Check for updates

Tutorial

Advances in Methods and Practices in Psychological Science
July-September 2026, Vol. 9, No. 3, pp. 1–10
© The Author(s) 2026
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/25152459261477396
www.psychologicalscience.org/AMPPS

Pairmutate: A Web Application for Testing Dyadic Associations Through Comparison With Pseudo Dyads

Chiara Carlier, Liesse Frérart, Peter Kuppens, and Eva Ceulemans
Department of Psychology and Educational Sciences, KU Leuven, Leuven, Belgium

Abstract

Recently, a lot of attention in the behavioral sciences has been going to how feelings, thoughts, or behaviors overlap, covary, or synchronize across time within dyadic relationships, such as romantic couples or parent–child pairs. This is often investigated by computing and testing an association measure within each dyad. However, two individuals may demonstrate similar responses and thus association simply because they are in the same experimental conditions, have similar social or cultural backgrounds, or just share basic biological functions rather than that the observed similarity can be attributed to true interpersonal processes. This "common input" phenomenon is known to increase the probability of false positives when testing dyadic associations. Permutation testing allows one to account for it by comparing associations based on data from true dyads with associations based on data from pseudo dyads. Pseudo dyads are formed by coupling partners from different dyads in the sample. Although this approach has been proven valuable in dyadic research, it is conceptually and computationally complex and thus not readily accessible. In this tutorial, we explain the logic behind testing dyadic associations using pseudo dyads and introduce and demonstrate a newly developed web application called "Pairmutate" for this purpose.

Keywords

dyads, permutation, testing, associations, synchrony, software, shiny

Received 12/19/25; revision accepted 7/24/26

A central tenet in relationship research is that the feelings, thoughts, behaviors, or central or peripheral physiology of people involved in a relationship line up. This has been conceptualized as synchronization, covariation, or interdependence (Butler, 2011, 2015) and has been documented in, for instance, romantic partners, parent–child dyads, patients and their therapists, and so on (see e.g., Bar-Kalifa et al., 2023; Bodner et al., 2018; Coutinho et al., 2019; Lee et al., 2017; Sels et al., 2016; Sels, Ruan, et al., 2020; Tschacher & Meier, 2020; Weidmann et al., 2016). Dyadic associations are often investigated by computing association measures between variables measured in these persons across time and testing their statistical significance (e.g., Moulder et al., 2018; Ramseyer & Tschacher, 2011; Sels, Cabrieto, et al., 2020).

A well-known problem, however, is that the computed association between data from two different individuals may be contaminated by methodological or situational artifacts such that two individuals may demonstrate similar feelings, thoughts, or behaviors simply because they are in the same experimental conditions, have similar social or cultural backgrounds, or just share basic biological functions rather than it can be attributed to the relationship these individuals share. This phenomenon has been discussed in varying fields as the "common input problem," "stereotype accuracy," or "pseudo synchrony" and is known to increase the probability of false positives in detecting dyadic associations (Burgess, 2013; Hamilton, 2020; Kenny et al., 2006; Moulder et al., 2018).

Corresponding Authors:
Chiara Carlier, Department of Psychology and Educational Sciences, KU Leuven, Leuven, Belgium
Email: chiara.carlier@kuleuven.be

Liesse Frérart, Department of Psychology and Educational Sciences, KU Leuven, Leuven, Belgium
Email: liesse.frerart@kuleuven.be

Creative Commons NonCommercial CC BY-NC: This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 License (https://creativecommons.org/licenses/by-nc/4.0/), which permits noncommercial use, reproduction, and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).

A strong way to detect and control for such artifacts when conducting statistical inference is permutation testing, in which the associations in true dyads are compared with the associations observed in pseudo dyads. These pseudo dyads are formed by randomly coupling partners from different dyads. This random-coupling strategy keeps the common input intact but eliminates the contribution of the dyadic relationship that exists between true partners. The association between true dyads will thus be declared significant only if their dyadic relationship adds additional association strength on top of what is expected based on the common input as can also be observed in pseudo couples.

Although conceptually elegant, correctly conducting permutation testing is complex and requires researchers to program the procedure themselves. To lower the threshold to perform permutation testing in dyadic research, in the present article, we aim to (a) explain the key principles of permutation testing for dyadic associations and potential challenges, such as multiple measurements per person and the sometimes required manipulations of the order of these measurements to retain the common input and (b) introduce an accompanying easy-to-use web application called "Pairmutate" to perform these tests. We illustrate this with two dyadic-data examples in which we investigate the significance of the individual dyadic associations and the sample average, respectively.

Example Data

To illustrate the permutation testing and usage of the Pairmutate app throughout the article, we make use of two empirical examples, referred to as the "scenario data" and the "experience-sampling data." The scenario data include 10 romantic couples[1] that engaged in a computer task in which they read 24 scenarios developed to elicit a variety of negative emotions. Partners separately rated to which extent they expected to feel a fixed list of emotions in those circumstances (in an order randomized per participant) on a scale from 0 (*not at all*) to 100 (*very much*). For this illustration, we want to investigate whether there are individual differences in how partners' irritation covaries over situations and identify for which (types of) dyads this is the case.

The experience-sampling data set includes 10 romantic couples who engaged in a dyadic experience-sampling study in which they answered a variety of questions about their current emotions, environment, and recent experiences multiple times a day for a week in response to a notification of an app on their smartphone. Each partner separately filled out these small surveys, but they were both prompted at the same time. For this illustration, we want to investigate whether partners in general rated the same moments as relaxing and more so than can be expected by chance.

Note that both questions (individual differences and general association) are not mutually exclusive but rather complementary, much like an analysis in which both fixed and random effects are investigated. It is, for instance, possible to first test whether there is a significant average association and then test whether this significance holds for most dyads in the sample.

Testing for Relationship-Driven Associations in Dyads Is Not Trivial

To explain why testing for relationship-driven associations in true dyads is not trivial, we first recapitulate how one would usually test, for instance, how irritation covaries over time or across repeated measurements in a specific dyad. One would typically start by quantifying the relationship between the irritation ratings of both partners. This can be done by computing an association index, such as a Pearson product–moment correlation (although many other similarity indices are possible, which may also account for other patterns than linear covariation; for a review, see Carlier et al., 2024), between the irritation ratings of the first partner and the second partner over situations. Next, one wants to know whether the obtained association value is indicative of meaningful relationship-driven similarity, or in other words, (statistically) significant. This can be done by performing a null-hypothesis test. The null hypothesis in this case would be that the association between the irritation ratings of both partners equals zero. When the test results suggest that it is very unlikely that this null hypothesis holds true, it is rejected, meaning that the association between the irritation ratings of both partners differs from zero and is thus significant.

