Personality and Individual Differences 184 (2022) 111182

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

Comparing self-reported emotions and facial expressions of joy in heterosexual romantic couples☆

Katja M. Pollak[a,*,1], Sally G. Olderbak[a,*,1], Ashley K. Randall[b,2], Kevin K.H. Lau[b,2], Nicholas D. Duran[b,3]

[a] Ulm University, Germany
[b] Arizona State University, United States of America

ARTICLE INFO

Keywords:
Emotion
Facial expression
Self-report
Romantic partners
Dyadic data

ABSTRACT

To better understand individual differences in the expression of emotion within intimate relationships, we evaluated and compared patterns in facial expressions of joy against patterns in self-reported expressions of emotions in romantic couples. Using conversational data from 44 heterosexual romantic couples discussing four different topics, we examined the impact of stress on emotion expression, similarity in emotion expression between partners, and the influence of one partner's facial expressions on the self-reported expressions of the other partner. Overall, we found large differences between patterns in facial expressions of joy and patterns in self-reported emotions. First, using social relations analysis and generalizability analysis, we found that self-reported positive and negative emotions changed between stressful and non-stressful conversational topics, whereas facial expressions of joy remained stable. Second, we found similarities between romantic partners were common for self-reported positive emotions, less common for self-reported negative emotions, and uncommon for facial expressions. Finally, using Actor-Partner Interdependence Models, we found facial expressions of joy were unrelated to self-reported positive and negative emotions, and were non-significant predictors of partner's self-reported emotions. Our results challenge the use of *only* one methodology when measuring emotional experiences, as patterns observed for self-reported emotional data and facial expression data were not the same.

1. Introduction

To better understand individual differences in the expression and regulation of emotion within intimate relationships, researchers often conduct quasi-naturalistic studies observing interactions between romantic partners, measuring their self-reported emotion, physiological responses, and observing their non-verbal behavior. However, insights into the role of emotion, such as the impact of stress, similarity between partners in their felt emotion, or the impact of felt emotion, is often based on self-reported emotion (e.g., Griffin & Li, 2016) or physiological data (e.g., Reed, Randall, Post, & Butler, 2013), with fewer researchers investigating facial emotion expression data. Facial expressions are important indicators of non-verbal communication, as such it is important to incorporate this vital source of information.

Facial expressions of emotion are only weakly related, or in some cases unrelated, to self-reported emotion (Fernández-Dols & Crivelli, 2013; Reisenzein, Studtmann, & Horstmann, 2013), and are more variable and context-dependent than has been previously suggested (Barrett, Adolphs, Marsella, Martinez, & Pollak, 2019). This calls into question whether our understanding of the role of emotional processes in romantic relationships from self-reported data can be extended to that based on facial expressions of emotion. We will address this gap in the

☆ Author note:This research was supported by the Institute for Social Science Research at Arizona State University, awarded to Ashley K. Randall and Nicholas D. Duran, and the Margarete von Wrangell-Habilitationsprogramm für Frauen, awarded to Sally Olderbak. Portions of these findings were presented at the 2018 conference of the *International Association for Relationship Research*. The codebook of this study is available upon request by emailing ashley.k.randall@asu.edu.

* Corresponding authors.
E-mail addresses: katja.pollak@uni-ulm.de (K.M. Pollak), sally.olderbak@uni-ulm.de (S.G. Olderbak), ashley.k.randall@asu.edu (A.K. Randall), khlau2@asu.edu (K.K.H. Lau), nduran4@asu.edu (N.D. Duran).
[1] Institute of Psychology and Education, Ulm University.
[2] Counseling and Counseling Psychology, Arizona State Unviersity.
[3] School of Social and Behavioral Sciences, Arizona State University.

https://doi.org/10.1016/j.paid.2021.111182
Received 28 May 2021; Received in revised form 3 August 2021; Accepted 4 August 2021
Available online 20 August 2021
0191-8869/

literature from three perspectives, which we outline next.

1.1. Theoretical background

1.1.1. Impact of stress

Engaging in stressful conversation topics is linked to increases in self-reported negative emotions and decreases in self-reported positive emotions (Gottman, 1994). Stress is also linked to changes in facial expressions of emotion (Lerner, Dahl, Hariri, & Taylor, 2007). However, this research has utilized mostly artificial conditions (e.g., Trier Social Stress Test; Kirschbaum, Pirke, & Hellhammer, 1993) rather than the more naturalistic designs used in relationship research, such as asking romantic partners to discuss a stressful or conflict topic (Gottman, 1994). Hence, it is unclear if the paradigm most often used in relationship research to induce emotion impacts changes in facial expressions of emotion, particularly in the same way it impacts self-reported expressions of emotion. Thus, we will investigate patterns in changes in both facial expressions of emotion and self-reported emotion between romantic partners across different conversational topics, namely non-stress, external stress, internal stress, and enjoyment topics.

