08, $p < .05$). Again, direct accuracy was unrelated to experience (tense arousal: estimate = -0.06, *SE* = 0.11, $p = .56$; pleasant mood: estimate = 0.02, *SE* = 0.11, $p = .83$,).

As reported in the Methods section, experience with a particular colleague was operationalized as the dyadic minimum in the physicians' experience at this ICU (i.e., the work experience of the less senior physician of a dyad). In a control analysis, we next controlled for additional effects of the *more* senior physicians' work experience (i.e., for effects over and above the *less* senior physician's work experience). We again predicted affect judgments with the dyadic-minimum experience score and the abovementioned two interactions – one with truth and one with bias - as predictors. As additional predictors, we now included the dyadic-maximum

experience score (operationalized via the work experience of a dyad's *more* senior physician) and equivalent interactions – one with truth and one with bias. This did not change the results. The previously described interaction of the dyadic-minimum experience score with bias remained significant (tense arousal: estimate = -0.23, *SE* = 0.12, $p < .05$; pleasant mood: estimate = -0.19, *SE* = 0.09, $p < .05$), whereas the interaction of the dyadic-maximum score with bias was not significant (tense arousal: estimate = -0.004, *SE* = 0.01, $p = .71$; pleasant mood: estimate = -0.002, *SE* = 0.01, $p = .83$). Again, there were no interactions of experience with the partner's self-reported affect ("truth"; all *p*s > .70).[2]

Outcomes: Interaction Satisfaction and Handover Effectiveness

Empathic accuracy and satisfaction with the interaction. Perceiver's accuracy did not predict their own satisfaction with the interaction (tense arousal: estimate = 5.98, *SE* = 7.18, $p = .41$; pleasant mood: estimate = 12.70, *SE* = 14.44, $p = .38$). We also found no association of empathic accuracy for pleasant mood with the other partner's interaction satisfaction (estimate = 12.82, *SE* = 14.51, $p > .05$). We found a significant effect of physician's empathic accuracy for tense arousal, however: The more accurate one physician judged the partner's tense arousal, the higher was the partner's interaction satisfaction (estimate = 18.86, *SE* = 7.26, $p < .01$).

We considered the possibility that satisfaction was more likely to emerge if the partner correctly perceived low tension than when he or she accurately perceived high tension. We split the sample at the median for physician's self-reported tense arousal and repeated the analyses separately for handovers during which the target was relatively high versus relatively low in tense

[2] We tested whether these effects were related to physicians' status by controlling for differences in experience. We repeated the analyses and additionally included a three-way-interaction of perceiver's experience in the ICU, the partner's experience, and the perceiver's self-reported affect (that is, we tested for interactions of both partners' experience with bias). This did not change results, the three-way-interaction was non-significant for both affective states (tense arousal: estimate = -.0122, *SE*=.026, *p*=.6681; serenity: estimate = -.013, *SE*=.019, *p*=.49), thus lending no support for a potential role of status differences in the reported effects.

arousal. For handovers with relatively low tense arousal, there was still a marginal effect of empathic accuracy when predicting the other partner's interaction satisfaction (estimate = 18.72, SE = 10.00, DF = 201, t = 1.87, p = .06), whereas there were no such associations in the case of relatively high tense arousal (estimate = -0.23, *SE* = 10.04, *df* = 212, *t* = -.02, *p* = .98). That is, satisfaction with the interaction was higher if both partners agreed that there was no or little tense arousal, than when they disagreed in this regard. In contrast, empathic accuracy was not predictive of interaction satisfaction if the target was relatively high in tense arousal.

Empathic accuracy and handover effectiveness. There were no effects of empathic accuracy for either partner's perceived handover effectiveness (all *ps* > .20).