More technically speaking, in this illustration, the Pearson correlation is Fisher-z transformed and compared with the sampling distribution of correlations under the null hypothesis (the similarity is zero) to check whether the observed correlation is likely to stem from this sampling distribution. If all assumptions (e.g., normality, independence, and identical distribution of the scores per variable) are met, the sampling distribution under the null hypothesis would be a normal distribution, centered around a Fisher-z-transformed correlation of zero because this would be the expected correlation when the irritation of both partners would not be linearly related.

Depending on whether the researcher performs a one- or two-sided test, the observed correlation is said to be significant if it is equal to or more extreme than the left, right, or middle 95% of this sampling distribution. Here, significance means that the probability is very

small that the observed correlation could have been observed when there is no linear relation between the irritation ratings of the partners; hence, the hypothesis that the true correlation between partners' irritation is zero is very unlikely. The sampling distribution under the null hypothesis is thus a crucial ingredient of the hypothesis test because this will define which values of the correlation are significant and which values are not.

However, when testing association strength in dyads, some of the required assumptions might be violated, altering the expected sampling distribution under the null hypothesis. Take, for instance, the case of our scenario data. All participants were presented the same scenarios, developed to elicit similar emotions in every one of them. This causes the assumption of identically distributed irritation scores per partner to be violated. Rather, the irritation scores are more likely drawn from scenario-specific distributions (i.e., characterized by different means) because participants will rate some scenarios as more irritating than others. This, in turn, will affect the mean of the sampling distribution under the null hypothesis, which will no longer be zero. Rather, the more the scenario-specific distributions differ in means, the more positive the mean of the sampling distribution is. This problem is often labeled the "common-input problem" because a correlation between partners may arise solely from being exposed to the same input instead of arising from interpersonal interactive effects between the partners (Moulder et al., 2018; Ramseyer & Tschacher, 2011).

In addition to the common-input problem, both partners have to provide multiple ratings across time. This may cause the assumption of independent irritation ratings to be violated as well. Ratings from the same person could be related through a certain response tendency or become serially related when ratings closer in time are more alike. When such (serial) dependence is present in the data, the variance of the expected sampling distribution under the null hypothesis might increase, which inflates Type-1-error rates (Bodner et al., 2021).

Permutation Testing to Account for the Common-Input Problem

As a solution to these issues, permutation-testing strategies have been proposed (Corsini, 1956; Helm et al., 2012; Kenny et al., 2006; Moulder et al., 2018). Assuming that data have been obtained in multiple dyads, these strategies allow the user to build a correct sampling distribution (i.e., the sampling distribution of the association values if the partners did not interact and/or had no dyadic relationship) by permuting (i.e., shuffling) the data in such a way that the dyadic relationship is destroyed (e.g., being lovers, roommates, coworkers) while all other data characteristics one has information on can be retained (e.g., common input, [serial] dependence; Good, 2006; Moulder et al., 2018). More specifically, one builds a sampling distribution by randomly coupling the data of partners that belong to different couples, yielding so-called "virtual twins," "pseudo couples," or "pseudo dyads." Because of being exposed to common input, these pseudo dyads can show nonzero dyadic associations as well. By comparing the observed association in true dyads with that from pseudo dyads, one can conclude whether the observed dyadic association is significantly stronger than what can be expected on top of the common input. However, how exactly this comparison is done depends on the specific research question at hand.

In the scenario data, we wanted to investigate whether there are individual differences in how partners' irritation covaries over situations (on top of the effect of common input). Therefore, the sampling distribution pertains to association values of individual couples. To obtain this distribution, we scrutinize all possible 10 × 9 pseudo dyads because every Partner 1 can be coupled with nine other Partner 2s. By always joining a Partner 1 with a Partner 2, the distinguishability of partners (i.e., in terms of gender) is taken into account. For each true couple, the number of pseudo dyads with an equal or more extreme association score than the true couple will be divided by the total number of pseudo dyads, yielding a p value.[2] If this p value is smaller than a chosen significance level (e.g., .05), the association in the true couple is said to be significantly different from that in the pseudo couples (Fig. 1). All obtained p values are one-sided. Performing two-tailed tests is not straightforward in our setting because the null distribution cannot be expected to be symmetrical around zero. If researchers prefer to conduct a more conservative "two-tail-like" test, they can opt to use a stricter significance level and thus compare the observed association value with a more extreme percentile of the null distribution.

For the experience-sampling data, we were interested in whether partners in general rated the same moments as relaxing more than what can be expected by chance; hence, the sampling distribution should consist of average associations. This sampling distribution will be based on samples of pseudo dyads of equal size as the sample with true dyads (Fig. 1). For each sample, a mean is computed, and to make sure that each mean has a similar weight, all samples must be of equal size. Therefore, a high number (i.e., 1,000) of samples of 10 pseudo dyads will be created to be able to compare the average association value in the original sample with a sampling distribution of average association values in samples of pseudo dyads.[3] To obtain a p value, we divide the number of pseudo-couple samples that yield an equal or