1.1.2. Similarity in emotion expressions

Romantic partners are similar in their self-reported emotions (e.g., Butner, Diamond, & Hicks, 2007). These similarities in emotion have several advantages, including a facilitated understanding of partners' emotions and a higher relationship satisfaction (Anderson, Keltner, & John, 2003; Gonzaga, Campos, & Bradbury, 2007). Using the Specific Affect Coding System (SPAFF), a behavioral coding system, similarities between romantic partners have also been found in nonverbal expressions of affect (Geist & Gilbert, 1996). However, similarities between romantic partners in self-reported emotions have not been compared to the similarities in facial expressions of emotions. Thus, we will assess and compare potential similarities between romantic partners in each conversation and examine whether conclusions about emotional similarities between romantic partners based on facial expression data match conclusions about emotional similarities based on self-reported data.

1.1.3. Partner effects

Lastly, we will examine the predictive quality of one romantic partner's facial expressions on the other partner's self-reported emotions. Given individuals use facial expressions to infer the felt emotion of another person, and perceiving this emotion often leads to the perceiver simultaneously feeling the same emotion (i.e., emotional contagion; Parkinson, 2011), many suggest that a similar process often occurs between romantic partners (Randall & Schoebi, 2018). Thus, we would expect that the emotional expressions of one partner will be associated with the self-reported emotion of the other partner. We will address this research question by investigating actor and partner effects of facial expressions of emotion on self-reported positive and negative emotions.

1.2. Present study

The purpose of this study is to extend previous research on interpersonal relationships that has mostly relied on self-reported data to a broader approach that combines self-reported data with facial expression data. In this way, we aim to gain a deeper understanding of how patterns of self-reported emotional data match those of facial expression data in romantic couples under naturally elicited stress and non-stress conditions.

2. Method

2.1. Participants

Data were collected from a community sample of different sex, committed couples between October 3, 2014 and March 8, 2015. Participants were recruited through Craigslist, Facebook, and electronic mailing lists in the Phoenix metropolitan area of the United States. Participating couples had to meet the following criteria: 1) 18 years or older; 2) in a romantic relationship for at least 6 weeks; and 3) both partners were willing and available to participate. The sample size for the present study was based on sample sizes from studies using similar designs (e.g., Reed et al., 2013). According to Kline (2005) who suggests recruiting five persons for each parameter in a Structural Equation Model (SEM), this sample size was also enough for the SEM analyses for the second research question.

A total of 73 couples ($N = 146$ individuals) inquired about the study; among these, 67 couples were screened as eligible, 54 couples completed a baseline questionnaire, and 44 couples ($M_{\text{age}} = 30.90$, $SD = 7.81$) completed all portions of the study. Most participants identified as White (73%) or Hispanic (16%). Participants were highly educated with approximately 82% holding a university degree or greater. Partners had been together, on average, for 6.13 years ($SD = 6.75$). One half of the couples reported that they were married and approximately 27% of the couples identified as having children.

2.2. Procedure

This study was approved by the Arizona State University's Institutional Review Board and all participants gave written informed consent. Data for this study were collected in three parts: 1) a screening survey, 2) a baseline questionnaire, and 3) a laboratory session. If deemed eligible, participants completed the baseline questionnaire independent from their partner, which took approximately one hour to complete. During the laboratory session, couples completed four video-recorded six-minute conversations: (1) baseline, (2) external stress, (3) internal stress, and (4) enjoyment.

During the *baseline* conversation, participants were instructed to sit in the room while the research assistant left the room. Thus, participants were left in the room to talk as they "normally" would. Conversation topics for the subsequent conversations were taken from participants' responses on the baseline questionnaires. The *external stress topic* was selected based on partners responses on the External Stress Scale (Borders, Randall, & Bodenmann, 2016). This conversation topic was counterbalanced across dyads; for example, in Couple 1 the female partner discussed her greatest external stress, in Couple 2 the male partner, etc. The *internal stress conversation* topic was chosen based on the areas of internal stress (e.g., difficult habits of the partner) that partners ranked similarly on the Multidimensional Stress Scale (Bodenmann, 2007). Lastly, the *enjoyable conversation* topic was chosen based on the areas of enjoyment that partners ranked similarly on the Enjoyable Conversations Scale (Gottman et al., 2003).

Participants were instructed to sit facing each other at an approximate 45-degree angle. One handheld Cannon VIXIA HF R500 video camera was placed in front of each participant to film the face and torso. Following the conversations, participants were asked to independently respond to several self-report items to report how they felt due to the previous conversation. Upon completion, couples were debriefed, and each partner was compensated $35 USD.

2.3. Measures

2.3.1. Self-reported emotions

Participants rated 15 emotion items using a 5-point Likert-type scale (0 = *not at all* to 4 = *a very large amount*). Three items reflecting positive (i.e., happy/joyful, loving/affectionate/caring, and positive) and four items reflecting negative (i.e., angry/irritated/annoyed/frustrated, sad, put-down/hurt/rejected, and negative) emotions were averaged to create scales of positive and negative emotions respectively. Internal consistency for the negative scale ranged from $\alpha = .46$ to $.85$ and for the positive scale from $\alpha = .71$ to $.87$.

