Everyday empathic accuracy in younger and older couples: Do you need to see your partner to know their feelings?

Antje Rauers, Elisabeth Blanke, and Michaela Riediger

Max Planck Institute for Human Development, Berlin, Germany

This is a pre-copyedited, author-produced version of an article accepted for publication in *Psychological Science* following peer review. The version of record

Rauers, A., Blanke, E., & Riediger, M. (2013). Everyday empathic accuracy in younger and older couples: Do you need to see your partner to know his or her feelings?. *Psychological Science*, 24(11), 2210-2217. Copyright © 2013 The Author(s) doi: 10.1177/0956797613490747

is available online at: https://doi.org/10.1177/0956797613490747.

Author Note

Antje Rauers, Elisabeth Blanke, and Michaela Riediger, Max Planck Research Group "Affect Across the Lifespan," Max Planck Institute for Human Development, Berlin, Germany.

Correspondence concerning this article should be addressed to Antje Rauers, Friedrich Schiller University Jena, Department of Developmental Psychology, Am Steiger 3, Haus 1, 07743 Jena, Germany, email: antje.rauers@uni-jena.de, phone: +493641 945203.

Abstract

On average, older adults are less accurate than younger adults in recognizing emotions from faces or voices. We challenge the view that such differences in emotion-recognition tasks reflect differences in empathic accuracy (the ability to infer others' feelings): Empathic accuracy does not only rely on sensory cues (e.g., emotional expressions), but also on knowledge about a person. Using smart phones, we assessed empathic accuracy in younger and older couples' daily lives. Younger but not older adults' empathic accuracy profited from the presence of their partners: Younger adults' accuracy was higher than older adults' when their partners were present and could be observed. During the partners' absence, however, when judgments relied exclusively on knowledge, no age differences emerged, and both age groups were still more accurate than chance. We conclude that across adulthood, sensory information and knowledge differentially support empathic accuracy. Laboratory emotion-recognition tasks may therefore underestimate older adults' empathic competencies.

Keywords: Empathic accuracy, experience-sampling, age differences

Everyday empathic accuracy in younger and older couples: Do you need to see your partner to know their feelings?

Think about a close friend or relative. Can you tell how this person is feeling right now, even though he or she is not even present at the moment? The current study suggests that your judgment would likely be better than chance. Whereas previous studies mostly studied people's ability to interpret emotional expressions in the laboratory, we examined how accurately people judge another individual's current feelings in everyday life. The ability to correctly judge another person's internal states, such as his or her current feelings, is referred to as empathic accuracy (Ickes, 1993; Zaki, Bolger, & Ochsner, 2009). This ability has been considered as the cognitive component of empathy (e.g., Richter & Kunzmann, 2011). It implies correctly inferring what another person is feeling, without necessarily sharing the target's emotional state or sympathizing with him or her. Empathic accuracy is considered to be crucial for social interactions, job performance, and personal well-being (Hall, Andrzejewski, & Yopchick, 2009). Alarmingly, empirical findings have conveyed the impression that this important ability may not remain reliable throughout life, but may decrease with aging, the decline beginning as early as middle adulthood. This impression derives from a large body of laboratory studies showing that middle-aged and older adults, compared to younger adults, do worse at reading emotions from sensory stimuli (e.g., facial expressions or voices; Lambrecht, Kreifelts, & Wildgruber, 2012; Riediger, Voelkle, Ebner, & Lindenberger, 2011; Ruffman, Halberstadt, & Murray, 2009; Ruffman, Henry, Livingstone, & Phillips, 2008). However, we propose that age differences in the specific ability to label sensory stimuli, such as emotional expressions, do not necessarily reflect age differences in the ability to infer another person's feelings.