Discussion

At the end of daily shift handovers, 36 resident physicians in ICUs of five Swiss hospitals rated their affective states and those of their handover partners. Physicians ascribed to themselves and their handover partners low levels of tense arousal and higher levels of pleasant mood. Empathic accuracy (i.e., the association between perceiver's judgment of partner's affect and the partner's own self-reported affect), was significantly different from zero but modest in size. For both incoming and outgoing physicians, bias (i.e., the tendency to rely on one's own affect when judging the partner) was considerably higher than accuracy. While there were no effects of physicians' roles on either accuracy or bias, we found an effect of experience on bias: Both general work experience on the ICU and experience with a specific colleague were associated with increased bias. Physicians' accuracy for tense-arousal (but not for pleasant mood) predicted the other partner's satisfaction with the handover interaction, but accuracy was unrelated to perceived handover effectiveness.

Physicians' Accuracy

Physicians' level of accuracy in judging the states of their coworkers were more accurate than chance, even while engaging in the rather demanding task of a handover interaction. While the findings have to be replicated to gain more confidence in them, the pattern of results suggests that physicians have fairly good insight into their counterpart's affective states, but that their inferences also rely on projections of their own affective states. Importantly, the strategy of relying on one's own affect ("bias") can, in general, boost accuracy if current affect is in fact similar (i.e., positively correlated, Kenny & Acitelli, 2001; West & Kenny, 2011). In the current study, however, there was no similarity between dyad members' affective states, possibly because the situation was highly dissimilar for the dyad partners (with one physician about to end his or her shift, and the other one starting it). Therefore, physicians' bias did not help their accuracy on average and even reduced it in the case of tense-arousal judgments.

These findings from a real-life context are in line with evidence documenting moderate levels of success in paradigms where the perspective of others has to be taken (e.g., Fussell & Krauss, 1992). They also confirm that individuals use themselves as a starting point or anchor for developing models of others' typical emotional expressivity (Nickerson, 1999; Wilhelm & Perrez, 2004). The heuristic has its limitations, however, when putting too much weight on the anchor prevents sufficient adjustment (Epley et al., 2004) and thus leads to more bias that may prevent accurate judgment. Typically, the use of heuristics (and resulting biases) increases when resources are scarce. Even more than other real-life contexts, shift handovers at ICUs are characterized by exchanges of highly complex information under time constraints and with many distractions. Therefore, less resources for effortful deliberations about the counterpart's affective states may be available than in many other situations, particularly in comparison to the laboratory. In fact, in a real-life context like shift handovers, judging the partner's state is only a secondary task (cf. Hsu et al., 2008). When the primary task, i.e., handing over patient

information, imposes increased demands on cognitive resources, this may be reflected in a lower level of empathic accuracy. This dynamic cannot be extrapolated from laboratory research in which empathic inferences are the primary task.

Communication Roles and Physicians' Accuracy

There were no indications of differences in empathic accuracy as a function of physicians' roles – an effect that we had hypothesized from presumed role-based differences in cognitive resources and goals for the handover (Carroll et al., 2012; Manser et al., 2013). We hypothesized that outgoing physicians should be more tired and their cognitive capacity more exhausted because of task requirements. Incoming physicians, in turn, were expected to be more highly motivated to prepare well for the starting shift since the quality of their work depends in part on the information received from their partner. Nevertheless, we did not find these expected effects of role. One may speculate that a preoccupation of the physician with his or her own affective state masks the impact of one's role during the transition in and out of work. Alternatively, it may reflect the competence and professionalism of the physicians involved, who are able to regulate themselves in order to compensate for their fatigue. Overall and independently of their roles, physicians achieved an above-chance level of accuracy in judging their handoff partners' affective states but also relied to a substantial degree on their own states in judging their partner's.

One consequence of the high level of projection in judging colleagues' states – particularly for outgoing physicians – may be that physicians handing over information are less able to tailor their information to their incoming colleague because they fail to recognize the latter's needs (Manser et al., 2013). This follows from research by Chang and colleagues (2010) who found that outgoing physicians systematically fail to realize a breakdown of communication

in the course of handoffs. They overestimate how well they communicated. Therefore, they are less likely to verify whether the incoming colleagues actually understood the information handed over to them, leaving them with an incomplete state of knowledge. A similar finding was reported by Carroll and colleagues (2012): Nurses tended to overestimate knowledge of the incoming handover partners and left out too much information that the incoming nurses then requested as a means of compensation. Given the degree of accuracy reached by the physicians in our study, a complete breakdown of communication as described by Chang and colleagues (2010) is not likely to have happened in the handoffs we investigated.