2.3.2. Facial expressions of emotions

All videos were coded with Emotient SDK 4.1 (iMotions, 2016), a machine classifier that detects a face, tracks facial landmarks, and based on the location of these landmarks, computes evidence scores for seven emotions (anger, disgust, fear, joy, contempt, surprise, and sadness). Evidence scores are log odds of a human coder identifying that emotion as being present. For example, a score of 1 on joy means that expression is 10 times more likely to be categorized by a human coder as joy, a score of 2 means it is 100 times more likely, and so on. Evidence scores for the expression of joy, for example, indicate the full facial expression of joy and do not solely rely on smiles. Smiles, however, are a strong determinant of the facial expression score for joy. Emotient was shown to reliably and accurately code prototypical facial expressions of emotion for persons of different ethnicities and ages, across varying video conditions including picture resolution, lighting, head angle, and presence of glasses or hair (Emotient, 2016).

We chose to focus our analysis on expressions of positive and negative emotions. Of the seven emotions coded by Emotient, we decided to focus our analysis on three emotions - anger, sadness, and joy. We chose joy as an indicator of positive emotions, anger as an indicator of high-arousal negative emotions, and sadness as an indicator of low-arousal negative emotions. Evidence scores for anger, joy, and sadness were averaged across a conversation, for the full duration of the conversation, separately for each emotion, person, and conversation (Olderbak, Hildebrandt, Pinkpank, Sommer, & Wilhelm, 2014). Except for six data points, all anger scores (average across all conditions: $M = -2.17$ [-4.63 to -0.19], $SD = 0.89$) and sadness scores (average across all conditions: $M = -1.66$ [-3.80 to -0.13], $SD = 0.70$) were below 0 for every individual in every condition, indicating that on average these emotions were not expressed. Hence, that data was excluded. The joy scores, instead, were more common (average across all conditions: $M = -0.88$ [-5.11 to 2.18], $SD = 1.48$). While negative scores also indicate this emotion was not expressed, those scores could still be interpreted within the context of joy expression meaning joy was not being expressed. Thus, we decided to focus our analysis on facial expressions of joy, which can be interpreted from absent to present.

3. Results

The data can be found in the associated project on the Open Science Framework (https://osf.io/2vph7/?view_only=2f2b6c8f35c54c0a821929521fe87c4b). Three additional manuscripts have been published from this dataset (Grafsgaard, Duran, Randall, Tao, & D'Mello, 2018; Lau, Randall, Duran, & Tao, 2019; Randall, Tao, Leon, & Duran, in press). The syntax and results can be found in the online supplementary material. All analyses were performed using R 3.6.1 (R Development Core Team, 2001). Before the analyses, we removed univariate outliers (±3.5 SDs) for facial expressions of joy and self-reported negative and positive emotions within each condition, across men and women. In total, four data points from the self-reported negative emotions were excluded (two in the baseline and two in the enjoyment condition).

4. Impact of stress

First, we investigated the extent to which stress, induced through discussing stressful topics, changed facial expressions of joy and self-reported positive and negative emotions. This was investigated through social relations analysis (Kenny and La Voie, 1984), a method which enabled us to decompose the observed variance in facial expressions of joy. Specifically, through this decomposition, we could assess whether the observed variance is because different couples vs. different persons vs. different conversations were examined (vs. their interactions). Using the R package *VCA* (Schuetzenmeister & Dufey, 2018), we ran separate social relations analyses for self-reported positive emotions, self-reported negative emotions, and facial expressions of joy (see Table 1). We then estimated the generalizability of emotion scores across conditions through a generalizability analysis (Shavelson, Webb, & Rowley, 1989).

Table 1
Estimates of variance components in percent.

	Facial expressions of joy	Self-reported positive emotions	Self-reported negative emotions
Couple	64*	38*	39*
Person (nested within couple)	19*	7*	8*
Conversation	2*	30*	25*
Couple crossed with conversation	12*	21*	24*
Person (nested within couple) crossed with conversation + error	2	5	4

We cannot compute a separate estimate for *person (nested within couple) crossed with conversation* because it is confounded with the error component.
* Indicates this coefficient is statistically significant ($p < .05$)

4.1.1. Self-reported emotions

The couple facet contributed the most to variance in self-reported emotions for both positive and negative emotions (38% and 39% respectively), followed by conditions (30% for positive and 25% for negative emotions), indicating that different conversation topics impacted self-reported emotions. As is illustrated in Fig. 1, changes in self-reported emotions matched the conditions, i.e., less positive, and more negative emotions were reported after the stressful conversations compared to the enjoyment and the baseline conversations. This finding is further illustrated by the relatively low generalizability of self-reported positive (generalizability coefficient of .29) and negative emotions (generalizability coefficient of .33) across conditions.

4.1.2. Facial expressions of joy

Most of the variance in facial expressions of joy was attributable to the couple (64%), whereas only a very small amount (2%) was attributable to different conditions. This means there was no meaningful difference in facial expressions of joy between conversational topics (see Fig. 1) and that the generalizability of facial expressions for joy across conditions was high (generalizability coefficient of .83).

Overall, we found self-reported emotions changed across different conditions while facial expressions of joy remained stable.