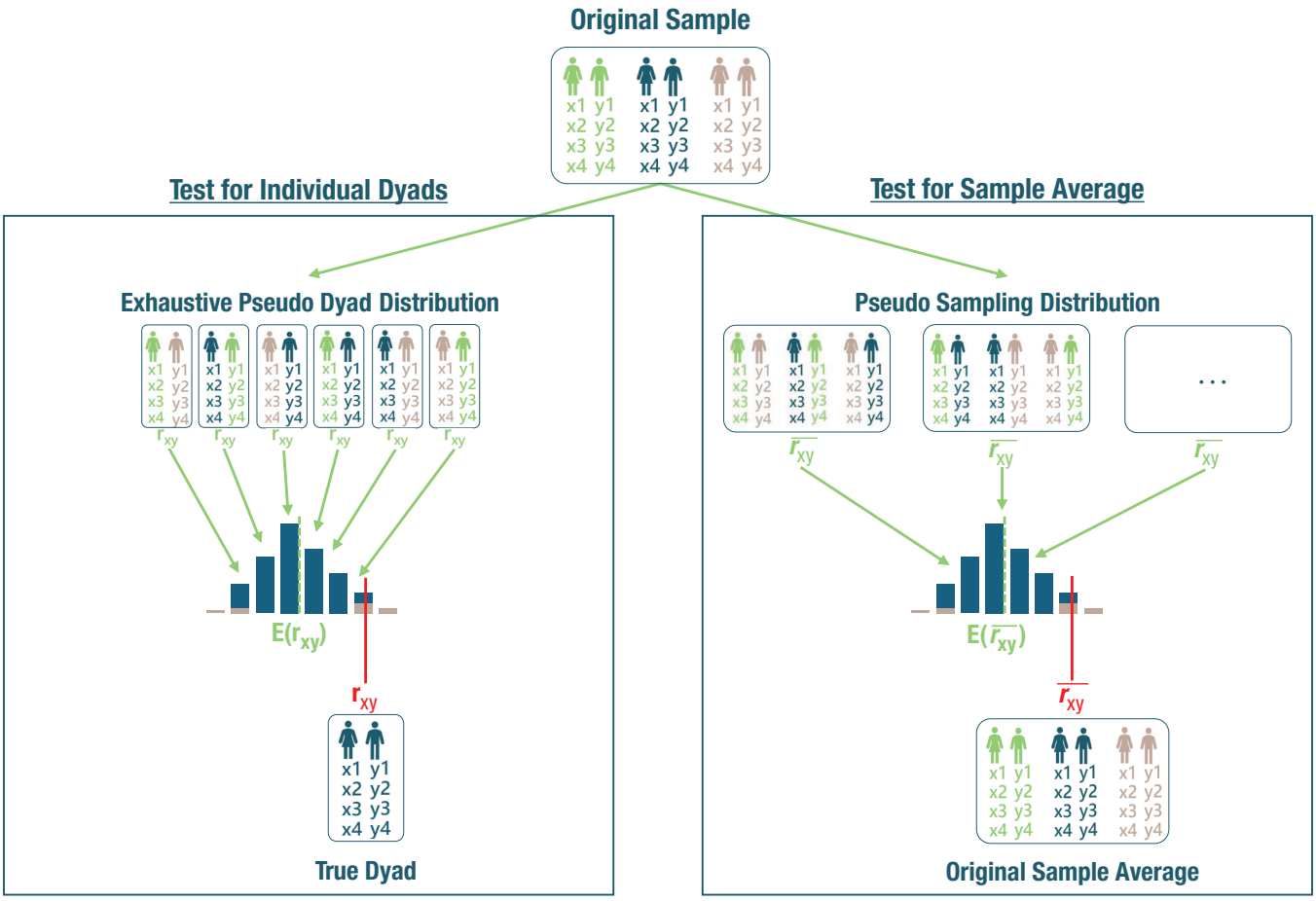


Fig. 1. Permutation testing of a Pearson product–moment correlation. (Left) The inference process for the observed correlation for one dyad. (Right) The inference procedure for the average correlation within the sample.

more extreme average association than the original average association by the total number of samples of pseudo dyads. Here again, all obtained *p* values are one-sided.

In both our examples, partners of different dyads are randomly coupled while preserving the original order in which their ratings were stored. This means that the first observation of Partner 1 is coupled to the first observation of Pseudo Partner 2 and so on. For chronologically ordered data, this ensures that the serial dependence between the ratings is also maintained in the pseudo dyads; hence, the Type-I-error rate of the test will not be inflated because of serial dependence. However, there are cases in which the order of the ratings requires some further attention to fully account for common-input problems.

When Common Input Is Differently Ordered Across Dyads

In the experience-sampling data, one could think of the days of the week as one source of common input because people in general might be more relaxed on weekends than on weekdays (Lavee & Ben-Ari, 2007). Yet dyads started the study on different days. If one would keep the ratings of participants in their original order while randomly assembling pseudo dyads, a Tuesday rating could be matched with a Saturday rating, and the possible common input of week versus weekend days is thus not controlled for (Fig. 2). To account for this, one must first reorder the days of the participant to align with those of the new pseudo partner. Within each day, the original order (and thus serial dependence) of the ratings is still retained (Fig. 2).

More specifically, as shown on the right side of Figure 2, a partner that started on Friday is linked to a partner that started on Monday. From the partner that started on Friday, the data from Monday to Thursday are moved to the start of the time series to match the day order of the person who started on Monday. The data from Friday to Sunday are then appended after these days. This is, of course, under the assumption that all participants have the same number of days. If, for

