Psychometrika (2025), **90**, 1536–1557
doi:10.1017/psy.2025.10040

APPLICATION AND CASE STUDIES - ORIGINAL

Enhancing Empathic Accuracy: Penalized Functional Alignment Method to Correct Temporal Misalignment in Real-Time Emotional Perception

Linh H. Nghiem[1], Jing Cao[2], Chrystyna D. Kouros[3] and Chul Moon[2]

[1]School of Mathematics and Statistics, University of Sydney, Sydney, NSW, Australia; [2]Department of Statistics and Data Science, Southern Methodist University, Dallas, TX, USA; [3]Department of Psychology, Southern Methodist University, Dallas, TX, USA

Corresponding author: Chul Moon; Email: chulm@smu.edu

(Received 13 September 2024; revised 5 August 2025; accepted 6 August 2025; published online 5 September 2025)

Abstract

Empathic accuracy (EA) is the ability to accurately understand another person's thoughts and feelings, which is crucial for social and psychological interactions. Traditionally, EA is assessed by comparing a perceiver's moment-to-moment ratings of a target's emotional state with the target's own self-reported ratings at corresponding time points. However, misalignments between these two sequences are common due to the complexity of emotional interpretation and individual differences in behavioral responses. Conventional methods often ignore or oversimplify these misalignments, for instance by assuming a fixed time lag, which can introduce bias into EA estimates. To address this, we propose a novel alignment approach that captures a wide range of misalignment patterns. Our method leverages the square-root velocity framework to decompose emotional rating trajectories into amplitude and phase components. To ensure realistic alignment, we introduce a regularization constraint that limits temporal shifts to ranges consistent with human perceptual capabilities. This alignment is efficiently implemented using a constrained dynamic programming algorithm. We validate our method through simulations and real-world applications involving video and music datasets, demonstrating its superior performance over traditional techniques.

Keywords: cognitive study; functional data analysis; regularization; square root velocity function; warping function

1. Introduction

The ability to perceive and understand the emotions and thoughts of others, broadly referred to as empathy, plays an important role in human society by facilitating cooperation and social cohesion (De Waal, 2008). While empathy encompasses multiple components, including sharing in another's emotional experience and concern for others, empathic accuracy (EA) refers to the specific skill of accurately inferring what another person is thinking and feeling in a given moment (Ickes, 1997). EA is typically measured behaviorally by comparing a perceiver's rating of a target's emotional state to the target's own self-reported emotional experience. Given its importance to social interactions and quality of life, EA has become a focal point of research across various fields. For example, in social science, EA's role was examined in developing and maintaining healthy social relationships (Sened et al., 2017).

© The Author(s), 2025. Published by Cambridge University Press on behalf of Psychometric Society.
This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (https://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press

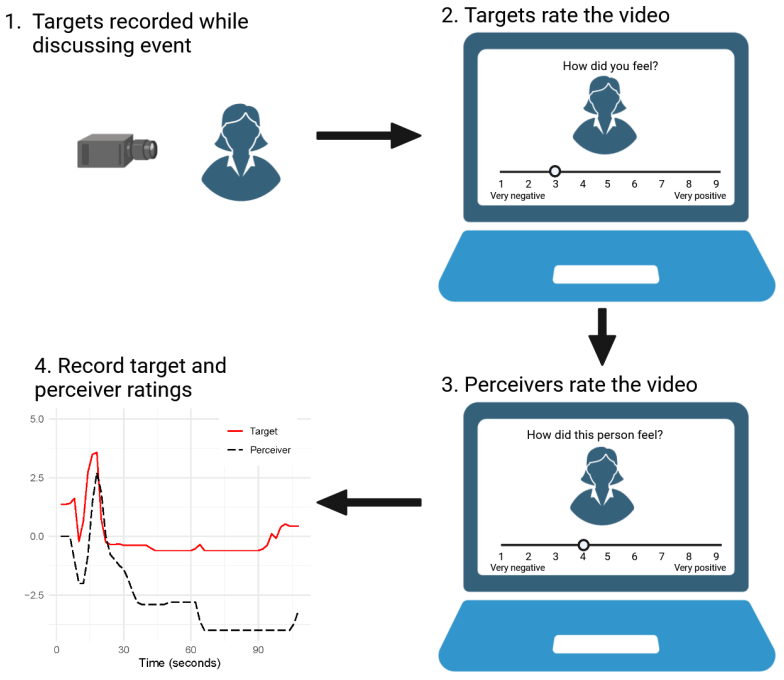


Figure 1. Example of real-time EA data collection procedure.

In clinical research, EA has been used as an index to differentiate individuals with certain psychiatric disorders from healthy controls (Lee et al., 2011). However, the validity of these studies critically depends on the quality and accuracy of EA measurement.

There are two types of studies commonly used to examine EA. One is the non-real-time EA study design, where perceivers provide their response to stimuli after the stimuli have been conducted. The outcome of their overall empathy can be the categories of emotion (e.g., happiness, anger, sadness, etc.) or the extent of emotion on a Likert-type scale (Ekman, 1992; Schweinle et al., 2002). The other EA study design is the real-time assessment of perceivers' empathy on an audio or video stimuli without pausing (i.e., the recorded affective states of targets) (Jospe et al., 2020; Zaki et al., 2008), where perceivers provide continuous feedback on their perceptions of the target's emotional state while the stimuli is unfolding. Illustrated in Figure 1, social targets varying in trait emotional intensity were videotaped while discussing emotional autobiographical events. Perceivers watch these videos and report the perceived emotions every two seconds using, for example, a 9-point Likert scale (e.g., 1 = extremely negative; 9 = extremely positive). Compared with the non-real-time EA studies, the real-time design provides more granular information on the dynamic nature of empathy in everyday interactions and detects subtle changes in emotional responses that might be missed in non-real-time assessments.

In this article, we focus on analyzing data from real-time EA study designs. For such designs, correlational analysis (Mackes et al., 2018; Zaki et al., 2009, among others) is a predominant statistical method for examining EA. This approach computes a monotonic transformation of the Pearson correlation between the observed perceivers' responses with targets' self-reported emotion rating. Linear models have also been introduced to investigate the influence of additional factors or unobserved variables on EA. For example, Tabak et al. (2022) proposed a latent variable model that decomposes EA into three separate dimensions: bias, discrimination, and variability. Bias measures the systematic difference between perceiver's ratings and target's ratings; discrimination measures perceiver' sensitivity in relation to target's ratings; and variability measures the variance of random error in perceiver's perceptions.