Experience and Physicians' Accuracy

We examined the effect of experience both of the ICU in general and with specific handover partners. Regarding general ICU experience, we found greater experience was associated with bias for tense arousal and bias for pleasant mood. This converges with the finding of Chang and colleagues (2010) that there was no improvement in handover communication over time. They speculated that increasing experience alone does not suffice for performance to improve without formal instruction. Either experience on the job alone is not sufficient to achieve accuracy (e.g., Roulin et al., 2015) or it may lead to a decrease in the effort to read colleagues' states during handovers. Negative effects of experience on accuracy have been reported, before (e.g., Gesn & Ickes, 1999; Kilpatrick et al., 2002; Thomas, Flechter, & Lange, 1997) and are also known from research on job interviews. For instance, Dipboye and Jackson (1999) reported that initial information had a more pronounced bias on questioning for experienced than for less experienced interviewers.

On the one hand, the result has to be treated with caution because the sample size is limited and we did not sample to maximize differences in experience. On the other hand, our

confidence in the possibility to interpret the finding increased after a replication of the negative effect with a second facet of experience: familiarity with a specific colleague. Like work experience on the ICU, experience with a specific colleague was also associated with higher bias when judging tense arousal as well as with higher bias when judging pleasant mood. Experience with a specific colleague was operationalized as the least time that either of the dyad members had worked at the ICU (i.e., the work experience of the less senior physician in a dyad). To support the interpretation of this dyadic-minimum score as reflecting the dyad members' acquaintanceship, we additionally controlled for effects of the more senior physician's work experience, which did not have any effects over and above the dyadic-minimum score.

Hence, integrating results for work experience and experience with a particular colleague, there appears to be a more general pattern. The more the perceiver can rely on general knowledge already acquired (Cronbach, 1955) about her work and partners, the less closely she may monitor the contextual information of a specific handover and may focus instead on the transmitted information itself. If this effect can be established, it has implications for the way work is organized for teams. Studies of transactive memory look at this phenomenon from another angle but with the same result. They state that prolonged cooperation builds up transactive memory which, in turn, frees cognitive resources (e.g., Austin, 2003). It is not clear, however, how these resources are deployed. They could be used to deepen the analysis of the colleague but regularly may be shifted to the informational task at hand.

Alternatively, experience may be related to levels of power and there is evidence on power-related differences in interpersonal sensitivity/perspective taking (see e.g., Galinsky et al., 2003; Galinsky et al., 2006; Schmid Mast et al., 2009). In our sample, however, status differences were not substantial and it is unlikely that they drove the effect of experience.

Physicians' Accuracy and Partner's Satisfaction

We expected accurate judgments of colleagues' affective states to enable physicians to better adapt their communication behavior and hence the quality of information transmission during the handover to improve as a function of empathic accuracy. The quality of transferred information, in turn, may have implications for both, physician strain and patient safety in the subsequent shift. To examine these implications, we related differences between perceiver-judged and partner-reported affective states to partners' satisfaction with the handover process and to the degree to which all questions were clarified. As the only significant effect emerging from these analyses, empathic accuracy for tense arousal predicted the other partner's satisfaction with the interaction. While this result is in line with the notion that empathic accuracy may allow perceivers to adequately adjust their behavior to their partners in a given situation (e.g., Zaki et al., 2008) and to have better relationships in general (Hall et al., 2009), our confidence in the validity of this result was limited by a follow-up analysis suggesting that this effect was driven by a subsample of situations with low (rather than high) arousal. It is therefore conceivable that satisfaction did not emerge as a function of empathic accuracy, but instead resulted from relaxed handover situations in which both partners agreed that tense arousal was low.