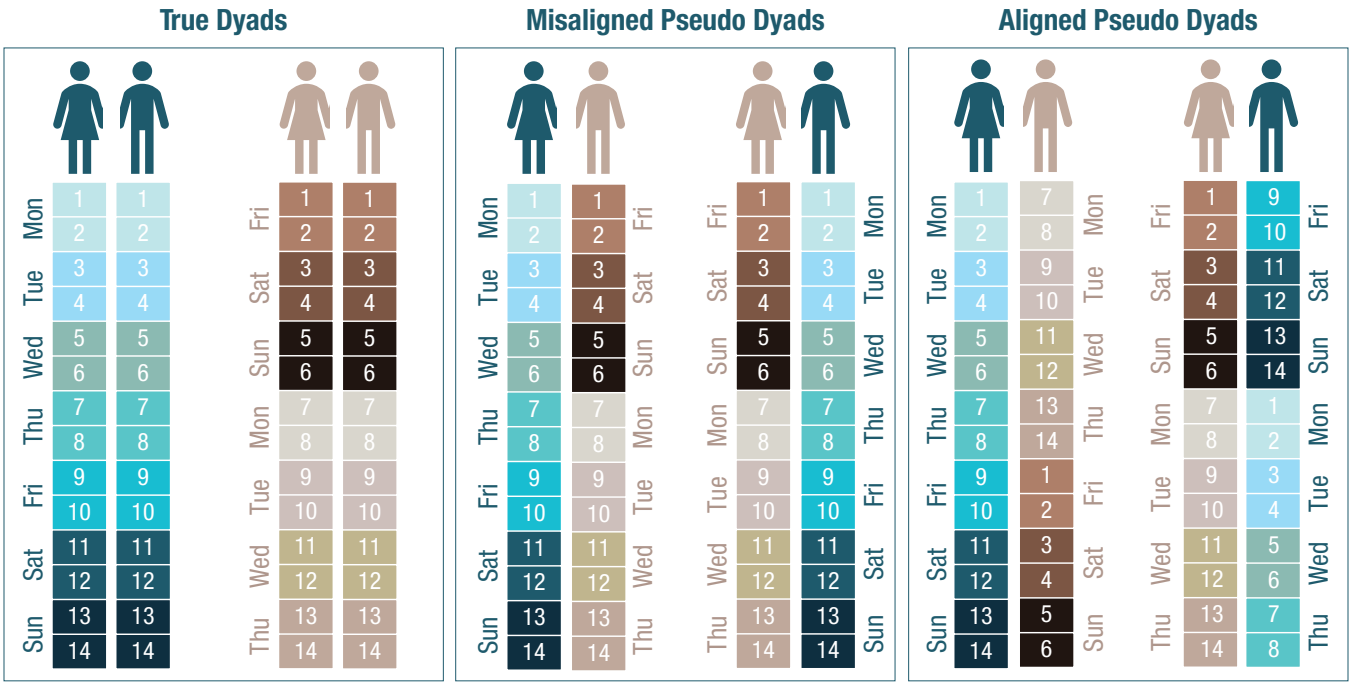


Fig. 2. Controlling for the common input by aligning the days of the week between pseudo partners. Blue versus brown colors represent true-dyad memberships, color shades represent days of the week, and numbers represent observation numbers.

instance, one participant stopped the study after 6 days instead of 7 days, then the 7th day will be removed for the pseudo partner as well.

Likewise, in experimental studies, counterbalanced conditions may need to be realigned when forming pseudo dyads. In the remainder of this article and in the Pairmutate app, we refer to blocks of data that need to be aligned similarly within pseudo partners (e.g., weekdays, conditions, trials) as "segments."

A Dedicated Software Application: Pairmutate

Our examples showcase that testing associations in dyads can be challenging because of the common-input, serial-dependence, and alignment issues and that using classical statistical software can lead to incorrect inferences. To help researchers navigate this challenge, we developed a dedicated user-friendly Shiny application, Pairmutate, that facilitates conducting permutation analyses to evaluate dyadic associations with the option to align segments to control for common input.

Pairmutate can be accessed at ppw-okpiv.pages.gitlab.kuleuven.be/researchers/u0149080/pairmutate and consists of three main sections: data import, analysis, and results. In the data-import section, users can choose between repeated-measures data with a fixed order or with need for alignment and perform a check of their variables. In the analysis section, users can select the desired permutation test (individual dyads or sample average) and download the permuted data set. Finally, the results section displays the significance-test results for the selected variables and association measure and the associate sampling distribution. Results can be obtained for three popular association measures: Pearson product–moment correlation, Spearman's ρ, and mean absolute difference (Carlier et al., 2024; Cronbach & Gleser, 1953).

Data import

The uploaded data table must be a .csv file containing the data in a dyadic format. This implies that each row corresponds to a measurement occasion for the full dyad, containing scores on all relevant variables. If the data contain multiple measurement occasions, all should preferably be included in the data file (e.g., Beep Number 1–10 or Stimulus 1–20) even if some have missing values for one or more of the variables of interest (Table 1). Users must also specify whether their data are in a fixed order (e.g., every participant has ratings for Stimulus 1–20 stored in the same order) or whether their data need alignment (e.g., participants started the study on different days or went through conditions in a different order). Depending on the indicated data type, the uploaded data table must contain a certain set of columns with designated column names. Table 1 displays the required columns for the needs-alignment option:

Table 1. Excerpt of the Experience-Sampling Data

dyadnr	obsnr	segmnr	segm	segmObsnr	ppnr.1	relaxed.1	ppnr.2	relaxed.2
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .
2	34	4 (Day 4)	1 (Monday)	4 (Observation 4 of Day 4)	2	25	102	NA
2	35	4	1	5	2	13	102	NA
2	36	4	1	6	2	8	102	18
2	37	4	1	7	2	11	102	39
2	38	4	1	8	2	28	102	66
2	39	4	1	9	2	70	102	37
2	40	4	1	10 (Observation 10 of Day 4)	2	66	102	84
2	41	5 (Day 5)	2 (Tuesday)	1 (Observation 1 of Day 5)	2	40	102	26
2	42	5	2	2	2	24	102	75
2	43	5	2	3	2	67	102	19
2	44	5	2	4	2	62	102	16
2	45	5	2	5	2	29	102	12
2	46	5	2	6	2	NA	102	20
2	47	5	2	7 (Observation 7 of Day 5)	2	21	102	25
. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .	. . .

Note: dyadnr contains the identification numbers given to the true dyads in this study; obsnr represents the measurement-occasion number; segmnr represents the segment number per dyad (in this example, 5 = Day 5 and thus fifth segment of the study for this dyad, etc.); segm contains the identification numbers of each segment (in this example, 1 = Monday, 2 = Tuesday, etc.); segmObsnr represents the measurement-occasion number within a segment and starts anew with every new segment; ppnr.1 and ppnr.2 contain the identification numbers given to Partner 1 and Partner 2 of each true dyad, respectively; [variables].1 and [variables].2 are the variables of interest for Partners 1 and Partners 2, respectively (in this example, the variable of interest is "relaxed" for both partners).