In general, failing to understand emotional expressions in the laboratory does not necessarily imply deficits in inferring others' emotions in daily life (Funder, 1987; Krueger & Funder, 2004). Empathic judgments (i.e., assumptions about another individual's current feelings) in daily life do not only rely on the adequate perception of sensory cues, such as emotional expressions, but also on acquired knowledge (Ickes, 1993; Sze, Goodkind, Gyurak, & Levenson, 2012). For instance, people may take into account their knowledge about a familiar person (e.g., how tense this person usually feels at work) when making empathic judgments. This knowledge can support empathic accuracy even when the other person is currently absent and cannot be seen, heard, or talked to at the moment (Wilhelm & Perrez, 2004). Everyday empathic accuracy thus encompasses competencies beyond the ability to interpret sensory information as provided by the target's emotional expressions. This fact seems obvious. Surprisingly, however, age-comparative studies have predominantly targeted people's ability to interpret sensory stimuli. Much less is known about age differences in the ability to infer another person's current feelings; an ability that may well draw on interpretations of emotional expressions, but that may also rely on one's knowledge about an interaction partner. We argue that this conventional focus may not only be selective, but may also underestimate older adults' everyday competencies. Building on the two-component model of intellectual functioning (Baltes, 1987; Lindenberger & Baltes, 2000), we propose that the skills to use either sensory cues or acquired knowledge for empathic judgments age differentially. Sensory functioning (i.e., the speed and accuracy in processing sensory information) declines across adulthood (Baltes & Lindenberger, 1997; Lin et al., 2011). In comparison, acquired knowledge is less sensitive to aging (Charness & Krampe, 2008; Salthouse, 2003). Empathic judgments from sensory cues—an ability that is targeted when subjects are asked to label emotional expressions in traditional emotion-recognition tests—may thus be particularly difficult for older adults, compared to younger adults. In contrast, the ability to use acquired knowledge for empathic judgments should be less affected by aging. We therefore predicted that in daily life, age differences in empathic accuracy would be larger for judgments supported by sensory cues, and smaller for judgments derived exclusively from knowledge.

We tested this prediction by comparing two everyday scenarios: empathic judgments when the target person is present and empathic judgments when the target person is absent. Across both types of situations, people may use knowledge about a target (e.g., this person's typical mood in the morning) to judge his or her current affect. However, only during the target's momentary presence are additional sensory cues to the target's current affect available (e.g., facial or verbal information). Here, people can use sensory cues to adjust their empathic judgments. In contrast, when a target person is absent and cannot be observed, heard, or talked to, empathic judgments about this person have to rely exclusively on knowledge acquired before the time of the judgment. Across both types of situations, we repeatedly assessed empathic judgments in cohabitating heterosexual couples' daily lives. Because of the age-related decline in sensory functioning, we expected sensory cues, as potentially provided in the partner's presence, to be more supportive for younger than for older adults' empathic accuracy. When their partners were present, we therefore hypothesized that younger adults would be more accurate than older adults. However, we expected younger adults' advantage over older adults to cease during the partner's absence, when no sensory cues to the target's affect were available. Here, we expected that both age groups' use of acquired knowledge would allow valid assumptions about the partner's current feelings, resulting in reduced age differences, and greater-than-chance accuracy for both age groups. In short, we expected to find greater age differences when the partner was present than when the partner was absent.

Method

Participants

We recruited *N* = 100 heterosexual couples (i.e., 200 persons) from the Berlin area by means of newspaper advertisements and a recruitment company. There were two age groups: *N* = 100 younger adults, forming 50 couples (age range 20–30 years, *M* = 25.94, *SD* = 2.94) and *N* = 100 older adults, forming 50 couples (age range 69– 80 years, *M* = 74.20, *SD* =2.89). All couples cohabitated, with 6% of the younger couples, and 96% of the older couples being married. The relationship duration was 0.82–11.45 years (*M* = 4.52, *SD* = 2.54) for the younger couples and 14.77– 61.26 years (*M* = 48.62, *SD* = 10.16) for the older couples. Fifty-nine percent of the younger, and 51% of the older adults had graduated from high school or a higher educational institution.

Conventional emotion-recognition task

We included a conventional emotion-recognition task to ensure that the sample was representative with respect to the well-established age difference in facial-emotion recognition as measured in the laboratory. We used 36 pictures from the FACES database (Ebner, Riediger, & Lindenberger, 2010) showing younger, middle-aged, and older men and women posing angry, happy, sad, disgusted, neutral, and fearful facial expressions. For each facial expression, participants were asked to choose one of six emotion labels. As expected, younger adults outperformed older adults in this test (Figure 1; theoretical accuracy range: 0–1; younger adults: *M* = .71, *SD* = .09; older adults: *M* = .61, *SD* = .09), which we statistically tested in SAS (9.1; MIXED procedure) by predicting emotion-recognition scores with age group (parameter estimate of the age effect for men: - .10, $SE = .02$, $p < .0001$; estimate for women: -.11, $SE = .02$, $p < .0001$). To account for dyadic interdependencies among romantic partners, we treated partners as nested within couples, and couples as the unit of analysis (Campbell & Kashy, 2002). We ran parallel models for men and women, and requested different parameter estimates for men and women (Raudenbush, Brennan, & Barnett, 1995).