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press

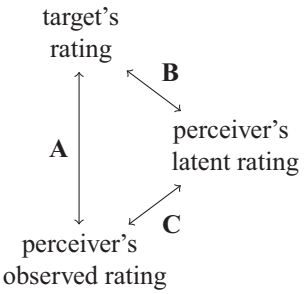


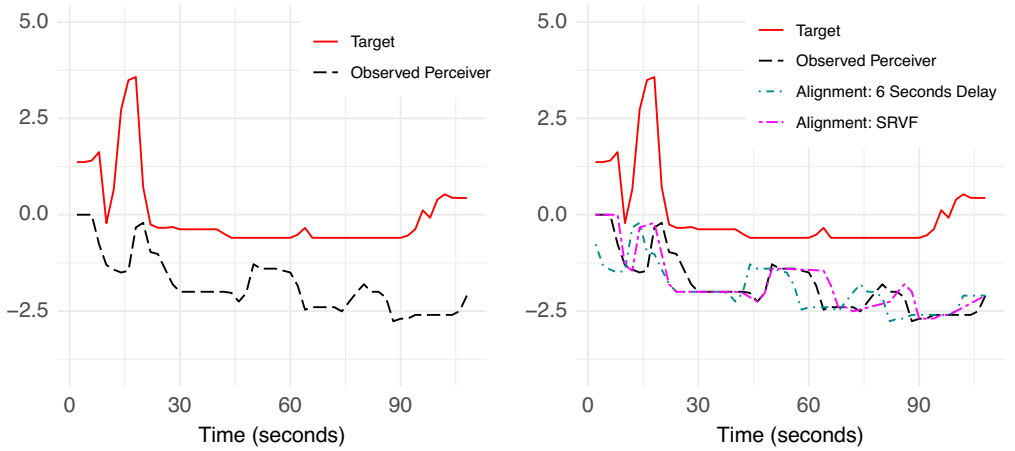
Figure 2. Illustration of different relations between a target's rating, a perceiver's latent rating, and the perceiver's observed rating. Discrepancy A denotes the difference between a target's rating and a perceiver's observed rating. Disagreement B captures the inconsistency between target's rating and perceiver's latent rating, which is the true focus of EA. Misalignment C refers to the divergence between perceiver's observed rating and their latent rating, often due to distortions in expressing their internal judgment. Discrepancy A can arise from both Disagreement B and Misalignment C. Most conventional EA methods mistakenly assess Discrepancy A, thereby conflating measurement error with genuine empathic inaccuracy.

A key assumption in traditional correlational and linear model analyses of EA is that perceivers' and targets' rating sequences are perfectly aligned—that is, a perceiver's rating at a given time point is directly compared to the target's rating at that same moment. However, this assumption often fails in practice due to the complex cognitive processes involved in interpreting another person's emotional state and the time required to produce a behavioral response. Scherer's multi-stage model of emotion decoding Scherer (2003) highlights how perceivers actively interpret dynamic cues such as facial expressions, gestures, and vocalizations to infer emotions, a process that naturally introduces temporal delays. Additionally, the act of recording a response, such as pressing a key or moving a joystick, can vary in duration, further contributing to misalignment. To address these issues, we posit that each perceiver has an underlying latent rating that reflects their true empathic understanding, independent of these timing distortions.

Figure 2 illustrates the relations between a target's rating, a perceiver's latent rating, and the perceiver's observed rating. Discrepancy A, the difference between the target's rating and the perceiver's observed rating, can arise from two sources: Disagreement B, which reflects the true empathic inaccuracy between the target's rating and the perceiver's latent rating, and Misalignment C, which captures the temporal mismatch between the perceiver's latent and observed ratings. Crucially, EA is intended to measure a perceiver's ability to correctly infer another person's emotional state—that is, to quantify Disagreement B—not the speed or timing of their response. A perceiver who accurately identifies a target's emotion, even with a slight misalignment, should not be penalized as less empathically accurate. However, traditional EA methods often ignore Misalignment C, relying solely on comparisons between the target's and perceiver's observed ratings (i.e., measuring Discrepancy A). This could potentially lead to biased estimates of EA and inflated variability due to even minor timing discrepancies. Our work addresses this limitation by explicitly correcting for Misalignment C, thereby yielding more accurate estimates of Disagreement B and preserving the psychological validity of EA assessments.

Note that common approaches to address misalignment in EA studies involve introducing a fixed response delay, assuming consistent emotional expression patterns across individuals. This method shifts perceivers' response time series backward by a predetermined amount (Huang et al., 2015; Khorram et al., 2019; Nicolle et al., 2012). However, Scherer (2003) countered this assumption, arguing that emotional expressions are diverse and context-dependent. In addition, a review of event-related potential (ERP) studies spanning 40 years found that emotional stimuli elicit differences in neural processing speed based on valence and arousal level Olofsson et al. (2008). The findings suggest that stimuli with higher motivational relevance receive priority in neural processing. Consequently, the misalignment between perceivers' and targets' ratings is more complex than a simple, fixed time shift applied to all participants. It would be inappropriate to apply a fixed time delay in EA studies, as it fails

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press



(a) Emotion rating functions (b) Aligned functions

Figure 3. (a) An example of misaligned rating sequences between a perceiver and a target. The solid red line represents the target's ratings and the black dashed perceiver's ratings. (b) Aligned ratings for the perceiver. The green dashed line shows the 6-second delay adjustment, and the purple dashed line shows the aligned ratings using the penalized SRVF representation.

to account for the variability in emotional processing across different moments. Figure 3a illustrates an example of misalignment between perceiver and target ratings in an EA study (Devlin et al., 2014). While a delay in perceiver responses is evident, it is not the sole cause of misalignment. For instance, the perceiver's prolonged sustained response from 10 to 15 seconds, in contrast to the target's brief dip at 10 seconds, highlights the complex nature of these discrepancies.

To accommodate a wider range of misalignment patterns beyond simple time shifts, time series alignment methods aim to preserve key structural features in the data, such as peaks and valleys, ensuring more accurate analysis. One widely used technique is dynamic time warping (DTW), which aligns time series by stretching or compressing the time axis to match similar patterns (Berndt & Clifford, 1994; Sakoe & Chiba, 1978). However, DTW can sometimes introduce distortions by forcing unnatural alignments between sequences (Marron et al., 2015; Srivastava et al., 2011; Zhao et al., 2020). To mitigate this, smoothness penalties have been proposed (Ramsay & Silverman, 2005). However, it may also lead to biased alignments (Guo et al., 2022). Alternatively, landmark-based methods align time series by identifying and matching distinctive features like peaks and valleys (Kneip et al., 2000). While potentially effective, these methods are highly sensitive to noise and may lose important information due to the discretization of continuous functions into a limited set of landmarks (Marron et al., 2015; Wang & Gasser, 1997). Moreover, such approaches are ill-suited for real-time emotion rating data in EA studies, where there is no clear consensus on the number or the location of meaningful landmarks.