Chang and colleagues (2010) reported a similar discrepancy regarding the level of agreement on information communicated and rated quality of handovers. An important quest for future research is, therefore, the determination of the conditions under which empathic accuracy affects outcomes during the handover or in the course of the subsequent shift (e.g., performance or strain). The present study suggests that training and job experience may be important dimensions for empathic accuracy.

Strengths and Limitations

We were able to establish cooperation in five ICUs for a considerable number of handovers although each ICU only has a fixed number of collaborators and there are many demands on physicians' time. Assessed in rich real-life contexts (Howland & Rafaeli, 2010) and based on external, realistic criteria (Trull & Ebner-Priemer, 2009), the accuracy measures possess high ecological validity. In addition, our study enabled us to relate differences in empathic accuracy to differences in physicians' work experience on the ICU and their experience with particular colleagues. But it is still important to replicate these findings in future studies with different – and bigger – samples characterized by a larger range of work experience and maybe also in different high-reliability organizations such as air-traffic control. Moreover, the empathic-accuracy paradigm compares judgments of handover partners' affective states with the self-reported affective states of the latter. While the correspondence of the two is labelled "truth" the partners' own self-reports may be subjected to biases, as well.

Finally, we reported on findings for two affective states. It is still conceivable that results could be different for other affect dimensions. For example, affects that are incompatible with physicians' roles such as disgust may be less clearly expressed and may be more difficult to read because display rules in the hospital culture prescribe the suppression of specific emotions (cf. Rafaeli & Sutton, 1989; Smith & Kleiman, 1989). In addition, the degree to which empathic accuracy is related to partner's satisfaction may vary with the affective dimension being judged.

Conclusions

The ability to infer another person's state is an important prerequisite of successful communication processes. In five intensive care units, physician residents inferred their colleagues' affective states during shift handovers with moderate accuracy, while also projecting their own affective states onto their colleagues – a strategy of limited utility in a situation where

states of incoming and outgoing physicians differed substantially. The finding that judgments did not improve with work experience, but rather became more biased, may have implications for organizing handovers and for formal training interventions (interpersonal skills can be improved through training, e.g., Elfenbein, 2006; Schlegel, Vicaria, Isaacowitz, & Hall, 2017). This study only provides limited evidence that empathic accuracy is related to handover outcomes, and more research is necessary before speculating about practical implications. It is, however, a first indication that handover satisfaction may be related to handover partners' social skills, which should be considered when conceptualizing and measuring handover success.

References

Abraham, J., Kannampallil, T., Patel, V. L. (2014). A systematic review of the literature on the evaluation of handoff tools: implications for research and practice. *Journal of the American Medical Informatics Association*, 21(1), 154–162. doi: 10.1136/amiajnl.2012.001351.

Accreditation Council for Graduate Medical Education (2017). *Common Program Requirements VI.E.3.c. https://www.acgmecommon.org/2017_requirements (accessed 11 SEP 2017).*

Almeida, D. M., & Kessler, R. C. (1998). Everyday stressors and gender differences in daily distress. *Journal of Personality and Social Psychology*, *75*(3), 670-680.

Arora, V., Johnson, J., Lovinger, D., Humphrey, H. J., & Meltzer, D. O. (2005). Communication failures in patient sign-out and suggestions for improvement: a critical incident analysis. *Quality and Safety in Health Care*, *14*(6), 401-407. doi: 10.1136/qshc.2005.015107.

Austin, J. R. (2003). Transactive memory in organizational groups: the effects of content, consensus, specialization, and accuracy on group performance. *Journal of Applied Psychology*, *88*(5), 866-878. doi: 10.1037/0021-9010.88.5.866.

Baayen R. H., Davidson D. J., & Bates D. M. (2008). Mixed- effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412. doi:10.1016/j.jml.2007.12. 005.

Bangerter, A. (2002). Maintaining interpersonal continuity in groups: The role of collective memory processes in redistributing information. *Group Processes & Intergroup Relations*, *5*(3), 203-219. doi: 10.1177/1368430202005003002.

Baumeister, R. F., Bratslavsky, E., Finkenauer, C., & Vohs, K. D. (2001). Bad is stronger than good. *Review of General Psychology*, *5*(4), 323-370. doi: 10.1037//1089-2680.5.4.323.