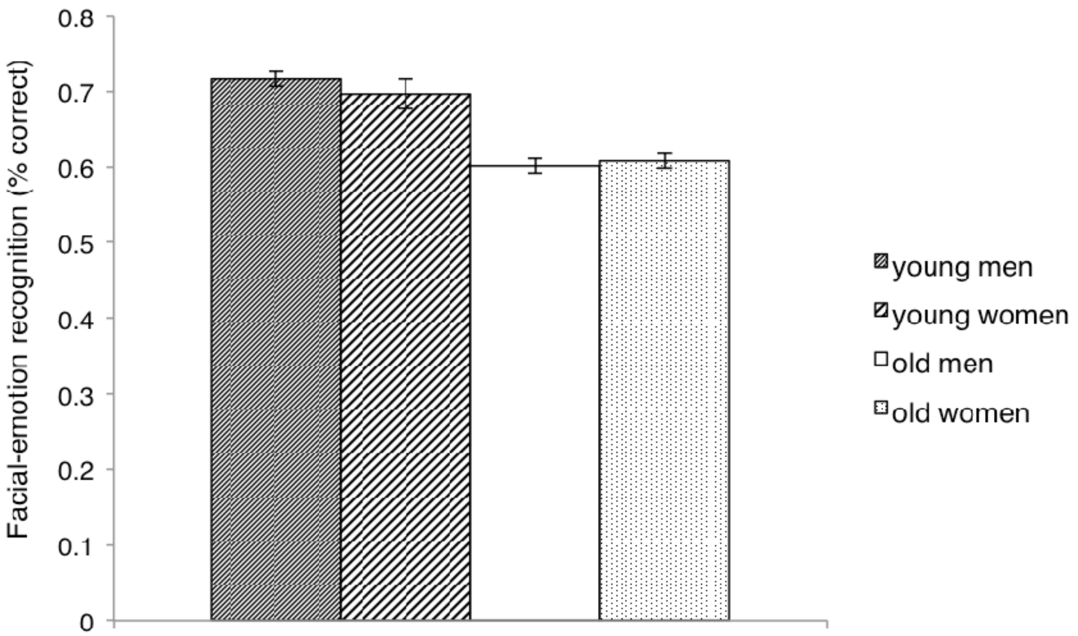

Figure 1. Age differences in emotion-recognition performance as tested in a conventional laboratory task using facial expressions as sensory cues. Bars represent model-implied means for all subsamples; error bars represent 2 standard errors from estimated subsample means.

Experience sampling using smart phones

We used a smart-phone based experience-sampling technology to capture empathic judgments in the participants' daily lives. Experience sampling collects data in people's natural living environments, capturing experiences and thoughts during, or in close temporal proximity to, their occurrence (Hoppmann & Riediger, 2009). Participants were provided with a Nokia E50 smart phone and responded to pseudo-randomized prompts. Cohabitating partner's schedules were synchronized so that partners were paged simultaneously. We instructed couples not to talk to each other while completing the assessment. On average, participants provided 86.76 measurements (range: 72–94, SD = 3.48) over the course of 15 days with six daily assessments.

Building on earlier accuracy paradigms (Ickes, Stinson, Bissonnette, & Garcia, 1990; Wilhelm & Perrez, 2004), we repeatedly asked both partners to rate their own current affect (self-rating) and their partner's current affect (judgment). At each mobile assessment, participants first provided *self-ratings* on eight affect items. We included four positive affect items (happy, enthusiastic, balanced, content) and four negative affect items (angry, downcast, disappointed, nervous) to cover low- as well as high-arousing affects for each valence, and to represent prototypical emotions typically displaying sufficient variability in both younger and older adults' daily lives. Using a seven-point scale from zero (not at all) to six (very much), participants indicated how much they experienced each of these affects at the moment. We subtracted ratings for negative-affect items from the ratings for positive-affect items to compute a measure of affect balance for each measurement occasion (theoretical range: -6 to 6; younger adults: M = 1.14, SD = .91; older adults: M = 1.45, SD = .83). Participants then provided *judgments* of their partner's current affect, using the same affect items and answer format as for their self-ratings. Again, we subtracted negative-affect ratings from positive-affect ratings to compute a measure of the affect balance participants ascribed to their partners (theoretical range: -6 to 6; younger adults: M = 1.14, SD = .81; older adults: M = 1.39, SD = .80). Self-ratings and judgments displayed severe positive skewness when untransformed. We applied an inverse logarithmic transformation to both variables to approach normality (Tabachnick & Fidell, 2007).