Due to the high-frequency nature of the observed EA rating data, we treat each observed curve as a sample path of a continuous function in the time domain, i.e., functional data. Such an approach of representing high-frequency data as functional is common in the literature (Kokoszka & Reimherr, 2017). From this perspective, misalignment between two observed ratings could be explained by a smooth warping function that distorts the time domain of the perceiver relative to that of the target. Hence, the target and the perceiver's rating functions can be aligned by estimating this smooth warping function from the observed data, for example, by minimizing an $\mathbb{L}^2$ distance between the target and the estimated aligned response function (Ramsay & Li, 1998). Recently, the square root velocity function (SRVF) representation has been employed for aligning functions (Srivastava et al., 2011), and has been increasingly applied across various fields, including biology, medicine, geology, and signal processing (Bharath et al., 2018; Laga et al., 2014; Mitchell et al., 2025; Su et al., 2014; Zhao et al., 2020). As we will review in Section 2, this SRVF representation leverages the Fisher–Rao metric's invariance property, and

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press

enables a consistent separation of horizontal component (also known as phase) from vertical component (also known as amplitude) of functions, making visualization and summarizing variability in functional datasets more effective (Xie et al., 2017).

Building upon the SRVF-alignment framework, this article introduces a novel penalized SRVF-based alignment method for unsynchronized rating sequences in EA studies. Our approach introduces both practical and methodological innovations. Practically, it is the first method in EA research to accommodate a wide range of misalignment patterns (e.g., delays, compressions, and stretches), moving beyond the limitations of fixed-delay adjustments. Methodologically, we incorporate a novel penalty term that constrains temporal shifts within bounds consistent with human perceptual capabilities, thereby preventing excessive or unrealistic alignments (Gunes & Pantic, 2010; Levenson, 1988; Marioryad & Busso, 2014; Ringeval et al., 2015). This is important because not all temporal discrepancies should be corrected; some may reflect genuine empathic inaccuracy rather than misalignment. To address this, our penalized alignment method selectively adjusts only short-term misalignments—those occurring within a few seconds—treating them as Misalignment C (as shown in Figure 2). In contrast, larger discrepancies, which may indicate a lack of empathic understanding (Disagreement B), are preserved.

To highlight the contribution of our method, Figure 3b compares the proposed penalized SRVF method with a fixed 6-second delay adjustment. Although the 6-second delay adjustment aligns the peaks between the two sequences, it, unfortunately, eliminates the brief 5-second sustain at the start of the perceiver's sequence, which originally matched up with the target's self-rating sequence. In contrast, the proposed penalized SRVF-based method has aligned the peaks while keeping the initial sustain in the perceiver's sequence in place, demonstrating its flexibility in handling complex misalignment patterns. By enabling a more precise alignment, our method yields a more accurate estimation of EA, avoiding the pitfalls of underestimation when misalignment is ignored and overestimation when no penalty is applied.

The remainder of the article is structured as follows. Section 2 provides background information on EA and existing alignment methods. The proposed methodology is detailed in Section 3. To evaluate the proposed method, Section 4 presents a simulation study and comparisons to alternative approaches. Real-world applications of assessing EA in social and music contexts are explored in Section 5. Finally, Section 6 offers a discussion of the findings and concludes the article.

2. Background

2.1. Elastic functional data analysis

Functional data often exhibit both vertical and horizontal differences, where the latter is known as phase variation and characterized by misaligned geometric features such as peaks and valleys in the time domain (Tucker et al., 2014; Wallace et al., 2014; Wu & Srivastava, 2014). Let $x, a, y : [t_0, t_T] \to \mathbb{R}$ be the function of the target rating, the function of the perceiver's latent rating, and the function of the perceiver's observed rating, respectively. To account for both vertical and horizontal difference between these two functions, we assume a data generation process that $y(t) = f(x(\tilde{t}), \varepsilon(\tilde{t}))$, where $\tilde{t} = \psi(t)$, $f : \mathbb{R}^2 \to \mathbb{R}$ is a link function, $\psi : [t_0, t_T] \to [t_0, t_T]$ is a time warping function, $\circ$ denotes the composition operator, and ε is a random noise function. Essentially, the process of generating the perceiver's observed rating y is decomposed into a transformation step and a warping step, as depicted in the following expression (2.1).

$$x \xrightarrow{\text{Transform}} a = f(x, \epsilon) \xrightarrow{\text{Warp}} y = a \circ \psi. \tag{2.1}$$

Note that in this transformation step, the link f matches the target function x at a time point t to the perceiver's latent function a at the same time t. This correspondence is distorted by a warping function ψ in the second warping step, so the target function x at time t now is matched to the perceiver's

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press

observed function y at $\psi(t)$. The warping function ψ is usually assumed to belong to the set of smoothing functions Γ_I, where

$$\Gamma_I = \{\psi \mid \psi(t_0) = t_0, \psi(t_T) = t_T,\ \psi'(t) \text{ exists }, \psi'(t) \geq 0,\ \gamma = \psi^{-1} \text{exists}\}.$$

Relating this process to Figure 2, the transformation step models the Disagreement B, while the warping step models the Misalignment C, and the Discrepancy A accumulates both steps together. In the context of measuring EA, the perceiver's latent function a represents a rating from the perceiver that is aligned with x, i.e., that *can be compared with the target x point-to-point in time*, and a measure of EA is a similarity measure between x and a.

The data generation process (2.1) motivates the following workflow for quantifying EA. Because $y = a \circ \psi$, we can write $a = y \circ \gamma$, where the inverse warping function $\gamma = \psi^{-1} \in \Gamma_I$ is assumed to exist since $\psi \in \Gamma_I$. Hence, we first conduct an alignment step to obtain an estimated inverse warping function $\hat{\gamma}$ and an estimated latent function $\hat{y} = y \circ \hat{\gamma}$ from the observed target and perceiver functions x and y. Then, we could estimate EA by a similarity measure between x and $\hat{y}$.