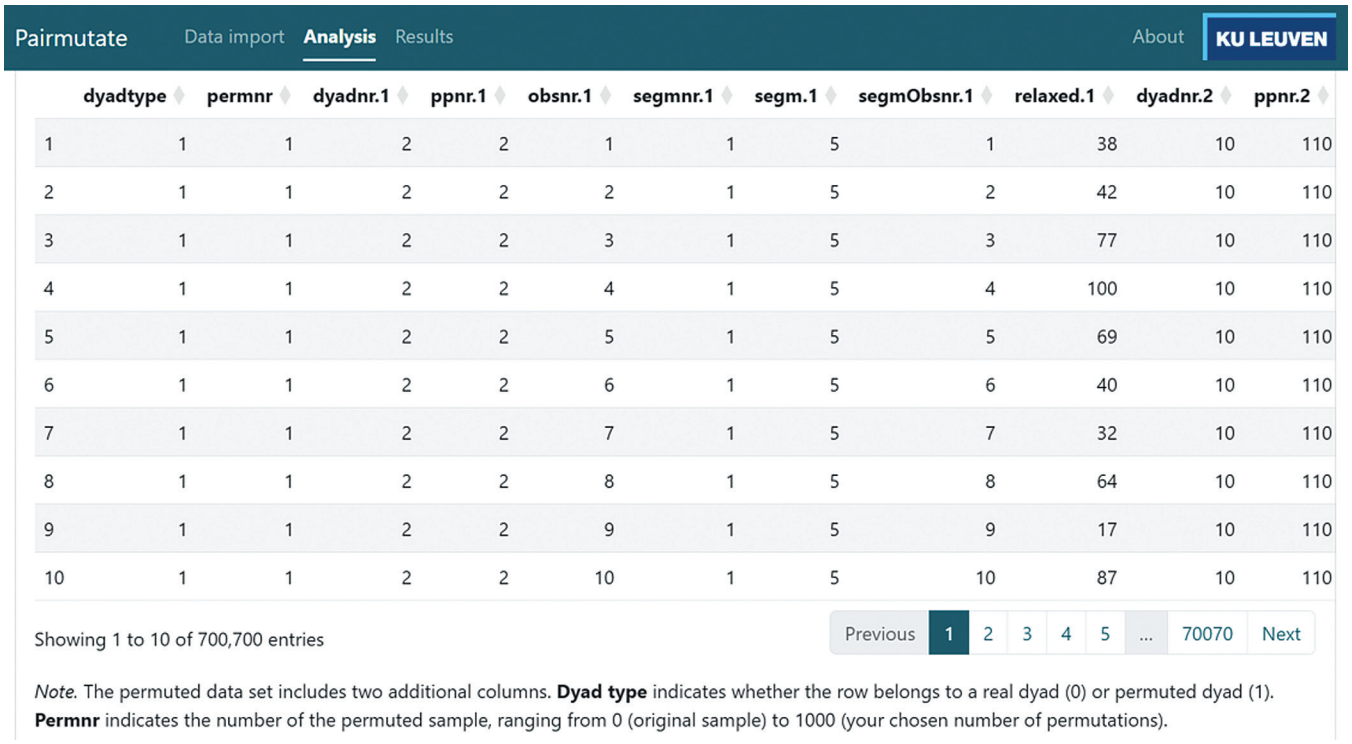
For the fixed-order option, the same columns are required except for the three segment ones. This means that the researcher has to carefully preprocess the data beforehand to resolve any anomalies and to put the data table in the correct format, for example, by using the tools by Revol et al. (2023). Before proceeding, the user must review the characteristics of the uploaded data to identify any potential issues or unexpected patterns (e.g., required columns are missing or are in an incorrect format).

Turning to our example data, the scenario data contain all ratings for Scenarios 1 to 24 in the same order for all dyads; hence, the first option, "fixed order," is indicated. For the experience-sampling data, dyads started on different days of the week (e.g., segm 5 for dyad 2 but segm 2 for dyad 10). These days need to be aligned before forming pseudo dyads to control for common input; hence, the second option, "needs alignment," is selected. If all participants had started the study on Monday and ended on Sunday, no alignment would be needed, and the option "fixed order" could be selected. Both example data sets made use of a semicolon as separator, which should be indicated when uploading the data.

Analysis

In the "Analysis" section of the app, the researcher selects whether they want to test the individual dyadic associations or the sample average. To recapitulate, when testing the individual dyads, all possible pseudo dyads will be scrutinized to provide a p value for each true dyad; testing for the sample average pertains to creating N samples of the same number of pseudo dyads to provide a p value for the sample average. For the latter, users need to provide the desired number of permutation samples, which is usually a high number, such as 1,000, and choose a number (called "seed") to be able to recreate this pseudorandom process and reproduce the results. The duration of compilation of the new permuted data set depends on, among other things, the size of the original data set, the desired test, and when testing the sample average, the number of permutations. Upon completion, the permuted data set is displayed, and users can either download it as a .csv file for further analysis in their preferred software or proceed to the "Results" section of the app.

For the scenario data, the test for individual dyads was indicated because we were interested in the individual differences in the covariation of irritation ratings and thus would prefer an analysis per dyad. Because the number of possible pseudo dyads exponentially increases with the number of dyads in the sample, the run time of this permutation can get quite long with larger samples. Therefore, the app allows the choice between the exhaustive sample of pseudo dyads (more complete but takes longer) or a self-provided number of desired pseudo dyads (reduces processing time), which will then be subsampled from the exhaustive list.



	dyadtype	permnr	dyadnr.1	ppnr.1	obsnr.1	segmnr.1	segm.1	segmObsnr.1	relaxed.1	dyadnr.2	ppnr.2
1	1	1	2	2	1	1	5	1	38	10	110
2	1	1	2	2	2	1	5	2	42	10	110
3	1	1	2	2	3	1	5	3	77	10	110
4	1	1	2	2	4	1	5	4	100	10	110
5	1	1	2	2	5	1	5	5	69	10	110
6	1	1	2	2	6	1	5	6	40	10	110
7	1	1	2	2	7	1	5	7	32	10	110
8	1	1	2	2	8	1	5	8	64	10	110
9	1	1	2	2	9	1	5	9	17	10	110
10	1	1	2	2	10	1	5	10	87	10	110

Fig. 3. Screenshot of the Pairmutate app "Analysis" section for the experience-sampling data.

For this illustration, we went with the exhaustive option. The permuted data set includes an extra column, "dyadtype," to indicate whether a row belongs to a real dyad (0) or a pseudo dyad (1; Fig. 3). In the experience-sampling data, the test for the sample average was selected because we wanted to investigate whether partners in general rated the same moments as relaxing; we chose 1,000 permutations and set a reproducibility seed at 123. After permuting, the data set includes two additional columns: "dyadtype" and "permnr." Because the original data set (i.e., the true dyads) is appended to the end of the permuted data set, again the "dyadtype" column is added to indicate whether a row belongs to a real dyad (0) or a pseudo dyad (1). "Permnr" is added to indicate the number of the permuted sample, ranging from 0 (original sample) to 1,000 (the chosen number of permutations; Fig. 3).