After the affect ratings, participants indicated whether or not their partner was present at the moment (present being defined as “I can see my partner right now”). On average, younger couples were together in 37% of all time points (range = 11–90%, between-person *SD* = 17), and older couples, in 59% of all time points (range: 18–96%, between-person *SD* = 24), respectively. The partners’ mutual reports on their co-presence diverged in 10% of the assessments. As these cases were likely caused by differences in the partners’ temporal delays in responding to the (synchronized) assessment instruments, we only included those measurements (*N* = 15,661) where both partners converged in their reports on being together or apart.

Statistical analyses: Modeling empathic accuracy

The basis for modeling empathic accuracy was the truth-and-bias model of judgment (West & Kenny, 2011), which we implemented in multilevel modeling using the MIXED procedure in SAS 9.1. The truth-and-bias model of judgment can be used to predict a person’s judgment about a target person (e.g., judgments about the target’s current affect) by some criterion of accuracy (e.g., the target’s self-rated current affect), while controlling for potential bias (e.g., the judging person’s own current affect). In our sample, each participant took on two roles for each measurement occasion: judging the partner’s affect (i.e., the role of the *rater*), and having their own affect judged by the other partner (i.e., the role of the *target*). To model empathic accuracy, we predicted the rater’s judgment with the target’s self-ratings across repeated assessments. Higher parameter estimates reflect higher empathic accuracy. We controlled for the rater’s self-rated affect, as romantic couples tend to be similar in their affects (Hoppmann, Gerstorf, Willis, & Schaie, 2011; Schoebi, 2008). Simply assuming that the partner's affect is similar to one's own may therefore enhance accuracy (West & Kenny, 2011). The present analyses account for this potential strategy and estimate participants' empathic accuracy above and beyond assumed similarity. We centered the two continuous predictors (i.e., both partners' self-ratings) and the dependent variable (i.e., the rater's judgment) at the personal mean of the target's self-rating. That is, we subtracted the personal mean of the target's self-ratings from each individual rating (West & Kenny, 2011). Further main predictors were the participant's age group (coded 0 = young/1 = old) and the partner's presence (coded 0 = absent/1 = present). To test our hypothesis on the age-differential role of the partner's presence for empathic accuracy, we included a three-way interaction between age group, the partner's presence, and the target's self-rating, as well as all lower-level interactions.[1]

Results

Our hypotheses were supported by the data. When predicting the rater's judgment of the target's affect, the three-way interaction of the target's self-rated affect, the partner's presence, and age group was significant for both men and women (parameter estimate of the interaction for men = -0.12, SE = 0.04, $p < 0.05$; estimate for women = -0.11, SE = 0.05, $p < 0.05$). This interaction effect was comparable for men and women (i.e., a contrast test by gender was not significant; estimate = -.01, $SE = .07$, $p = .84$). As illustrated by Figure 2 (left and middle panel), the interaction effect signifies that age differences in empathic accuracy differed depending on the presence or absence of the partner.