Bates, D. W., & Gawande, A. A. (2003). Improving safety with information technology. *New England Journal of Medicine*, *348*(25), 2526-2534. doi: 10.1056/NEJMsa020847.

Biehl, B., Dangel, S., & Reiser, A. (1986). Profile of mood states. In CIPS (Collegium International Psychiatriae Scalarum) (Ed.), *Internationale Skalen für Psychiatrie.* Weinheim, Germany: Beltz.

Blanke, E. S., Rauers, A., & Riediger, M. (2016). Does being empathic pay off?—Associations between performance-based measures of empathy and social adjustment in younger and older women. *Emotion, 16*(5), 671-683. doi: 10.1037/emo0000166.

Bodenhausen, G. V. (1990). Stereotypes as judgmental heuristics: Evidence of circadian variations in discrimination. *Psychological Science*, *1*(5), 319-322. doi: 10.1111/j.1467-9280.1990.tb00226.x.

Bogenstätter, Y., Tschan, F., Semmer, N. K., Spychiger, M., Breuer, M., & Marsch, S. (2009). How accurate is information transmitted to medical professionals joining a medical emergency? A simulator study. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *51*(2), 115-125. doi: 10.1177/0018720809336734.

Bolger, N., & Laurenceau, J. P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. New York, NY: Guilford Press.

Carroll, J. S., Williams, M., & Gallivan, T. M. (2012). The ins and outs of change of shift handoffs between nurses: a communication challenge. *BMJ quality & safety*, *21*(7), 586-593. doi: 10.1136/bmjqs-2011-000614.

Carver, C. S. (2003). Pleasure as a sign you can attend to something else: Placing positive feelings within a general model of affect. *Cognition and Emotion*, 17, 241–261. doi: 10.1080/02699930302294.

Chang, V. Y., Arora, V. M., Lev-Ari, S., D'Arcy, M., & Keysar, B. (2010). Interns overestimate the effectiveness of their hand-off communication. *Pediatrics*, *125*(3), 491-496. doi: 10.1542/peds.2009-0351.

Cohen, M. D., & Hilligoss, P. B. (2010). The published literature on handoffs in hospitals: deficiencies identified in an extensive review. *Quality and Safety in Health Care*, qshc-2009. doi: 10.1136/qshc.2009.033480.

Cohen, S., Schulz, M.S., Weiss, E., & Waldinger, R.J. (2012). Eye of the beholder: The individual and dyadic contributions of empathic accuracy and perceived empathic effort to relationship satisfaction. *Journal of Family Psychology, 26*, 236-245.

Colvin, C. R., Vogt, D., & Ickes, W. (1997). Why do friends understand each other better than strangers do? In W. Ickes (Ed.), *Empathic accuracy* (pp. 163-193). New York, NY, U.S.A.: Guilford Press.

Cronbach, L. (1955). Processes affecting scores on "understanding of others" and "assumed similarity.". *Psychological Bulletin*, *52*(3), 177-193. doi: 10.1037/h0044919.

Dipboye, R. L., & Jackson, S. L. (1999). Interviewer experience and expertise effects. In R. W. Eder & M. H. Harris (Eds.), *The employment interview handbook* (pp. 259–278). Thousand Oaks, CA: Sage Publications.

Elfenbein, H. A. (2006). Learning in emotion judgments: training and the cross-cultural understanding of facial expressions. *Journal of Nonverbal Behavior, 30*, 21–36.

Elfenbein, H. A., & Ambady, N. (2002). Predicting workplace outcomes from the ability to eavesdrop on feelings. *Journal of Applied Psychology*, *87*(5), 963-971. doi: 10.1037//0021-9010.87.5.963.

Elfenbein, H. A., & Ambady, N. (2003). When familiarity breeds accuracy: cultural exposure and facial emotion recognition. *Journal of Personality and Social Psychology*, *85*(2), 276-290. doi: 10.1037/0022-3514.85.2.276.