4.2. Similarity in emotion expressions

Next, we examined similarities between males and females in self-reported emotions and facial emotion expressions of joy. To adjust for potential differences in variances between males and females, which impact the magnitude of correlations, we used SEM, which allowed us to standardize variances (Kenny, Kashy, & Cook, 2006). Specifically, we calculated one separate SEM for each estimate of emotion: self-reported positive emotions, self-reported negative emotions, and facial expressions of joy. For each SEM, we used Maximum Likelihood estimation and Full-Information Maximum Likelihood to handle missing values.

4.2.1. Self-reported positive emotions

For the first SEM, we included self-reported positive emotions for men and women for each conversation as eight separate manifest variables. Each manifest variable was then predicted by a single latent variable and the latent variances were fixed to 1. The latent variables were allowed to correlate with one another. The full correlation matrix is provided in the supplementary material. For ease of reading, we will only discuss those relevant to our research question. We found

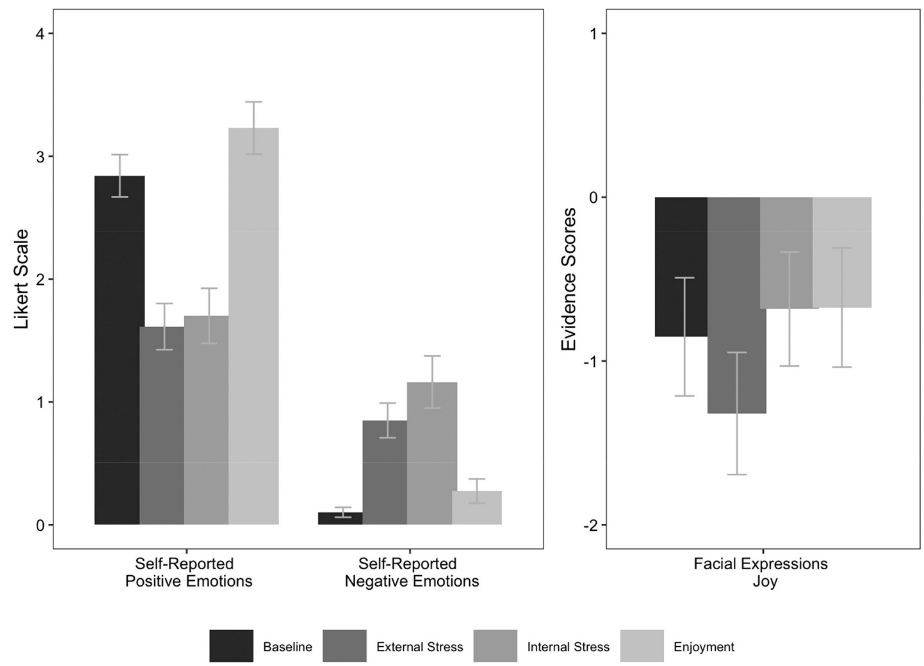

Fig. 1. Average self-reported positive and negative emotions, and facial expressions of joy across conditions.
Note. Error bars represent 95% confidence intervals. Self-Reported Emotions can range between 0 and 4, evidence scores can be interpreted as likelihoods (negative scores indicate that participants did not express joy, see Method section for details).

significant moderate correlations between males and females in every condition except for the enjoyment condition (see Fig. 2).

4.2.2. *Self-reported negative emotions*

We followed the same procedure for comparing self-reported negative emotions between males and females. We found no significant correlations between males and females in any of the condition except for the internal stress condition. During this condition, males' and females' self-reported negative emotions were significantly and positively correlated (see Fig. 2).