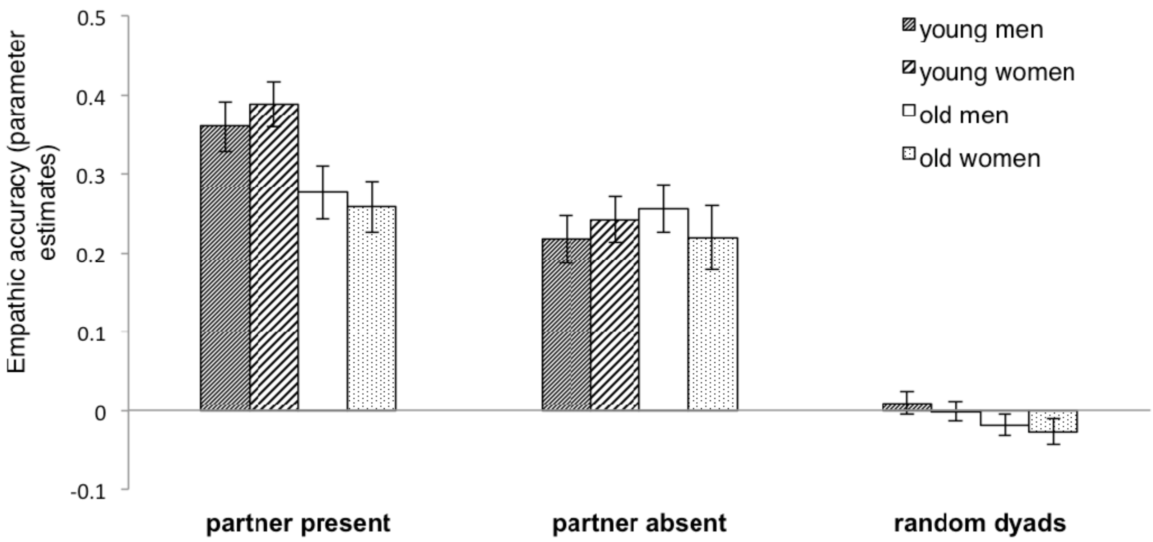

Figure 2. Empathic accuracy for romantic partner when present, for romantic partner when absent, and among random dyads. Bars show model-implied estimates for empathic accuracy, separately for each subsample and type of situation. Error bars depict 2 standard errors for parameter estimates. We ran complementary models with reversed dummy codes for dichotomous variables to generate each group of bars.

We followed up on the three-way interaction by repeating the analyses separately for younger and older adults (while omitting the variable of age group from the model). Again, the rater's judgment of the target's affect served as dependent variable. Predictors were the target's self-rated affect, the partner's presence, the interaction of these two variables, and the rater's self-rated affect. In line with our hypothesis, the interaction effect of the target's self-rated affect and the partner's presence was significant for both younger men (estimate = .17; SE = .03; p < .001) and younger women (estimate = .17, SE = .03, p < .001), but neither significant in older men (estimate = -.01, SE = .03, p = .71) nor older women (estimate = .01, SE = .04, p = .80). Put differently, younger adults' empathic accuracy profited from the partner's presence, while older adults' empathic accuracy did not vary significantly with the partner's presence or absence. We next followed up on the reported findings by running two separate analyses for situations with the partner present versus absent (while omitting the variable of the partner's presence from the model). There were no age differences when the partner was absent: Here, the interaction effect of the target's self-rated affect and age group was not significant (men: estimate = .05; SE = .05, p = .28; women: estimate = -.00, SE = .05, p = .99). Interestingly, empathic accuracy when the partner was absent was significantly better than zero (estimate for younger men = .22, SE = .03, p < .0001; estimate for younger women = .24, SE = .03, p < .0001; estimate for older men = .26, SE = .03, p < .0001; estimate for older women = .22, SE = .04, p < .0001), which is consistent with the assumption that knowledge about one's partner can serve as a basis for empathic judgments when the partner is momentarily absent. In contrast, when the partner was present, there were age differences: The interaction effect of the target's self-rated affect and age group was significant in women (estimate = -.012, SE = .04, p < .01) and marginally significant in men (estimate = -.08, SE = .04, p = .06). There were no gender differences in any of these follow-up analyses (all p for gender contrasts > .47).

Next, we tested if participants' judgments were specifically tailored to their partners, or if they instead relied on more general clues (e.g., on the time of day, or the target's gender and age). We randomly swapped partners within age groups, thus creating artificial age-homogeneous, cross-sex dyads. We then again predicted the rater's judgment with the random partner's self-rating, again controlling for the rater's own self-rating. The estimates for empathic accuracy were not different from zero, indicating that there was no empathic accuracy in these random dyads (see Figure 2, right panel; estimate for men = .01, SE = .02; p = .65; estimate for women = -.03, SE = .02, p = .20).

In a last step, we explored the role of the couples' relationship duration for our findings because on average, older couples had a longer relationship duration than younger couples. When repeating the analyses while controlling for relationship duration, our central finding—the three-way interaction of the target's self-rated affect, the partner's presence, and age group—remained significant, both for men (estimate = -.12, SE = .04, $p < .05$) and for women (estimate = -.11, SE = .05, $p < .05$). As relationship duration had a bimodal distribution in our sample, thus violating assumptions of normality, we additionally split the sample into three equal subsamples, dividing the sample at the 33.3rd, and the 66.6th percentile for relationship duration. The first group with the shortest relationship duration (n = 66) consisted of younger adults only, the medium group (n = 66) consisted of 51.4% younger adults, and the third group with the longest relationship duration (n = 68) consisted of older adults only. We then repeated the analyses while controlling for group membership (dummy coded with medium relationship duration as the reference group). Again, our central finding—the three-way interaction of the target's self-rated affect, the partner's presence, and age group—remained significant, both for men (estimate = -.12, SE = .04, $p < .05$) and for women (estimate = -.11, SE = .05, $p < .05$) . In a last step, we followed up on the general idea that relationship duration may influence the effect of the partner's presence on empathic accuracy. That is, we explored the explanatory value of relationship duration independent of participants' age. Within the age groups, the range in relationship duration was 0.82–11.45 and 14.77– 61.26 years, respectively. For these two available ranges, we tested whether the effect of the partner's presence on empathic accuracy depended on the couples' relationship duration. Here, we used relationship duration as a continuous predictor, but ran the analyses separately for both age groups (within the age groups, the distribution of relationship duration approached normality). We predicted the rater's judgment with the three-way interaction of the target's self-rated affect, the partner's presence, and the couples' relationship duration (group-mean centered), as well as all lower-level interactions and main effects. The three-way interaction was not significant, neither for younger adults (estimate for men = -.00, SE = .01, p = .98; estimate for women = .02, SE = .01, p = .09) nor for older adults (estimate for men = .00, SE = .06, p = .99; estimate for women = -.03, SE = .13, p = .79). This indicates that the effect of the partner's presence on empathy accuracy was independent of the couples' relationship duration, both within the range of 0.82—11.45 years, and within the range of 14.77—61.26 years. Taken together, these control analyses suggest that our central finding of conditional age differences in empathic accuracy (depending on the partner's presence) was not due to differences between younger and older participants' relationship duration.