Results

In the "Results" section, the user can indicate the variable of interest for each dyad member and select the desired association measure. The computation might take a few minutes. Once completed, the results are displayed with a visualization of the obtained sampling distribution on the left and the significance test results on the right.

In the scenario data, we wanted to test the covariation of irritation ratings per dyad. To assess this, the Pearson product–moment correlation was calculated between "irritated.1" and "irritated.2" because this is a good measure of covariation. The distribution (Fig. 4) displays the distribution of the correlations of all true dyads ($M = .57$; Fig. 4, light brown) and the distribution of the correlations of all possible permuted pseudo dyads ($M = .44$; Fig. 4, dark teal). Note that the mean correlation of the pseudo dyads is rather large, which indicates that common input is indeed at play here. More specifically, some of the scenarios in this study were designed to elicit more irritation than other scenarios such that irritation ratings can be correlated even among strangers. As shown on the right side of Figure 4, four out of the 10 dyads in the scenario data have a significant correlation (at significance level .05), indicating that their association goes beyond the correlation they would have had if they shared only common input and no relationship. This shows that the majority of the dyads did not individually have a significant association. It could now be interesting to see whether Dyads 3, 5, 7, and 9 have something in common that the others dyads do not have.

For the experience-sampling data, we were interested in whether partners in general rated the same moments as relaxing. To do this, the Spearman correlation between "relaxed.1" and "relaxed.2" was computed because this is a more robust measure of similarity. The light brown vertical line in Figure 5 depicts the mean correlation of the original sample as .24. The distribution depicts the sampling distribution of the mean correlations of the 1,000 permuted samples in dark teal in Figure 5. Because we

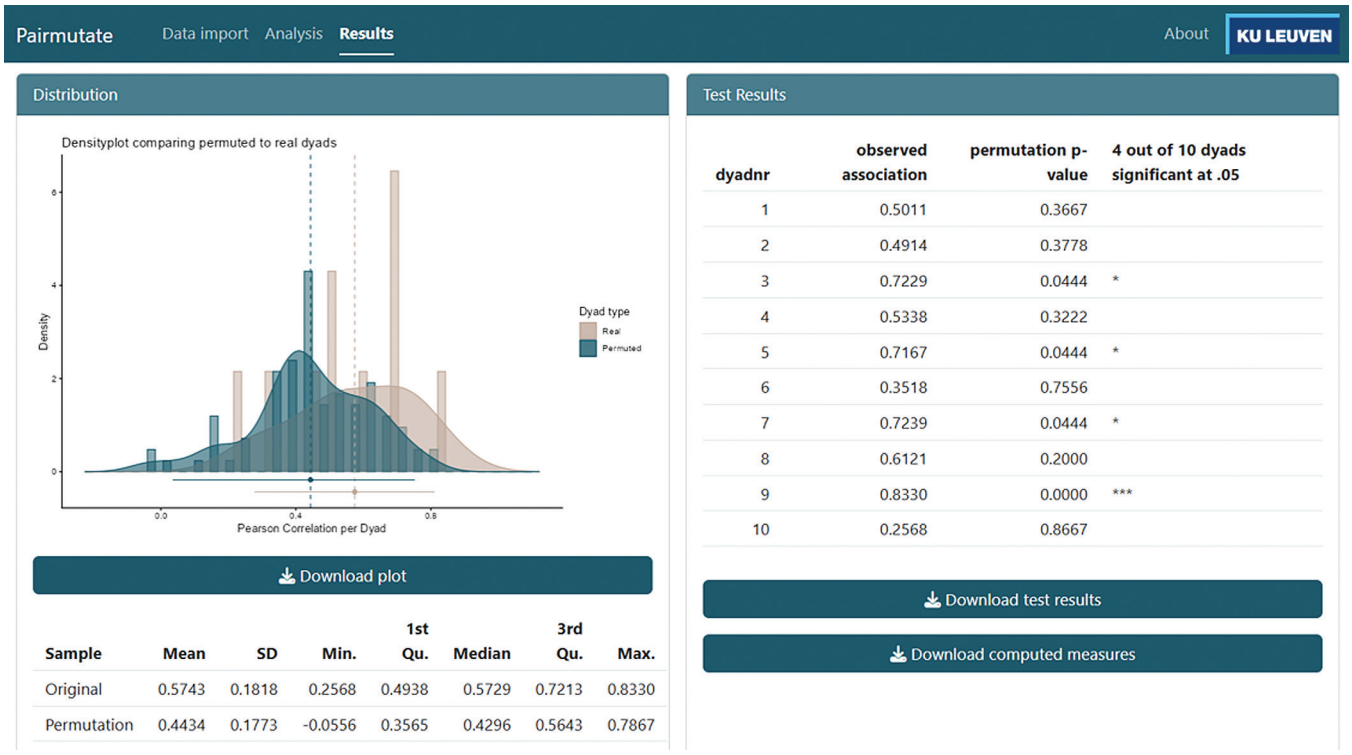


Fig. 4. Screenshot of the Pairmutate app "Results" section for the scenario data.

believe that relaxation is more likely to happen on weekend days versus weekdays and we want to account for this confound, we chose to align segments (days). Figure 5b demonstrates what would have happened if we had not indicated "needs alignment" for this data set but "fixed order." Then, the correlation in pseudo dyads would on average amount to .03, suggesting there is no common input at play. However, when we make sure that the relaxation ratings of the same days of the week are paired, the correlation in pseudo dyads on average becomes .16 (Fig. 5a), showing that on certain days of the week, most participants are a bit more relaxed than on other days. The correct analysis still demonstrates that the mean correlation of relaxation between dyad members is higher in the original sample compared with the permuted samples. In addition, the permutation test of the mean revealed that 30 out of the 1,000 permuted samples had a mean correlation higher than the original mean correlation of .24, implying that $p = .030$. When completing the permutations and analysis, it is possible to start over and continue with a new data set or analysis by clicking on the app title. This will navigate back to the data-import section and allow the user to clear all settings.