Epley, N., Keysar, B., Van Boven, L., & Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology*, *87*(3), 327-339. doi: 10.1037/0022-3514.87.3.327.

Frankel, R. M., Flanagan, M., Ebright, P., Bergman, A., O'brien, C. M., Franks, Z., ... & Saleem, J. J. (2012). Context, culture, and (non-verbal) communication affect handover quality. *BMJ Quality Safety*, *21*(Suppl 1), i121-i128. doi: 10.1136/bmjqs-2012-001482.

Fussell, S. R., & Krauss, R. M. (1992). Coordination of knowledge in communication: effects of speakers' assumptions about what others know. *Journal of personality and Social Psychology*, *62*(3), 378-391. doi: 10.1037/0022-3514.62.3.378.

Galinsky, A. D., Gruenfeld, D. H., & Magee, J. C. (2003). From power to action. *Journal of Personality and Social Psychology*, *85*(3), 453-466. doi: 10.1037/0022-3514.85.3453.

Galinsky, A. D., Magee, J. C., Inesi, M. E., & Gruenfeld, D. H. (2006). Power and perspectives not taken. *Psychological Science*, *17*(12), 1068-1074. doi: 10.1111/j.1467-9280.2006.01824.x.

Gemmiti, M., Hamed, S., Lauber-Biason, A., Wildhaber, J., Pharisa, C., & Klumb, P. L. (2017). Pediatricians' affective communication behavior attenuates parents' stress response during the medical interview. *Patient Education and Counseling*, *100(3)*, 480-486.

Gesn, P. R., & Ickes, W. (1999). The development of meaning contexts for empathic accuracy: Channel and sequence effects. *Journal of Personality and Social Psychology*, *77*(4), 746. doi: 10.1037/0022-3514.77.4.746.

Hall, J. A. (2011). Clinicians' accuracy in perceiving patients: its relevance for clinical practice and a narrative review of methods and correlates. *Patient Education and Counseling, 84(3)*, 319-324.

Hall, J. A., Andrzejewski, S. A., Murphy, N. A., Mast, M. S., & Feinstein, B. A. (2008). Accuracy of judging others' traits and states: Comparing mean levels across tests. *Journal of Research in Personality, 42(6)*, 1476-1489.

Hall, J. A., Andrzejewski, S. A., & Yopchick, J. E. (2009). Psychosocial correlates of interpersonal sensitivity: A meta-analysis. *Journal of Nonverbal Behavior*, *33*(3), 149-180. doi: 10.1007/s10919-009-0070-5.

Hall, J. A., Stein, T. S., Roter, D. L., & Rieser, N. (1999). Inaccuracies in physicians' perceptions of their patients. *Medical Care, 37(11)*, 1164-1168.

Howland, M., & Rafaeli, E. (2010). Bringing everyday mind reading into everyday life: Assessing empathic accuracy with daily diary data. *Journal of Personality*, *78*(5), 1437-1468. doi: 10.1111/j.1467-6494.2010.00657.x.

Horwitz, L. I., Meredith, T., Schuur, J. D., Shah, N. R., Kulkarni, R. G., & Jenq, G. Y. (2009). Dropping the baton: a qualitative analysis of failures during the transition from emergency department to inpatient care. *Annals of Emergency Medicine, 53(6)*, 701-710.

Horwitz, L. I., Moin, T., Krumholz, H. M., Wang, L., & Bradley, E. H. (2008). Consequences of inadequate sign-out for patient care. *Archives of Internal Medicine*, *168*(16), 1755-1760. doi: 10.1001/archinte.168.16.1755.

Hsu, K. E., Man, F. Y., Gizicki, R. A., Feldman, L. S., & Fried, G. M. (2008). Experienced surgeons can do more than one thing at a time: effect of distraction on performance of a simple laparoscopic and cognitive task by experienced and novice surgeons. *Surgical Endoscopy*, *22*(1), 196-201. doi:10.1007/s00464-007-9452-0.

Ickes, W. (2011). Everyday mind reading is driven by motives and goals. *Psychological Inquiry, 22*, 200–206. doi: 10.1080/1047840X.2011.561133.