Discussion

In this experience-sampling study, we investigated younger and older adults' empathic accuracy in daily life. Drawing on propositions of the two-component model of intellectual functioning (Baltes, 1987), we assumed divergent aging trajectories for the skills associated with using either acquired knowledge or sensory cues for empathic judgments. We had predicted that age differences in empathic accuracy would be greater when the partner was momentarily present (i.e., could currently be observed) than when he or she was absent.

This hypothesis was supported by our results. As expected, younger adults were more accurate than older adults when the partner was present, reflecting the pattern that typically emerges in laboratory studies investigating emotion recognition from sensory stimuli (Ruffman et al., 2008). In contrast, there was no age difference in accuracy when the partners were apart. This age-differential pattern of results was due to the fact that younger adults' empathic accuracy profited from the presence (versus absence) of their partners, whereas older adults' accuracy did not vary with the partner's presence or absence. Our results rule out the possibility that participants simply applied a stable representation of their partner's affect to all ratings. This strategy would not have resulted in accuracy, which we measured as the statistical covariation of two ratings over time. Importantly, all accuracy estimates were different from zero, indicating that both younger and older adults were more accurate than chance. Thus, romantic partners did not engage in random guessing, even when their partner was absent. This conclusion was further supported when we repeated our analysis after randomly swapping partners within age groups, thus creating artificial age-homogeneous, cross-sex dyads. Among those random dyads, empathic accuracy was not significantly different from zero.

The current study is the first to our knowledge to assess adult-age differences in everyday empathic accuracy. As opposed to the large body of past studies that focused on age differences in the ability to interpret emotional expressions in the laboratory, we investigated younger and older adults' ability to infer a real target's feelings in daily life—an ability that may draw on the target's emotional expressions, but also on additional information such as one's acquired knowledge about the target. The experience-sampling method that we used provides enhanced ecological validity, which is considered a particular asset in age-comparative research on empathic accuracy (Isaacowitz & Stanley, 2011). Ecologically valid assessments in the complexity of daily life, however, also imply limitations concerning questions of causality. For example, we did not measure participants' use of acquired knowledge and sensory information directly, but used the partner's momentary absence or presence as a proxy variable. Experimentally manipulating the cues to a target's affect poses a desirable route for future research. Furthermore, whereas the effects reported here are in line with the two-component model of intellectual aging (Baltes, 1987), our data do not provide an empirical answer to the question whether the observed results derive from aging-related changes or from other variables associated with age. For example, research suggests that empathic accuracy declines with relationship duration (Kilpatrick, Bissonnette, & Rusbult, 2002; Thomas, Fletcher, & Lange, 1997). Age group differences in relationship duration may thus have contributed to *mean levels* in empathic accuracy. Importantly, the literature does not suggest that relationship duration may *differentially* affect empathic accuracy (i.e., depending on the partner's presence), and exploratory analyses did not provide any support for this idea, either. It should however be noted that the variable of relationship duration had a bimodal distribution in our sample, with comparatively few cases in the medium range. Studying a wider range in both age and relationship duration than that covered by the present sample would be desirable in the future.