Limitations and Future Directions

When designing the app, we had to make several choices regarding which features to implement. As a result, the app has a number of limitations. First, for now, only a few statistical analyses are accommodated in the app, whereas many other dyadic-analysis strategies have been proposed in the literature. It is, however, possible to download the permuted data sets, which allows researchers to perform other dyadic analyses, such as actor-partner-interdependence modeling (Cook & Kenny, 2005) or fitting coupled oscillators (Boker & Nesselroade, 2002) on these permutated data, using other software packages.

Second, it is unclear for now which number of pseudo dyads and, by extension, true dyads is needed to create a stable distribution under the null hypothesis for the individual-dyad analysis. Future research is needed to investigate this issue in detail. However, note that the number of pseudo dyads exponentially increases with the number of true dyads and that a set of 1,000 pseudo dyads can already be reached with 33 true dyads

Conclusion

Common-input and serial-dependence issues in dyadic-association tests can be accounted for by comparing the association in true dyads with that in pseudo dyads. We provided a tutorial and dedicated software for performing the associated permutation analyses. The permutation logic and practical use of the Pairmutate app were illustrated with two real data examples.

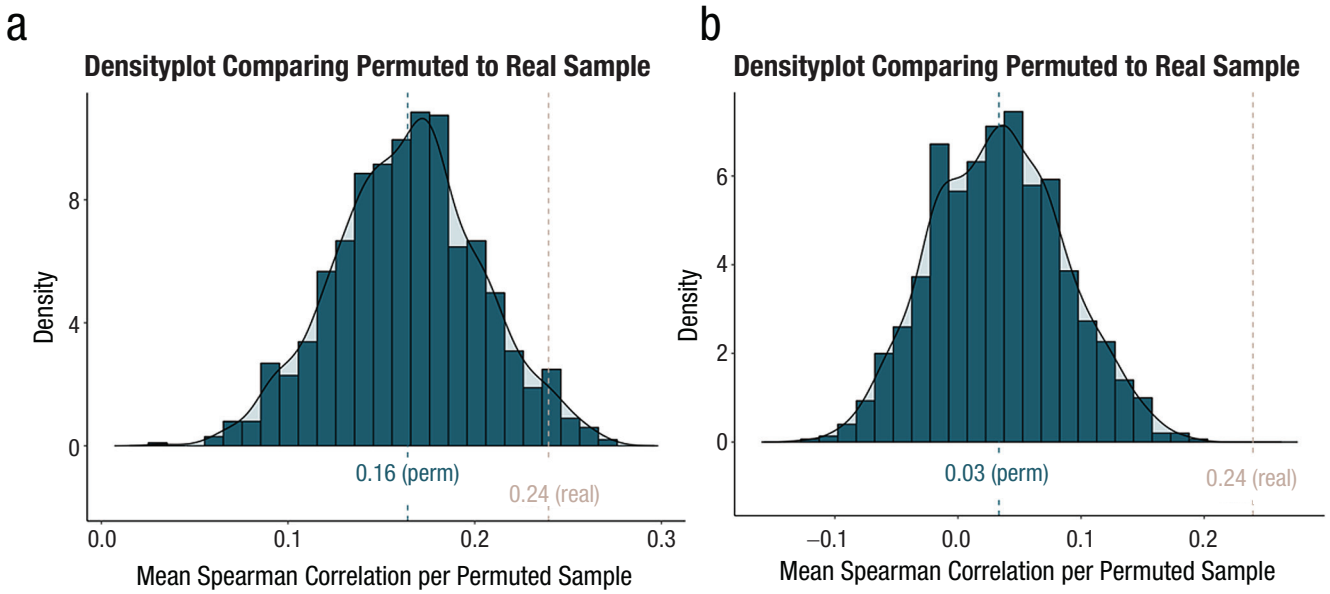


Fig. 5. Sampling distribution of mean correlation for (a) aligned and (b) nonaligned experience-sampling data.

Transparency

Action Editor: David A. Sbarra
Editor: David A. Sbarra
Author Contributions

C. Carlier and L. Frérart are joint first authors.

Chiara Carlier: Conceptualization; Data curation; Funding acquisition; Methodology; Software; Validation; Visualization; Writing – original draft; Writing – review & editing.

Liesse Frérart: Conceptualization; Data curation; Methodology; Software; Validation; Writing – original draft; Writing – review & editing.

Peter Kuppens: Conceptualization; Funding acquisition; Methodology; Supervision; Writing – original draft; Writing – review & editing.

Eva Ceulemans: Conceptualization; Funding acquisition; Methodology; Supervision; Writing – original draft; Writing – review & editing.

Declaration of Conflicting Interests

The authors declared that there were no conflicts of interest with respect to the authorship or the publication of this article.

Funding

The research reported in this article was supported by an FWO fellowship granted to C. Carlier (1153725N), KU Leuven Research Council Grants C14/23/062 and iBOF/21/090 awarded to E. Ceulemans and P. Kuppens, and a research grant from the EOS Excellence of Science program (G0I2422N) awarded to E. Ceulemans.

ORCID iD

Chiara Carlier https://orcid.org/0000-0003-4985-6240

Acknowledgments

We express our sincere gratitude to Kristof Meers for his invaluable help with the deployment of the web application.

All code and example data can be accessed on the Pairmutate GitLab repository: https://gitlab.kuleuven.be/ppw-okpiv/researchers/u0149080/pairmutate. The Pairmutate app can be accessed via https://ppw-okpiv.pages.gitlab.kuleuven.be/researchers/u0149080/pairmutate/.

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/25152459261477396

Notes

1. We selected only 10 couples for this study to make clear that this is just an example of how to use the app and to avoid the idea that any substantive conclusions would be made based on this toy example. However, because readers might be interested in trying out the app with a larger sample, we have added the full scenario data set on GitLab and in the Supplemental Material as well.
2. True dyads are thus compared with pseudo dyads only. This is done to make sure that no dyadic interaction or relationship exists between dyads in the comparison group because this is the subject of the test itself.
3. Likewise, here, pseudo samples can never contain a true dyad by accident; however, the exact same pseudo dyad can sometimes appear in more than one pseudo sample.