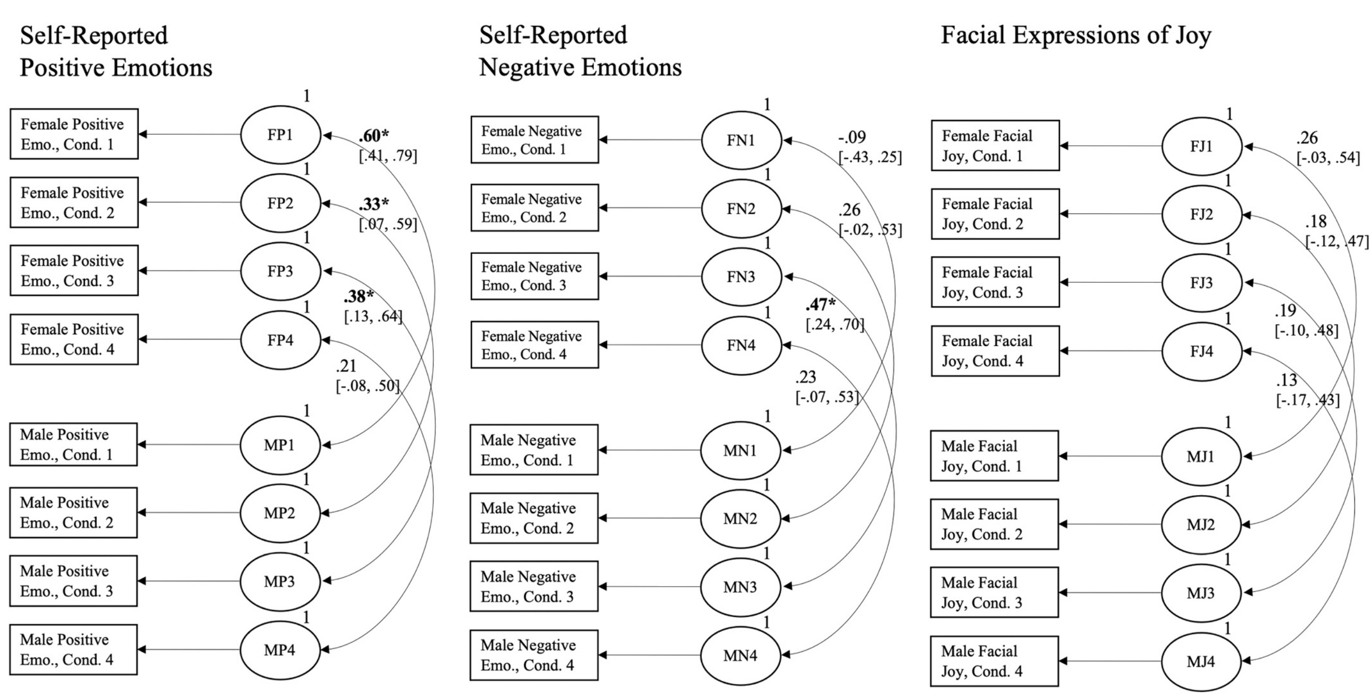

Fig. 2. SEM results examining the similarities between males and females in self-reported emotions and facial expressions of joy.
Note. * indicates this correlation is statistically significant ($p < .05$). 95% confidence intervals are presented in brackets. Emo. = Emotions, Cond. = Condition (1 = baseline, 2 = external stress, 3 = internal stress, 4 = enjoyment), FP = female positive, MP = male positive, FN = female negative, MN = male negative, FJ = female joy, MJ = male joy. In each of the SEMs, latent variances were fixed to 1 and all latent variables were allowed to correlate with each other. For reasons of parsimony, only the correlations of interest are plotted (see Results section for details).

4.2.3. Facial expressions of joy

We examined similarity in facial expressions of joy in the same way we examined similarity for self-reported emotions. Overall, we found facial expressions of joy were not significantly correlated between males and females for any of the conditions (see Fig. 2).

4.2.4. Overall similarities

Overall, these analyses show that similarities between romantic partners were common for self-reported positive emotions, less common for self-reported negative emotions and uncommon for facial expressions of joy (see Fig. 2, please also note the difference in magnitude between correlations between males' and females' self-reported emotions and the correlation of their facial expressions scores, e.g., for the first condition: $r_{\text{self-reported\_emotions}} = .60$, $r_{\text{facial\_expressions}} = .26$). Thus, our results suggest that differences between similarity based on self-reported emotions and similarity based on facial expression data exist.

4.3. Partner effects

To examine the effects of one partner's facial expressions of joy on their partners' self-reported emotions (partner effects), we estimated Actor-Partner Interdependence Models (APIM; Kenny et al., 2006). We sought to investigate these associations for each conversation; however, to avoid Type I errors and to conserve statistical power to detect significant effects, we decided to examine the effect of facial expressions of joy on self-reported emotions across conditions (vs. separate by condition). We estimated one APIM where facial expressions of joy predicted self-reported positive emotions. We then estimated a second APIM where facial expressions of joy predicted self-reported negative emotions. APIMs were modeled with multilevel modeling using the R package *nlme* (Pinheiro et al., 2020), with Restricted Maximum Likelihood Estimation. Errors were allowed to correlate within couples.

To determine whether females' actor and partner effects differ from the males' actor and partner effects, we ran each APIM twice. In the first run, we modeled the partners as distinguishable and in the second run, we modeled the partners as indistinguishable. We then compared the model fits between runs using the Akaike (AIC) and Bayesian information criteria (BIC).

We found that for both self-reported positive emotions ($AIC_{\text{distinguishable partners}} = 1041.89$, $AIC_{\text{indistinguishable partners}} = 1031.22$; $BIC_{\text{distinguishable partners}} = 1075.75$, $BIC_{\text{indistinguishable partners}} = 1050.08$) and self-reported negative emotions ($AIC_{\text{distinguishable partners}} = 774.85$, $AIC_{\text{indistinguishable partners}} = 761.02$; $BIC_{\text{distinguishable partners}} = 808.62$, $BIC_{\text{indistinguishable partners}} = 779.84$), the AIC as well as the BIC was lower for the model with indistinguishable partners compared with the model with distinguishable partners. This indicates that females' and males' actor and partner effects did not differ from each other. Thus, for the final analyses, we chose the more parsimonious model and present the results from the APIMs with indistinguishable partners (see Fig. 3).

We found no significant actor or partner effect of facial expressions of joy on self-reported positive emotions indicating that one's own and one's partner's facial expressions of joy were independent of self-reported positive emotions. For self-reported negative emotions, we also did not find significant actor or partner effects. However, the actor effect ($p = .07$; i.e., one's own facial expressions of joy predicting one's own self-reported negative emotions) might reach significance in studies with more statistical power.

Overall, our results indicate that facial expressions of joy of one romantic partner have no effect on the self-reported emotions of their partner.