While these empirical extensions of the present research are still pending, the present findings nevertheless offer two important implications for future research on empathic accuracy: First, in daily life, people are able to make relatively valid empathic judgments without using any sensory cues, which should be reflected in approaches to measuring empathic competencies. Second, individuals may differ in how much their empathic accuracy profits from various sources of information, which we exemplified on the basis of age-group differences. In essence, our results suggest using more comprehensive approaches to investigating empathic accuracy, both regarding the factors supporting empathic accuracy, and interpersonal differences therein. In particular, they support current calls to revisit the question of age differences in empathic competencies (Isaacowitz & Stanley, 2011; Richter, Dietzel, & Kunzmann, 2010). Coming back to the initial question whether you could tell a social partner's current feelings while this person is absent: Your judgment would likely be better than chance, and whereas many abilities deteriorate with aging, this particular ability may remain reliable throughout your life.

Authors’ contributions

A.R. and M.R. developed the study concept and study design. M.R. conceptualized the dyadic mobile-phone based assessment technology and coordinated its development. Testing and data collection were performed by A.R. and E.B. A.R. performed the data analysis and interpretation, consulting M.R and E.B. A.R. drafted the paper, and M.R. and E.B. provided critical revisions. All authors approved the final version of the paper for submission.

Acknowledgements

This research was funded by the Max Planck Society. We thank Gloria Luong for helpful comments on an earlier version of this manuscript, and Julia Delius for editorial assistance.

References

Atkins, D. C. (2005). Using multilevel models to analyze couple and family treatment data: Basic and advanced issues. *Journal of Family Psychology, 19*, 98-110. doi: 10.1037/0893-3200.19.1.98

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology, 23*, 611-626. doi: 10.1037/0012-1649.23.5.611

Baltes, P. B., & Lindenberger, U. (1997). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: A new window to the study of cognitive aging? *Psychology and Aging, 12*, 12-21. doi: 10.1037/0882-7974.12.1.12

Bolger, N., & Shrout, P. E. (2007). Accounting for statistical dependency in longitudinal data on dyads. In T. D. Little, J. A. Bovaird & N. A. Card (Eds.), *Modeling contextual effects in longitudinal studies* (pp. 285-298). Mahwah, NJ: Lawrence Erlbaum.

Campbell, L., & Kashy, D. A. (2002). Estimating actor, partner, and interaction effects for dyadic data using PROC MIXED and HLM: A user-friendly guide. *Personal Relationships, 9*, 327-342. doi: 10.1111/1475-6811.00023

Charness, N., & Krampe, R. T. (2008). Expertise and knowledge. In S. M. Hofer & D. F. Alwin (Eds.), *Handbook of cognitive aging: Interdisciplinary perspectives* (pp. 244-258). Thousand Oaks, CA: Sage.

Ebner, N. C., Riediger, M., & Lindenberger, U. (2010). FACES – A database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods, 42*, 351-362. doi: 10.3758/BRM.42.1.351

Funder, D. C. (1987). Errors and mistakes: Evaluating the accuracy of social judgment. *Psychological Bulletin, 101*(1), 75-90. doi: 10.1037/0033-2909.101.1.75

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

Hoppmann, C. A., Gerstorf, D., Willis, S. L., & Schaie, K. (2011). Spousal interrelations in happiness in the Seattle Longitudinal Study: Considerable similarities in levels and change over time. *Developmental Psychology, 47*, 1-8. doi: 10.1037/a0020788

Hoppmann, C. A., & Riediger, M. (2009). Ambulatory assessment in lifespan psychology: An overview of current status and new trends. *European Psychologist, 14*, 98–108. doi: 10.1027/1016-9040.14.2.98

Ickes, W. (1993). Empathic accuracy. *Journal of Personality, 61*, 587-610. doi: 10.1111/j.1467-6494.1993.tb00783.x

Ickes, W., Stinson, L., Bissonnette, V., & Garcia, S. (1990). Naturalistic social cognition: Empathic accuracy in mixed-sex dyads. *Journal of Personality and Social Psychology, 59*, 730-742. doi: 10.1037/0022-3514.59.4.730

Isaacowitz, D. M., & Stanley, J. T. (2011). Bringing an ecological perspective to the study of aging and recognition of emotional facial expressions: Past, current, and future methods. *Journal of Nonverbal Behavior, 35*, 261-278. doi: 10.1007/s10919-011-0113-6

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

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

Krueger, J. I., & Funder, D. C. (2004). Towards a balanced social psychology: Causes, consequences, and cures for the problem-seeking approach to social behavior and cognition. *Behavioral and Brain Sciences, 27*, 313-327. doi: 10.1017/S0140525X04000081

Lambrecht, L., Kreifelts, B., & Wildgruber, D. (2012). Age-related decrease in recognition of emotional facial and prosodic expressions. *Emotion, 12*, 529-539. doi: 10.1037/a0026827