References

Bar-Kalifa, E., Goren, O., Gilboa-Schechtman, E., Wolff, M., Rafael, D., Heimann, S., Yehezkel, I., Scheniuk, A., Ruth, F., & Atzil-Slonim, D. (2023). Clients' emotional experience as a dynamic context for client–therapist physiological synchrony. *Journal of Consulting and Clinical Psychology*, *91*(6), 367–380. https://doi.org/10.1037/ccp0000811

Bodner, N., Kuppens, P., Allen, N. B., Sheeber, L. B., & Ceulemans, E. (2018). Affective family interactions and their associations

with adolescent depression: A dynamic network approach. *Development and Psychopathology*, *30*(4), 1459–1473. https://doi.org/10.1017/S0954579417001699

Bodner, N., Tuerlinckx, F., Bosmans, G., & Ceulemans, E. (2021). Accounting for auto-dependency in binary dyadic time series data: A comparison of model- and permutation-based approaches for testing pairwise associations. *British Journal of Mathematical and Statistical Psychology*, *74*(S1), 86–109. https://doi.org/10.1111/bmsp.12222

Boker, S. M., & Nesselroade, J. R. (2002). A method for modeling the intrinsic dynamics of intraindividual variability: Recovering the parameters of simulated oscillators in multi-wave panel data. *Multivariate Behavioral Research*, *37*(1), 127–160. https://doi.org/10.1207/S15327906MBR3701_06

Burgess, A. P. (2013). On the interpretation of synchronization in EEG hyperscanning studies: A cautionary note. *Frontiers in Human Neuroscience*, *7*, Article 881. https://doi.org/10.3389/fnhum.2013.00881

Butler, E. A. (2011). Temporal interpersonal emotion systems: The "TIES" that form relationships. *Personality and Social Psychology Review*, *15*(4), 367–393. https://doi.org/10.1177/1088868311411164

Butler, E. A. (2015). Interpersonal affect dynamics: It takes two (and time) to tango. *Emotion Review*, *7*(4), 336–341. https://doi.org/10.1177/1754073915590622

Carlier, C., Karch, J. D., Kuppens, P., & Ceulemans, E. (2024). A comparison of measures for assessing profile similarity in dyads. *Psychologica Belgica*, *64*(1), Article 1. https://doi.org/10.5334/pb.1297

Cook, W. L., & Kenny, D. A. (2005). The actor–partner interdependence model: A model of bidirectional effects in developmental studies. *International Journal of Behavioral Development*, *29*(2), 101–109. https://doi.org/10.1080/01650250444000405

Corsini, R. J. (1956). Understanding and similarity in marriage. *The Journal of Abnormal and Social Psychology*, *52*(3), 327–332. https://doi.org/10.1037/h0043556

Coutinho, J., Oliveira-Silva, P., Fernandes, E., Gonçalves, O. F., Correia, D., Perrone Mc-Govern, K., & Tschacher, W. (2019). Psychophysiological synchrony during verbal interaction in romantic relationships. *Family Process*, *58*(3), 716–733. https://doi.org/10.1111/famp.12371

Cronbach, L. J., & Gleser, G. C. (1953). Assessing similarity between profiles. *Psychological Bulletin*, *50*(6), 456–473. https://doi.org/10.1037/h0057173

Good, P. I. (2006). *Resampling methods: A practical guide to data analysis* (3rd ed.). Birkhäuser.

Hamilton, A. (2020). *Hype, hyperscanning and embodied social neuroscience*. OSF. https://doi.org/10.31234/osf.io/rc9wp

Helm, J. L., Sbarra, D., & Ferrer, E. (2012). Assessing cross-partner associations in physiological responses via coupled oscillator models. *Emotion*, *12*(4), 748–762. https://doi.org/10.1037/a0025036

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

Lavee, Y., & Ben-Ari, A. (2007). Relationship of dyadic closeness with work-related stress: A daily diary study. *Journal of Marriage and Family*, *69*(4), 1021–1035. https://doi.org/10.1111/j.1741-3737.2007.00428.x

Lee, T.-H., Miernicki, M. E., & Telzer, E. H. (2017). Families that fire together smile together: Resting state connectome similarity and daily emotional synchrony in parent-child dyads. *NeuroImage*, *152*, 31–37. https://doi.org/10.1016/j.neuroimage.2017.02.078

Moulder, R. G., Boker, S. M., Ramseyer, F., & Tschacher, W. (2018). Determining synchrony between behavioral time series: An application of surrogate data generation for establishing falsifiable null-hypotheses. *Psychological Methods*, *23*(4), 757–773. https://doi.org/10.1037/met0000172

Ramseyer, F., & Tschacher, W. (2011). Nonverbal synchrony in psychotherapy: Coordinated body movement reflects relationship quality and outcome. *Journal of Consulting and Clinical Psychology*, *79*(3), 284–295. https://doi.org/10.1037/a0023419

Revol, J., Carlier, C., Lafit, G., Verhees, M., Sels, L., & Ceulemans, E. (2023). *Preprocessing ESM data: A step-by-step framework, tutorial website, R package, and reporting templates*. PsyArXiv. https://doi.org/10.31234/osf.io/hnu2t

Sels, L., Cabrieto, J., Butler, E., Reis, H., Ceulemans, E., & Kuppens, P. (2020). The occurrence and correlates of emotional interdependence in romantic relationships. *Journal of Personality and Social Psychology*, *119*(1), 136–158. https://doi.org/10.1037/pspi0000212

Sels, L., Ceulemans, E., Bulteel, K., & Kuppens, P. (2016). Emotional interdependence and well-being in close relationships. *Frontiers in Psychology*, *7*, Article 283. https://doi.org/10.3389/fpsyg.2016.00283

Sels, L., Ruan, Y., Kuppens, P., Ceulemans, E., & Reis, H. (2020). Actual and perceived emotional similarity in couples' daily lives. *Social Psychological and Personality Science*, *11*(2), 266–275. https://doi.org/10.1177/1948550619845927

Tschacher, W., & Meier, D. (2020). Physiological synchrony in psychotherapy sessions. *Psychotherapy Research*, *30*(5), 558–573. https://doi.org/10.1080/10503307.2019.1612114

Weidmann, R., Ledermann, T., & Grob, A. (2016). The interdependence of personality and satisfaction in couples. *European Psychologist*, *21*(4), 284–295. https://doi.org/10.1027/1016-9040/a000261