5. Discussion

5.1. Summary

Findings from this study support previous emotion research that used self-reported data as we (a) also found that self-reported emotions changed when romantic partners discussed stressful topics (Gottman, 1994) and (b) found that romantic partners were similar in their self-reported emotions (Butner et al., 2007). However, as expected, we found that patterns observed for self-reported expressions of emotion cannot be translated to facial expressions of joy easily. Specifically, while we found self-reported emotions changed due to different experimental conditions, this kind of change did not hold for facial

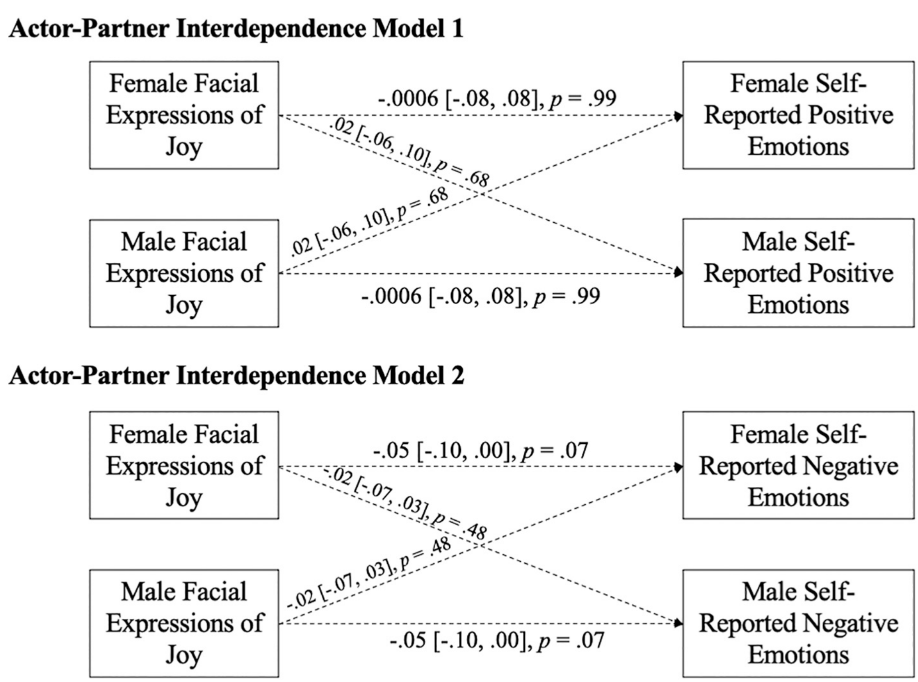

Fig. 3. APIM results predicting positive and negative emotions across conditions using multilevel modeling.
Note. Dotted lines indicate non-significant paths. We tested whether distinguishing between men and women improved model fit (see *Partner Effects* under *Results*). Because it did not, the two actor effects and the two partner effects were set to be equal. 95% confidence intervals are presented in brackets.

expressions of joy. Similarly, we found that emotional similarities between romantic partners were common for self-reported positive emotions and uncommon for facial expressions of joy. Unsurprisingly, the almost stable facial expressions of joy of one romantic partner could not predict the self-reported emotions of the other romantic partner. Our results thus indicate that conclusions on the impact of stressful conversations on emotional experiences, on similarities in emotional experiences, and on emotional partner effects can differ depending on how emotion is measured.

These findings can be interpreted within two perspectives. On the one hand, they could hint to participants not facially showing their felt emotions. This interpretation is in line with the recent conclusion put forward by Barrett et al. (2019) that a person's facial expressions of emotion often do not convey his or her emotional state. As such, facial expressions are different to self-reported emotions, and little is known about when and under which circumstances facial movements express emotions (Barrett et al., 2019). In fact, facial expressions may be displayed for many other reasons apart from communicating one's felt emotion (Krumhuber, Küster, Namba, & Skora, 2021). For example, individuals might intentionally mask their true feelings to prevent conflicts or to not hurt someone else's feelings and this might be especially true for romantic partners (Winterheld, 2017). In this way, one romantic partner's facial expressions might differ greatly from their genuinely felt emotion.

On the other hand, however, previous studies have shown that facial expressions change, for example, due to stress (Mayo & Heilig, 2019). Hence, the current findings could also be interpreted to indicate that the software used was not able to detect small changes in facial expressions of emotions. This interpretation would question the validity and the use of such automated software within these conditions and indeed, such automated software has already been critiqued for not picking up on spontaneous facial expressions (Krumhuber, Küster, Namba, & Skora, 2021). Future research is advised to use different sources to code facial expressions of emotions to see to what degree those different sources match. For example, human coders could be used as they have been found to recognize felt emotions in others at above-chance levels (e.g., Sternglanz & DePaulo, 2004).

5.2. Limitations and future research

One limitation of this study is due to large amounts of missing data from the video recordings due to participants turning their heads away from the camera (i.e., out of video frame). As such, we were unable to assess second-by-second facial emotional data across each conversation; rather, we averaged the scores for facial expressions of emotions within each condition. Average scores, however, have been shown to be robust even with 95% of missing data. Specifically, one simulation study used facial expression scores of joy, simulated different amounts of missing data in the facial expression scores, and found no difference in the resulting means of these scores (Olderbak, Hrycyk, Geiger, Fraude, & Foran, 2021). Thus, although large amounts of missing data existed in our data, mean facial expression scores – as we have used them in this study – seem to be robust.