Laurenceau, J.-P., & Bolger, N. (2005). Using diary methods to study marital and family processes. *Journal of Family Psychology, 19*, 86-97. doi: 10.1037/0893-3200.19.1.86

Lin, F. R., Ferrucci, L., Metter, E., An, Y., Zonderman, A. B., & Resnick, S. M. (2011). Hearing loss and cognition in the Baltimore Longitudinal Study of Aging. *Neuropsychology, 25*, 763-770. doi: 10.1037/a0024238

Lindenberger, U., & Baltes, P. B. (2000). Life span psychology theory. In A. E. Kazdin (Ed.), *Encyclopedia of Psychology* (Vol. 5, pp. 52-57). New York: Oxford University Press.

Raudenbush, S. W., Brennan, R. T., & Barnett, R. C. (1995). A multivariate hierarchical model for studying psychological change within married couples. *Journal of Family Psychology, 9*, 161-174. doi: 10.1037/0893-3200.9.2.161

Richter, D., Dietzel, C., & Kunzmann, U. (2010). Age differences in emotion recognition: The task matters. *Journals of Gerontology: Psychological Sciences and Social Sciences, 66B*, 48-55. doi: 10.1093/geronb/gbq068

Richter, D., & Kunzmann, U. (2011). Age differences in three facets of empathy: Performance-based evidence. *Psychology and Aging, 26*, 60-70. doi: 10.1037/a0021138

Riediger, M., Voelkle, M. C., Ebner, N. C., & Lindenberger, U. (2011). Beyond ‘‘happy, angry, or sad?’’ Age-of-poser and age-of-rater effects on multi-dimensional emotion perception. *Cognition and Emotion, 25*, 968-982. doi: 10.1080/02699931.2010.540812

Ruffman, T., Halberstadt, J., & Murray, J. (2009). Recognition of facial, auditory, and bodily emotions in older adults. *Journals of Gerontology: Psychological Sciences and Social Sciences, 64B*, 696-703. doi: 10.1093/geronb/gbp072

Ruffman, T., Henry, J. D., Livingstone, V., & Phillips, L. H. (2008). A meta-analytic review of emotion recognition and aging: Implications for neuropsychological models of aging. *Neuroscience & Biobehavioral Reviews, 32*, 863–881. doi: 10.1016/j.neubiorev.2008.01.001

Salthouse, T. A. (2003). Interrelations of aging, knowledge, and cognitive performance. In U. M. Staudinger & U. Lindenberger (Eds.), *Understanding human development: Dialogues with lifespan psychology* (pp. 265-287). Dordrecht, Netherlands: Kluwer Academic Publishers.

Schoebi, D. (2008). The coregulation of daily affect in marital relationships. *Journal of Family Psychology, 22*, 595-604. doi: 10.1037/0893-3200.22.3.595

Sze, J. A., Goodkind, M. S., Gyurak, A., & Levenson, R. W. (2012). Aging and emotion recognition: Not just a losing matter. *Psychology and Aging, 27*, 940-950. doi: 10.1037/a0029367

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Boston, MA: Allyn and Bacon.

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

West, T. V., & Kenny, D. A. (2011). The truth and bias model of judgment. *Psychological Review, 118*, 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*, 183-246. doi: 10.1023/B:SOCI.0000007339.48649.20

Zaki, J., Bolger, N., & Ochsner, K. (2009). Unpacking the informational bases of empathic accuracy. *Emotion, 9*, 478-487. doi: 10.1037/a0016551

Footnote

[1] In order to account for statistical dependencies in the dyadic longitudinal data, we followed recommendations that statistical models for longitudinal data from distinguishable dyads should accommodate two levels only, despite the conceptual appeal of three-level models (for example, observations nested within persons, nested within couples; Atkins, 2005; Kenny, Kashy, & Cook, 2006; Laurenceau & Bolger, 2005). Here, these two levels pertain to individuals crossed with observations, nested within couples. We used parallel multilevel models for men and women (Bolger & Shrout, 2007; Raudenbush et al., 1995), requesting two distinct random intercepts for men and women. Beyond the variance explained by these random effects, additional variance was explained when specifying the residual structure, using the Kronecker product structure (TYPE= UN@AR(1) in SAS). This way, we accounted for (a) interdependencies due to repeated measurements over time and (b) dyadic interdependencies among the partner's ratings at a given measurement (Bolger & Shrout, 2007; Kenny et al., 2006).