Ickes, W., Bissonnette, V., Garcia, S., & Stinson, L. L. (1990). Implementing and using the dyadic interaction paradigm. H. Clyde & M. S. Clark (Eds.), *Research methods in personality and social psychology. Review of personality and social psychology, Vol. 11* (pp. 16-44). Thousand Oaks, CA, US: Sage Publications.

Keller, S., Tschan, F., Beldi, G., Kurmann, A., Candinas, D., & Semmer, N. K. (2016). Noise peaks influence communication in the operating room. An observational study. *Ergonomics, 59(12)*, 1541-1552. doi: 10.1080/00140139.2016.1159736.

Kelly, J. R., & Barsade, S. G. (2001). Mood and emotions in small groups and work teams. *Organizational Behavior and Human Decision Processes, 86(1)*, 99-130.

Kenny, D. A., & Acitelli, L. K. (2001). Accuracy and bias in the perception of the partner in a close relationship. *Journal of Personality and Social Psychology*, *80*(3), 439. doi: 10.1037/0022-3514.80.3.439.

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic data analysis*. New York : Cambridge University Press.

Kilpatrick, S. D., Bissonnette, V. L., & Rusbult, C. E. (2002). Empathic accuracy and accommodative behavior among newly married couples. *Personal Relationships*, *9*(4), 369-393. doi: 10.1111/1475-6811.09402.

Klein, K. J., & Hodges, S. D. (2001). Gender differences, motivation, and empathic accuracy: When it pays to understand. *Personality and Social Psychology Bulletin*, *27*(6), 720-730. doi: 10.1177/0146167201276007.

Locker L., Hoffman L., Bovaird J. A. (2007). On the use of multilevel modeling as an alternative to items analysis in psycholinguistic research. *Behavioral Research Methods, 39*, 723–730. doi: 10.3758/BF03192962.

Manser, T., Foster, S., Flin, R., & Patey, R. (2013). Team communication during patient handover from the operating room - more than facts and figures. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *55*(1), 138-156. doi: 10.1177/0018720812451594.

Marangoni, C., Garcia, S., Ickes, W., & Teng, G. (1995). Empathic accuracy in a clinically relevant setting. *Journal of Personality and Social Psychology*, *68*(5), 854-869. doi: 10.1037/0022-3514.68.5.854.

Morgan, J., & Krone, K. (2001). Bending the rules of" professional" display: Emotional improvisation in caregiver performances. *Journal of Applied Communication Research, 29(4)*, 317-340.

Nickerson, R. S. (1999). How we know—and sometimes misjudge—what others know: Imputing one's own knowledge to others. *Psychological Bulletin*, *125*(6), 737-759. doi: 10.1037/0033-2909.125.6.737.

Parkinson, B. (1996). Emotions are social. *British journal of psychology*, *87*(4), 663-683. doi: 10.1111/j.2044-8295.1996.tb02615.x.

Pothier, D., Monteiro, P., Mooktiar, M., & Shaw, A. (2005). Pilot study to show the loss of important data in nursing handover. *British Journal of Nursing*, *14*(20), 1090-1102. doi: 10.12968/bjon.2005.14.20.20053.

Rafaeli, A. & Sutton, R. I. (1989). The expression of emotion in organizational life. In L. Cummings & B. Staw (Eds.), *Research in Organizational Behavior*, *Vol. 11*, 1-42.

Rameson, L. T., Morelli, S. A., & Lieberman, M. D. (2012). The neural correlates of empathy: experience, automaticity, and prosocial behavior. *Journal of Cognitive Neuroscience*, *24*(1), 235-245. doi: 10.1162/jocn_a_00130.

Roter, D. L., Frankel, R. M., Hall, J. A., & Sluyter, D. (2006). The expression of emotion through nonverbal behavior in medical visits. *Journal of General Internal Medicine, 21(S1*).

Roulin, N., Bangerter, A., & Levashina, J. (2015). Honest and deceptive impression management in the employment interview: Can it be detected and how does it impact evaluations? *Personnel Psychology*, *68*(2), 395-444. doi: 10.1111/peps.12079.