When interpretating these results, one should also have in mind that the self-reported emotions refer to global retrospective feelings measured after each conversation while facial expressions were measured during each conversation and then averaged. This might have reduced the coherence between self-reported emotions and facial expressions of emotions. Nevertheless, future research is encouraged to examine facial expressions in more detail by assessing second-by-second facial expressions of emotions to detect changes within different conversational conditions.

Further, although this study has based its sample size considerations on sample sizes from similar studies, the overall sample size with 44 romantic couples is rather low. Hence, we cannot exclude the possibility that some of our results might be false positive results. Future studies using larger and more diverse samples should be conducted to corroborate our findings.

Additional limitations include homogeneity of participant demographics related to race and socioeconomic status, and because of the cultural influences on emotion (Barrett et al., 2019), it is important to examine associations between facial expressions of emotions and self-reported emotions in same-gender couples and couples from diverse backgrounds.

5.3. Implications

From a practical point of view, our results challenge the *common view* that posits that facial expressions are indicative of someone's felt emotions (Barrett et al., 2019). Hence, when socially interacting, individuals should keep in mind that facial expressions serve more purposes than just the expression of emotions. Individuals are thus advised to not solely rely on the facial expressions of their counterpart to infer that person's felt emotion.

Further, our results yield important implications for relationship researchers interested in partners' interpersonal emotion dynamics. Specifically, our results challenge the utility of collecting *only* facial expression emotion data as an indicator of emotion (Barrett et al., 2019; Reisenzein et al., 2013). As evidenced in our study, individuals may be experiencing a range of emotions that are not as obvious or detectable through general facial expressions. As a result, the conclusions relationship researchers draw when studying emotions may depend on how emotion is measured.

With modern software, facial expressions of emotion can be assessed automatically. However, the accuracy of such automatic coding needs to be examined more carefully. If those automatic codings then prove to be valid mechanisms to detect facial expressions, this would open possibilities for the creation of real-time assessment models for diagnostic purposes and targeted interventions. To improve our capacity to detect meaningful patterns, what is ultimately required is an approach that integrates additional automatically extracted behavioral signals that carry emotional content, such as acoustic, linguistic, physiological, and bodily movements (D'Mello, Kappas, & Gratch, 2017; Grafsgaard, Duran, Randall, Tao, & D'Mello, 2018). There is growing evidence that such an approach is strongly associated with self-reports and external ratings of affective states (Calvo & D'Mello, 2010).

CRediT authorship contribution statement

Katja M. Pollak: Formal analysis, Writing – original draft, Writing – review & editing. **Sally G. Olderbak:** Formal analysis, Writing – original draft, Writing – review & editing. **Ashley K. Randall:** Conceptualization, Methodology, Writing – review & editing. **Kevin K.H. Lau:** Formal analysis, Writing – review & editing. **Nicholas D. Duran:** Conceptualization, Methodology, Writing – review & editing.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.paid.2021.111182.

References

Anderson, C., Keltner, D., & John, O. P. (2003). Emotional convergence between people over time. *J. Pers. Soc. Psychol., 84*(5), 1054–1068.

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychol. Sci. Public Interest, 20*(1), 1–68. https://doi.org/10.1177/1529100619832930.

Bodenmann, G. (2007). Multidimensional stress scale for couples. In *Unpublished scale*. Switzerland: University of Zurich.

Borders, J., Randall, A. K., & Bodenmann, G. (2016). *Partner's external stress inventory. Unpublished scale*. Arizona State University.

Butner, J., Diamond, L. M., & Hicks, A. M. (2007). Attachment style and two forms of affect coregulation between romantic partners. *Pers. Relat., 14*(3), 431–455.

Calvo, R. A., & D'Mello, S. K. (2010). Affect detection: An interdisciplinary review of models, methods, and their applications. *IEEE Trans. Affect. Comput., 1*(1), 18–37. https://doi.org/10.1109/T-AFFC.2010.1.

D'Mello, S., Kappas, A., & Gratch, J. (2017). The affective computing approach to affect measurement. *Emot. Rev., 10*, 174–183. https://doi.org/10.1177/1754073917696583.

Emotient. (2016). FACET 2.0 performance evaluation. http://imotions.com.

Fernández-Dols, J. M., & Crivelli, C. (2013). Emotion and expression: Naturalistic studies. *Emot. Rev., 5*(1), 24–29. https://doi.org/10.1177/1754073912457229.

Geist, R. L., & Gilbert, D. G. (1996). Correlates of expressed and felt emotion during marital conflict: Satisfaction, personality, process, and outcome. *Personal. Individ. Differ., 21*(1), 49–60.

Gonzaga, G. C., Campos, B., & Bradbury, T. (2007). Similarity, convergence, and relationship satisfaction in dating and married couples. *J. Pers. Soc. Psychol., 93*(1), 34–48.