Since misalignment between two functions is inherently related to the difference in how fast they move, a common way to conduct the alignment step is to compare how these functions change over time, which is mathematically described by their corresponding first derivative. Therefore, the general idea of the SRVF-based alignment methods is to minimize the distance between the first derivative of the target x and the estimated function $\hat{y}$. We briefly review the formulation of the SRVF representation here, where more details can be found in Srivastava & Klassen (2016).

For any absolute continuous function $f : [t_0, t_T] \to \mathbb{R}$, the SRVF of f is the function $q_f : [t_0, t_T] \to \mathbb{R}$, $q_f(t) = \text{sign}\{f'(t)\}\sqrt{|f'(t)|}$, where $f'(t) = df/dt$. As described in the previous paragraph q_f, this SRVF is defined based on the first derivative f'; the specific form of $q_f(t)$ is motivated to keep its norm unaffected by the warping, which is useful to separate a function into its amplitude and phase component (Srivastava & Klassen, 2016). Specifically, if f is warped by γ, the corresponding SRVF of $f \circ \gamma$ becomes $q_{f\circ\gamma} = (q_f \circ \gamma)\sqrt{\gamma'}$, but the squared $\mathbb{L}^2$ norm is preserved $\|(q_{f\circ\gamma})\|_2^2 = \|q_f\|_2^2$. Let q_x and q_y be the SRVFs of the target and perceiver functions, respectively. Then, the SRVF-based alignment method aims to find an optimal inverse warping function that minimizes this discrepancy, i.e.,

$$\hat{\gamma}_u = \underset{\gamma\in\Gamma_I}{\arg\inf} \|q_x - (q_y, \gamma)\|_2^2, \tag{2.2}$$

where we write $q_{f\circ\gamma} = (q_f, \gamma)$ to ease the notation. The optimal $\hat{\gamma}_u$ is expected to align two functions so that the transformed function $\hat{y} = y \circ \hat{\gamma}_u$ is aligned with x. The subscript u stands for "unpenalized," meaning the optimal $\hat{\gamma}_u$ is not subject to any other constraint than being in the space Γ_I. This unpenalized alignment has been implemented in the `fdasrvf` packages (Tucker, 2025) in both R and Python. After conducting the alignment step, in addition to the EA measure obtained by computing a similarity metric between $\hat{y}$ and x, we can also quantify the amount of warping made by each perceiver in relative to the target by a Fisher–Rao phase distance between the estimated warping $\hat{\gamma}_u$ and the identity warping $\gamma_{id}(t) = t$ as

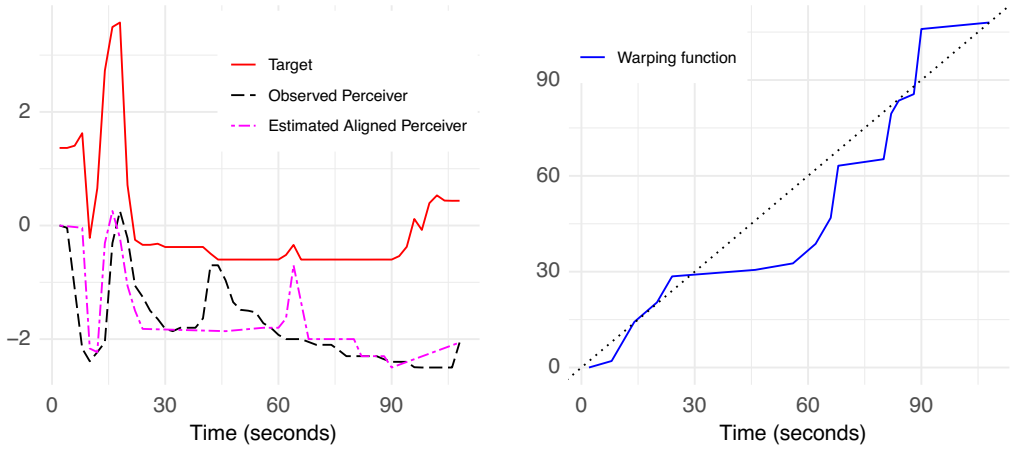
$$d_p(x, y) \approx \cos^{-1}\left(\int_0^1 \sqrt{\hat{\gamma}_u'(t)}\, dt\right), \tag{2.3}$$

which is a proper metric distance on the set Γ_I (Srivastava & Klassen, 2016).

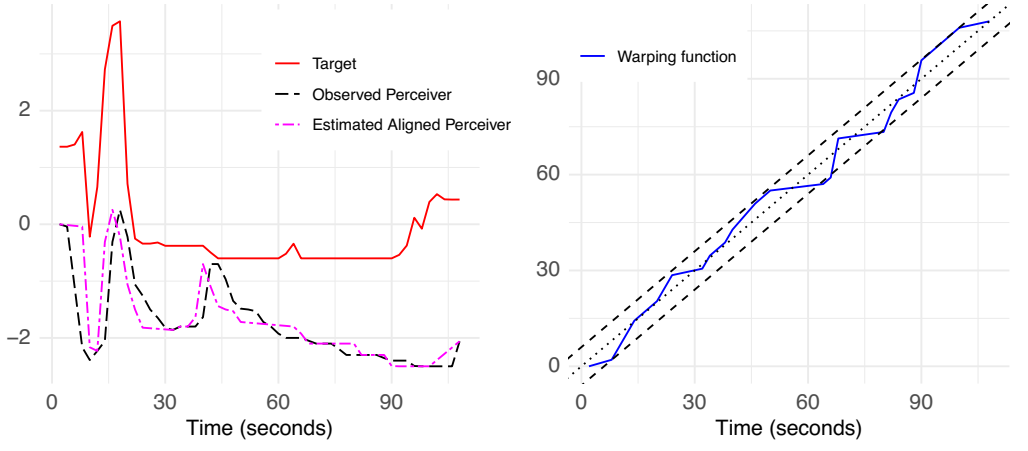
2.2. Unpenalized SRVF leads to over-alignment

While the SRVF representation leads to several theoretical benefits, one main disadvantage of the unpenalized SRVF for studying EA is that the estimated perceiver function $y \circ \hat{\gamma}_u(t)$ may be overaligned

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press



(a) Unpenalized alignment



(b) Penalized alignment

Figure 4. Example target and perceiver's emotion ratings of Devlin et al. (2014). (Left): target x (solid), perceiver's observed response y (dash), and estimated perceiver's response $\hat{y} = y \circ \hat{\gamma}$ (dot dash). (Right): estimated warping function $\hat{\gamma}$.

with the target x and thus could differ from the perceiver's latent function a. In other words, the unpenalized SRVF not only corrects for the Misalignment C in Figure 2, but also potentially removes inherent temporal Disagreement B.