Schlegel, K., Vicaria, I., Isaacowitz, D., & Hall, J. A. (2017). Effectiveness of a short audiovisual emotion recognition training program in adults. *Motivation and Emotion, XX, XX*.

Schmid Mast, M., & Hall, J. A. (in press). The impact of interpersonal accuracy for behavioral outcomes. *Current Directions in Psychological Science*.

Schmid Mast, M., Jonas, K, & Hall, J. A. (2009). Give a person power and he or she will show interpersonal sensitivity: the phenomenon and its why and when. *Journal of Personality and Social Psychology*, 97(5), 835-850. doi: 10.1037/a0016234.

Smith, A. C. & Kleiman, S. (1989). Managing Emotions in Medical School: Student's Contacts with the Living and the Dead. *Social Psychology Quarterly, 52*, 56-69.

Steyer, R., Schwenkmezger, P., Notz, P., & Eid, M. (1997). *Der mehrdimensionale Befindlichkeitsfragebogen* (MDBF) [The multidimensional affect rating scale (MDBF)]. Göttingen, Germany: Hogrefe.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*, 257–285.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th edition). London, England: Pearson Longman.

Tesluk, P. E., & Jacobs, R. R. (1998). Toward an integrated model of work experience. *Personnel Psychology*, *51*(2), 321-355. doi: 10.1111/j.1744-6570.1998.tb00728.x.

Thomas, G., Fletcher, G. J., & Lange, C. (1997). On-line empathic accuracy in marital interaction. *Journal of Personality and Social Psychology*, *72*(4), 839-850. doi: 10.1037/0022-3514.72.4.839.

Trull, T. J., & Ebner-Priemer, U. W. (2009). Using experience sampling methods/ecological momentary assessment (ESM/EMA) in clinical assessment and clinical research: Introduction to the special section. *Psychological Assessment, 21(4)*, 457-462. doi: 10.1037/a0017653.

West, T. V., & Kenny, D. A. (2011). The truth and bias model of judgment. *Psychological Review*, *118*(2), 357-378. doi: 10.1037/a0022936.

Wilhelm, P., & Perrez, M. (2004). How is my partner feeling in different daily-life settings? Accuracy of spouses' judgements about their partner's feelings at work and at home. *Social Indicators Research*, *67*(1-2), 183-246. doi: 10.1023/B:SOCI.0000007339.48649.20.

Zaki, J., Bolger, N., & Ochsner, K. (2008). It takes two - the interpersonal nature of empathic accuracy. *Psychological Science*, *19*(4), 399-404. doi: 10.1111/j.1467-9280.2008.02099.x.

Zaki, J., Weber, J., Bolger, N., & Ochsner, K. (2009). The neural bases of empathic accuracy. *Proceedings of the National Academy of Sciences*, *106*(27), 11382-11387. doi: 10.1073/pnas.0902666106.

Table 1

Fixed effects predicting the perceiver's judgment with both partners' self-reported affect using crossed-random-effects models.

	Tense arousal		Pleasant affect	
Fixed Effects (Predictors)	*Estimate*	*SE*	*Estimate*	*SE*
Constant	-0.009	0.029	0.451	0.695
Partner's self-reported affect	0.243***	0.044	0.212***	0.035
Perceiver's self-reported affect	0. 365***	0.048	0.374***	0.039
Variance Components				
Judge	0.020***	0.006	13.281***	3.865
Partner	0.005*	0.002	2.125*	0.958
N (observations)	456		516	

Note. We report fixed effects as unstandardized regression coefficients. Self-ratings and judgments for pleasant affect were squared and those for tense arousal were reflected and logarithmized. The direction of the constant for tense arousal is therefore reversed. We successively included a variance component (random intercept) for the perceiver and for the partner. For both affective states, these two steps each significantly improved the model fit. Additionally, including variance components for the dyad or the hospital did not further improve the model fit; these components were thus omitted from the models for the sake of parsimony. *** $p < .001$. * $p < .05$.