Gottman, J. M. (1994). *What predicts divorce? The relationship between marital processes and marital outcomes*. Lawrence Erlbaum Associates, Inc.

Gottman, J. M., Levenson, R. W., Gross, J., Frederickson, B. L., McCoy, K., Rosenthal, L., … Yoshimoto, D. (2003). Correlates of gay and lesbian couples' relationship satisfaction and relationship dissolution. *J. Homosex., 45*(1), 23–43.

Grafsgaard, J. F., Duran, N. D., Randall, A. K., Tao, C., & D'Mello, S. K. (2018). *Generative Multimodal Models of Nonverbal Synchrony in Close Relationships*. IEEE.

Griffin, W. A., & Li, X. (2016). Using Bayesian nonparametric hidden semi-markov models to disentangle affect processes during marital interaction. *PLoS One, 11*(5), Article e0155706. https://doi.org/10.1371/journal.pone.0155706.

Kenny, D. A., & La Voie, L. (1984). The social relations model (L. Berkowitz (ed.); Vol. 18, pp. 141–182). Academic Press. https://doi.org/10.1016/S0065-2601(08)60144-6.

iMotions (2016). Emotient SDK 4.1 [computer software]. Copenhagen, Denmark. Retrieved March 2016.

Kenny, D. A., Kashy, D. A., & Cook, W. L. (2006). *Dyadic Data Analysis*. Guilford Press.

Kirschbaum, C., Pirke, K.-M., & Hellhammer, D. H. (1993). The trier social stress test – a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology, 28*, 76–81.

Kline, T. (2005). *Psychological testing: A practical approach to design and evaluation*. Sage.

Lau, K. H., Randall, A. K., Duran, N. D., & Tao, C. (2019). Examining the effects of couples' real-time stress and coping processes on interaction quality: Language use as a mediator. *Frontiers in Psychology, 9*, 2598. https://doi.org/10.3389/fpsyg.2018.02598.

Lerner, J. S., Dahl, R. E., Hariri, A. R., & Taylor, S. E. (2007). Facial expressions of emotion reveal neuroendocrine and cardiovascular stress responses. *Biol. Psychiatry, 61*(2), 253–260.

Mayo, L. M., & Heilig, M. (2019). In the face of stress: Interpreting individual differences in stress-induced facial expressions. *Neurobiology of Stress, 10*, Article 100166. https://doi.org/10.1016/j.ynstr.2019.100166.

Olderbak, S., Hildebrandt, A., Pinkpank, T., Sommer, W., & Wilhelm, O. (2014). Psychometric challenges and proposed solutions when scoring facial emotion expression codes. *Behav. Res. Methods, 46*(4), 992–1006. https://doi.org/10.3758/s13428-013-0421-3.

Olderbak, S., Hrycyk, L., Geiger, M., Fraude, I., & Foran, H. (2021). Using Automated Facial Emotion Coding to Understand Couple Processes. *Manuscript in preparation*.

Parkinson, B. (2011). Interpersonal emotion transfer: Contagion and social appraisal. *Soc. Personal. Psychol. Compass, 5*(7), 428–439.

Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R Core Team. (2020). *nlme: Linear and Nonlinear Mixed Effects Models. R package version 3.1–148.* https://CRAN.R-project.org/package=nlme.

R Development Core Team. (2001). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Randall, A. K., & Schoebi, D. (2018). *Interpersonal emotion dynamics in close relationships*. Cambridge University Press.

Randall, A. K., Tao, C., Leon, G., & Duran, N. D. (2021). Couples' co-regulation dynamics as a function of perceived partner dyadic coping. *Anxiety, Stress & Coping*. in press.

Reed, R. G., Randall, A. K., Post, J. H., & Butler, E. A. (2013). Partner influence and in-phase versus anti-phase physiological linkage in romantic couples. *Int. J. Psychophysiol., 88*(3), 309–316. https://doi.org/10.1016/j.ijpsycho.2012.08.009.

Reisenzein, R., Studtmann, M., & Horstmann, G. (2013). Coherence between emotion and facial expression: Evidence from laboratory experiments. *Emot. Rev., 5*(1), 16–23. https://doi.org/10.1177/1754073912457228.

Schuetzenmeister, A., & Dufey, F. (2018). VCA: Variance component analysis. *R package version, 1*(3), 4.

Shavelson, R. J., Webb, N. M., & Rowley, G. L. (1989). Generalizability theory. *Am. Psychol., 44*(6), 922–932.

Sternglanz, R. W., & DePaulo, B. M. (2004). Reading nonverbal cues to emotions: The advantages and liabilities of relationship closeness. *J. Nonverbal Behav., 28*(4), 245–266. https://doi.org/10.1007/s10919-004-4158-7.

Winterheld, H. A. (2017). Hiding feelings for whose sake? Attachment avoidance, relationship connectedness, and protective buffering intentions. *Emotion, 17*(6), 965–980. https://doi.org/10.1037/emo0000291.

Krumhuber, E.G., Küster, D., Namba, S. & Skora, L. (2021). Human and machine validation of 14 databases of dynamic facial expressions. Behav. Res. Methods, 53, 686–701 (2021). https://doi.org/10.3758/s13428-020-01443-y.