Figure 4a shows one example from the study in Devlin et al. (2014) demonstrating the result of the previous SRVF alignment obtained from (2.2). In this study, the continuous ratings were recorded for 108 seconds and averaged over 2-second epochs. The alignment is obtained by using their SRVF representations q_x, q_y, and $(q_y, \hat{\gamma}_u)$. The estimated inverse warping function $\hat{\gamma}_u$ is plotted in the right panel of Figure 4a. When $\hat{\gamma}_u$ appears above the 45-degree line, it implies that the perceiver's response is delayed compared to the target, whereas $\hat{\gamma}_u$ below the 45 degree line indicates that the perceiver's response precedes the target. While it seems reasonable to expect that the perceiver's perception of a particular emotion would lag behind the target's actual expression of that emotion, there is evidence to suggest that people can make anticipatory perceptual judgements, especially when the stimuli are continuous and dynamic. For example, Thornton and Tamir Thornton & Tamir (2017) found that perceivers attend to emotion regularities and can predict up to two emotional transitions into the future.

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press

Koster–Hale and Saxe Koster-Hale & Saxe (2013) argued that the brain actively generates expectations about others' emotions, thoughts, and behaviors (so not just passively reacting to them). They refer to this as "predictive coding." In Figure 4a, the peak of the perceiver's response around $t = 45$ seconds is considered as a response preceding the target's self-rating around $t = 65$ seconds, and it is aligned accordingly by the unpenalized SVRF method.

Therefore, from the left plot of Figure 4a, the unpenalized SRVF method misaligns the peak of the perceiver's response, occurring at approximately $t = 40$ seconds, with the target's small peak at around $t = 65$ seconds, which is likely just noise. This alignment suggests an improbable scenario, where the perceiver predicts the target's emotional change 25 seconds in advance. Psychological research has consistently shown that reaction time-delay is limited to a few seconds: 0.5 to 4 seconds (Levenson, 1988), 3 to 6 seconds (Nicolle et al., 2012), 2 to 11 seconds (Marioryad & Busso, 2014), and 0.48 to 6.24 seconds (Ringeval et al., 2015). By disregarding this inherent limitation, unpenalized SRVF alignment overestimates synchronization between rating sequences, potentially leading to an unrealistic shape of the estimated warping function that exceeds the human exception bounds, and hence biased estimations of perceiver EA levels.

3. Method

3.1. Penalized elastic functional alignment

Penalized alignment has been proposed to control the amount of alignment (Guo et al., 2022; Mitchell et al., 2025; Wu & Srivastava, 2011) or to achieve smooth alignment (Srivastava & Klassen, 2016). To address the over-alignment issue inherent in the unpenalized SRVF method, an existing solution is to employ a penalized alignment approach by incorporating a penalty term into the unpenalized alignment optimization function (2.2). This results in the following objective function:

$$\|q_x - (q_y, \gamma)\|_2^2 + \lambda \mathcal{R}(\gamma), \tag{3.1}$$

where γ is the inverse warping function, $\lambda > 0$ is a penalty parameter, and $\mathcal{R}(\gamma)$ is a penalty function. Several penalty functions have been suggested in the literature, such as $\mathcal{R}(\gamma) = \|\sqrt{\gamma'} - \mathbf{1}\|_2^2$ and $\mathcal{R}(\gamma) = \cos^{-1}(\langle\sqrt{\gamma'}, \mathbf{1}\rangle)$, which are used to measure the differences between the SRVFs of γ and the identity warping $\psi_{id}(t) = \gamma_{id}(t) = t$ by the squared $\mathbb{L}^2$ norm and the arc length, respectively, where $\mathbf{1}$ is the constant function with value 1 and $\langle\cdot,\cdot\rangle$ denotes an inner product operator (Srivastava & Klassen, 2016).

The aforementioned penalty functions are inappropriate for current EA research. Primarily, it is challenging to select an optimal data-driven tuning parameter λ. Common cross-validation procedures that split the data into independent training and test sets do not preserve the geometric features of the data. Second, as reviewed in Section 2.2, psychological research indicates that misalignment in perceiver ratings occurs within a specific temporal window of a few seconds. Existing penalty functions, however, focus on controlling the overall amount of warping, which does not directly translate to constraining alignment at each individual time point as required for EA studies.

To address these limitations, we introduce a novel penalized alignment method that directly incorporates the established temporal boundary for maximum perceiver misalignment as a penalty term. Specifically, we construct the optimal inverse penalized warping function

$$\begin{aligned} \hat{\gamma}_p &= \arg\inf_{\gamma \in \Gamma_I} \|q_x - (q_y, \gamma)\|_2^2, \\ &\text{s.t. } \sup_t |\gamma(t) - \gamma_{id}(t)| \le \nu, \end{aligned} \tag{3.2}$$

where ν is the predefined upper limit of warping functions, corresponding to the maximum delay or advance observed in the perceivers' responses. Although the supremum norm limit ν plays the role

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press

of a tuning parameter, in practice, we often have prior knowledge about its value based on the research context, unlike the tuning parameter λ in the existing approach (3.1). Nevertheless, it is useful to perform a sensitivity analysis of the proposed method over a reasonable range of v. We denote $\hat{\gamma}_p$ as the estimated inverse warping function of penalized alignment, where the subscript p stands for "penalized." As $v \to 0$, $\hat{\gamma}_p \to \gamma_{id}$, so that no warping is allowed. On the other hand, if $v \geq \sup|\hat{\gamma}_u - \gamma_{id}|$, the constraint in (3.2) is inactive, then $\hat{\gamma}_p = \hat{\gamma}_u$. Consequently, any v smaller than $\sup|\hat{\gamma}_u - \gamma_{id}|$ induces a shrinkage effect, pulling the unpenalized warping towards the identity warping function, akin to penalized regression. This interpretable penalty mechanism enables our proposed penalized alignment to mitigate the risk of over-alignment, resulting in more plausible warping estimates and aligned responses.

We note that under the constraint (3.2), the Fisher–Rao phase distance d_p defined by (2.3) is still valid to measure the difference between the phase of two functions, with the exception that $\hat{\gamma}_u$ is replaced by $\hat{\gamma}_p$. The proof of Lemma 3.1 is given in Section S1 of the Supplementary Material.

Lemma 3.1. *The Fisher–Rao phase distance between x and y is estimated by* $d_p(x,y) = \cos^{-1}\left(\int_0^1 \left(\sqrt{\hat{\gamma}_p'(t)}dt\right)\right)$.

3.2. Computing the penalized SRVF alignment

A discrete approximation for the solution of the optimization problem specified in (3.2) can be found by using the following dynamic programming algorithm (DPA) (Srivastava & Klassen, 2016). Assume both the SRVF functions of the target and the perceiver q_x and q_y are observed at $T+1$ time points, $t_0 < t_1 < t_2 < \cdots \leq t_T$. Without loss of generality, we assume that $t_0 = 0$ and $t_T = 1$, and that these time points are equally spaced, i.e., $t_m = m/T$ for $m = 0, \ldots, T$. The inverse warping function γ matches the point (q_y, γ) with the point q_x, so γ can be viewed as a graph with a collection of points $(t_m, \gamma(t_m))$, from $(0,0)$ to $(1,1)$ in $\mathbb{R}^2$. We assume that within each interval $[t_m, t_{m+1}]$, the function $\gamma(t)$ can be approximated by a straight line, so the final estimate for $\hat{\gamma}$ is a piecewise linear path. Since γ is non-decreasing, the slope of this graph is always strictly between 0 and 90 degrees. Furthermore, the cost function in (3.2) can be approximated by

$$\int_0^1 \left\{q_x(t) - q_y(\gamma(t))\sqrt{\gamma'(t)}\right\}^2 dt \approx \sum_{m=0}^{T} \int_{t_m}^{t_{m+1}} \left\{q_x(t) - q_y(\gamma_m(t))\sqrt{\gamma_m'(t)}\right\}^2 dt, \tag{3.3}$$

where $\gamma_m(t)$ is a straight line connecting $(t_m, \gamma(t_m))$ and $(t_{m+1}, \gamma(t_{m+1}))$. The function on the right-hand side of (3.3) is additive over the graph, and hence enables the use of DPA. Our goal then is to find an optimal linear piecewise path from $(0,0)$ to $(1,1)$ in $\mathbb{R}^2$ that minimizes (3.3), subject to the constraint that $\sup_{t \in \{t_1, \ldots, t_T\}} |\gamma(t) - t| \leq v$. Using DPA, we can construct this path recursively as follows.

Given a feasible point (t_k, t_l), i.e., $|t_l - t_k| \leq v$ in the graph, let $\mathcal{N}_{k,l} = \{(k', l') \mid 0 \leq k' < k,\ 0 \leq l' < l,\ |k' - l'| \leq v\}$ denote the set of all nodes in the graph that are allowed to go to (t_k, t_l) by a straight line. Starting from $(0,0)$, if we have already determined and stored the cost of reaching nodes in $\mathcal{N}_{k,l}$, then the cost of reaching (t_k, t_l) is given by

$$H_{k,l} = \min_{(k',l') \in \mathcal{N}_{k,l}} \left(H_{k',l'} + \int_{t_k}^{t_{k'}} \left\{q_x(t) - q_y(\gamma(t))\sqrt{\gamma'(t)}\right\}^2 dt \right), \tag{3.4}$$

where we initialize $H_{0,0} = 0$ and $H_{0,l} = H_{k,0} = \infty$ for any $l \neq 0$ and $k \neq 0$. Let $(\hat{k}, \hat{l})$ be the nodes that minimize the right-hand side of (3.4) and repeat the process for every possible point (t_k, t_l) in the graph. Then, the optimal curve $\hat{\gamma}_p$ is obtained by connecting all such points using piecewise linear curves. Note that compared to the standard DPA algorithm to align the two functions (Srivastava & Klassen, 2016), we have modified the set of permitted nodes to account for the constraint imposed on the warping function.

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press

The algorithm is summarized in Algorithm 1.

Algorithm 1 SRVF alignments with sup constraints.

Set $\mathcal{N}_{k,\ell} = \{(k',l') \mid 0 \le k' < k,\ 0 \le l' < l, |k'-l'| \le v\}$
Set $H_{0,0} = 0$, $H_{0,l} = H_{k,0} = \infty$ for any $l \neq 0$ and $k \neq 0$.
for $(0,0 < (t_k,t_l) < (1,1)$ **do**
  Find (k',l') that minimizes the right-hand side of (3.4)
  Compute $H_{k,l}$
end for
Initialize $(t_k,t_l) = (1,1)$ ▷ Construction of $\hat{\gamma}_p$
while $(t_k,t_l) \neq (0,0)$ **do**
  Draw a straight line between (t_k,t_l) and $(t_{k'},t_{l'})$
  Set $(t_k,t_l) = (t_{k'},t_{l'})$
end while

Because the temporal window can vary according to the emotions and the modality (Gunes & Pantic, 2010; Gunes & Schuller, 2013), we used three thresholds of $v = 6, 8$, and 10 seconds for EA applications in Section 5. Figure 4b shows penalized alignment of the same example presented in Figure 4a, using an upper limit of warping function differences of six seconds ($v = 6$). Since $\sup|\hat{\gamma}_u - \gamma_{id}| = 23.39 > v = 6$ for unpenalized alignment, penalized alignment shrinks the estimated inverse warping function $\hat{\gamma}_p$ toward the identity warping function. Consequently, the resulting estimated perceiver latent function $\hat{y} = y \circ \hat{\gamma}_p$ does not exhibit peaks or valleys that deviate from the observed perceiver sequence by more than six seconds.

4. Simulation study

To demonstrate the performance of our functional alignment approach, we conducted a number of simulation studies. It is challenging to use real EA data to evaluate functional alignment methods because perceivers' latent ratings are unknown. However, we generated perceivers' latent responses from the real target ratings and used them to compare different alignment methods.

4.1. Simulation 1

In this simulation, we evaluated the effectiveness of various alignment methods. To approximate the settings in real–data applications, we used the four targets $x_j(t)$ directly derived from the real target data of Devlin et al. (2014) corresponding to four videos: high intensity positive, low intensity positive, high intensity negative, and low intensity negative, for $j = 1,\ldots,4$. We smoothed these raw data using the cubic smoothing spline and recorded the functional values for 300 evenly-spaced time points ($t = 0,1,\ldots,299$).

Next, we generated $n = 500$ perceivers' latent responses for each target function $x_j(t)$ using the following model

$$a_{ij}(t) = \epsilon_{ij}(t)\, x_j(t) + u_{ij}(t),$$

for $i = 1,\ldots,n$. Here, we set $\epsilon_{ij}(t) = K_h(W_{ij}(t) + 1)$, where $W_{ij}(t)$ is a one-dimensional Wiener process (i.e., Brownian motion) at time t (Mörters & Peres, 2010), $u_{ij}(t) = K_h(S_{ij}(t))$ with $S_{ij}(t)$ being a standardized random walk at time t that is obtained by cumulatively summing the standard normal $N(0,1)$ noise and applying a standardization transformation. We denote K_h as the Gaussian kernel smoothing with bandwidth h, and in this simulation, we used $h = 20$ to ensure both ϵ_{ij} and u_{ij} are smooth. Then, we generated the perceiver's observed response y_{ij} using the perceiver's latent rating a_{ij} by

$$y_{ij}(t) = (a_{ij} \circ \psi_{ij})(t), \tag{4.1}$$

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press

where $\psi_{ij} \in \{\psi \mid \psi \in \Gamma_I, \sup|\psi(t) - t| = \eta_{ij} \text{ for } t \in [0,299]\}$ is the warping function and η_{ij} is the true individual upper limit of the warping amount. We first generated the warping functions randomly using the *rgam* function in the R package `fdasrvf` Tucker (2025), and then rescaled them such that $\sup|\psi_{ij}(t) - t| = \eta_{ij}$. With that simulation configuration, the true correlation between a_{ij} and the target $x_j(t)$ has the mean around $\{0.65, 0.66, 0.67, 0.66\}$ with standard deviation $\{0.24, 0.27, 0.23, 0.25\}$ for all $j = 1, \ldots, 4$, respectively.

We considered five different methods to align the observed perceiver response to the target, including (1) no alignment, (2) optimal fixed delay, (3) unpenalized SRVF alignment, (4) the squared $\mathbb{L}^2$ norm penalty SRVF alignment, and (5) our proposed penalized SRVF alignment. Let $\hat{y}_{ij}(t) = y_{ij} \circ \hat{\gamma}_{ij}(t)$ denote the estimated perceiver's response, where $\hat{\gamma}_{ij}(t)$ denotes an estimated inverse warping function from one of the above methods. The no alignment option assumes the identity inverse warping $\hat{\gamma}_{ij}(t) = t$. For the optimal fixed delay method (2), we found the optimum amount of delay $0 \leq \delta \leq v$ that achieves the smallest $\mathbb{L}^2$ distance between q_{x_j} and $(q_{y_{ij}}, \hat{\gamma}_{ij})$, where $\hat{\gamma}_{ij}(t) = 0$ if $t = 0$, $\hat{\gamma}_{ij}(t) = t + \delta$ if $0 < t < 1 - \delta$, and $\hat{\gamma}_{ij}(t) = 1$ otherwise. The inverse warping functions of the unpenalized and penalized SRVF alignments were obtained by solving (2.2) and (3.2), respectively. The squared $\mathbb{L}^2$ norm penalty SRVF alignment implements the penalty $\hat{\gamma}_{ij} = \arg\inf_{\gamma \in \Gamma_I} \|q_{x_j} - (q_{y_{ij}}, \gamma)\|_2^2 + \lambda \|\sqrt{\gamma'(t)} - \mathbf{1}\|_2^2$, where $\mathbf{1}$ is the constant function with value one (Srivastava & Klassen, 2016). To the best of our knowledge, an optimal method for selecting λ has not been established in the literature, so we implemented the method with $\lambda = 0.01$. We leave the investigation of optimal selection strategies for λ to future research.

In the simulation, we set the alignment warping limit for the penalized SRVF to $v \in \{6, 8, 10\}$ seconds regardless of the true warping limit η_i to reflect the real-world cases, where the true warping limit is unknown. Also, to account for individual warping variations, we examined two different settings of η_{ij}, including a constant $\eta_{ij} = v$ seconds for all $i = 1, \ldots, n$ and $j = 1, \ldots, 4$, and a varying η_{ij} randomly generated from a Gamma distribution $\Gamma(k, \theta)$ with $k = v$ being the shape and $\theta = 1$ being the scale parameter of the Gamma distribution.

We evaluated the performance of the alignment methods with two metrics. First, we computed the average $\mathbb{L}^2$ distance between the perceiver's latent function a_{ij} and the estimated functions $\hat{y}_{ij}$ by $d_a = \|\hat{y}_{ij} - a_{ij}\|_2^2$. The closer d_a gets to zero, the more accurate estimation of the perceiver's latent response. Second, we computed the average bias between the true and the estimated correlations to the target, $n^{-1} \sum_{i=1}^{n} \{\rho(x_j, \hat{y}_{ij}) - \rho(x_j, a_{ij})\}$. Here, $\rho(x_j, \hat{y}_{ij})$ is a commonly used metric for measuring EA, and $\rho(x_j, a_{ij})$ can be considered as the true correlation that the alignment methods aim to achieve. We reported the results for each target separately.

Table 1 shows the performance metrics for the case $\eta_{ij} = 8$ and $\eta_{ij} \sim \Gamma(8, 1)$. Results from additional settings, which lead to similar conclusions, are provided in Section S2 of the Supplementary Material. Among the five alignment methods evaluated, the proposed penalized SRVF approach consistently outperforms the others in producing the least biased estimation of EA in all the considered simulation designs. In addition, the average amplitude distance d_a of the proposed penalized SRVF is the smallest for the high negative and high positive targets. For the low negative and low positive targets, the $\mathbb{L}^2$ penalized SRVF shows the lowest average d_a but yields much larger bias. Considering both metrics, the results imply that the proposed method makes the most accurate estimation of the phase shift.

The unpenalized SRVF, optimal fixed, and no alignment methods all yield less accurate estimates of a_{ij} compared to the proposed penalized SRVF. Among them, the unpenalized SRVF produces the largest value of d_a, primarily because it tends to over-align the perceiver's observed response y_{ij} to the target's rating x_j, leading to distorted estimates $\hat{y}_{ij}$. Additionally, the optimal fixed alignment method often results in the highest standard errors and bias in d_a, indicating that it provides inconsistent estimates of the perceiver's latent ratings.

4.2. Simulation 2

In practical applications, the true warping limit η is typically unknown. As a result, the alignment warping limit v must be chosen based on approximate prior knowledge, which may not perfectly match

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press

Table 1. Performance of different alignment methods in the simulation studies under different warping limits η, d_a between the estimated perceiver $\hat{y}$ and the true latent perceiver a, and the $(10\times)$ bias of the estimated correlation between the true latent perceiver and the target.

η	Target	Metric	Pen. SRVF	$\mathbb{L}^2$ SRVF	Unpen. SRVF	Opt. fixed	No alignment
8	High Neg	d_a	**4.81 (1.36)**	6.74 (3.69)	11.76 (5.1)	6.71 (3.54)	5.62 (1.46)
		Bias	**0.02 (0.39)**	0.67 (0.91)	1.37 (1.53)	−0.36 (1.12)	−0.17 (0.41)
	Low Neg	d_a	8.33 (5.12)	**4.23 (2.44)**	9.22 (4.04)	9.03 (5.23)	10.56 (6.1)
		Bias	**−0.68 (1.79)**	0.69 (1.17)	1.47 (1.75)	−0.82 (2.14)	−1.55 (2.05)
	High Pos	d_a	**5.78 (1.62)**	7.17 (3.88)	11.59 (5.24)	7.88 (3.35)	6.04 (1.72)
		Bias	**−0.01 (0.87)**	0.63 (1.44)	0.71 (2.25)	−0.21 (1.74)	−0.13 (0.88)
	Low Pos	d_a	5.92 (1.86)	**4.84 (3.23)**	8.3 (3.52)	7.82 (3.83)	6.59 (1.86)
		Bias	**−0.13 (0.61)**	0.55 (0.84)	0.67 (1.1)	−0.76 (1.59)	−0.31 (0.65)
$\Gamma(8,1)$	High Neg	d_a	**4.75 (1.82)**	6.64 (3.56)	11.30 (5.03)	6.57 (3.72)	5.52 (1.96)
		Bias	**0.05 (0.43)**	0.72 (0.92)	1.27 (1.44)	−0.28 (1.03)	−0.14 (0.45)
	Low Pos	d_a	8.02 (4.98)	**4.24 (2.60)**	9.55 (4.24)	9.26 (5.51)	10.49 (5.95)
		Bias	**−0.6 (1.73)**	0.68 (1.27)	1.42 (1.78)	−0.86 (2.27)	−1.46 (1.98)
	High Pos	d_a	**5.69 (1.93)**	7.16 (3.91)	12.18 (5.81)	7.7 (3.41)	5.94 (2.00)
		Bias	**−0.09 (0.83)**	0.59 (1.26)	0.55 (2.52)	−0.32 (1.62)	−0.20 (0.84)
	Low Pos	d_a	5.98 (2.20)	**4.98 (3.17)**	8.40 (3.53)	7.96 (4.17)	6.57 (2.08)
		Bias	**−0.11 (0.73)**	0.68 (0.98)	0.77 (1.31)	−0.66 (1.98)	−0.29 (0.72)

Note: The lowest absolute bias and the lowest d_a are highlighted for each row. Standard errors are included in the parentheses.

the true value. To examine the impact of this mismatch, we conducted a simulation study using the same data generation process described in Section 4.1. Specifically, we evaluated 21 different true warping limits $\eta \in \{0, 0.8, \ldots, 8, \ldots, 15.2, 16\}$ seconds, while fixing the alignment warping limit at $\nu = 8$ seconds.

Figure 5 represents two performance metrics (d_a and bias) of the penalized SRVF method across 21 different true warping limits η. The results illustrate the method's robustness to variation in the true warping limit. Notably, when η falls within approximately two seconds of the alignment limit $\nu = 8$ seconds, the penalized SRVF method achieves low d_a and minimal bias, indicating accurate alignment and estimation. These findings suggest that even without precise knowledge of the true warping limit, selecting ν within a reasonable range yields reliable performance, underscoring the method's practical utility in real-world applications.

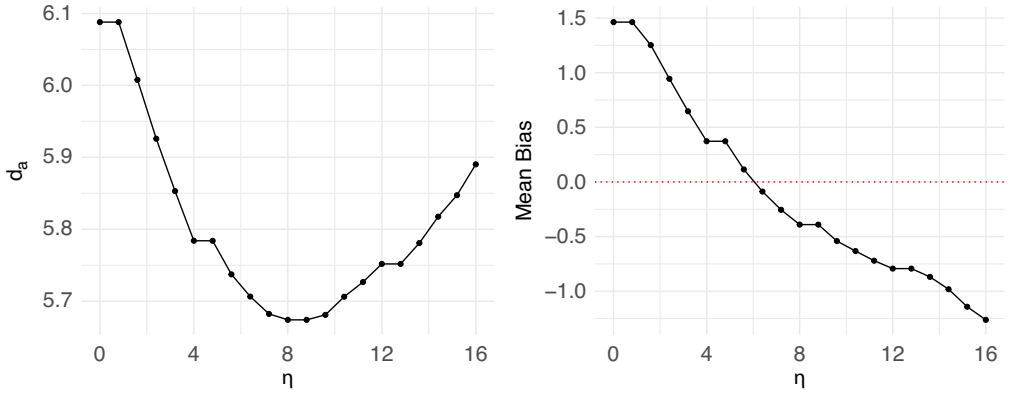
4.3. Simulation 3

In this simulation, we evaluated the performance of the alignment methods under different levels (i.e., high/medium/low) of EA. Perceivers' latent ratings $(i = 1, \ldots, 500)$ were generated following the idea proposed by Matuk et al. (2022),

$$a_i(t) = Q^{-1}\left(\beta_{0i}(t) + \beta_{1i}(t) q_x(t)\right) + \varepsilon_i^a(t), \tag{4.2}$$

where $Q^{-1}(q_f)$ denotes the inverse transformation from an SRVF q_f to the original function f and q_x is the SRVF of the target rating for the low intensity positive video x. The parameters, $\beta_{0i}(t) = \alpha_{0i}\sin(\pi t/50)$ and $\beta_1(t) = \alpha_{1i}\sin(\pi t/100)$, were chosen based on the setting in Ghosal et al. (2020), with $\varepsilon_i^a \sim N(0, 0.1^2)$.

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press



(a) d_a (b) Mean Bias

Figure 5. The penalized SRVF alignment results under 21 different true warping limits $\eta \in \{0, 0.8, \ldots, 8, \ldots, 15.2, 16\}$ seconds when the upper limit of alignment is $\nu = 8$ seconds. The red dotted line in the mean bias plot marks the unbiased level.

Table 2. Comparison results based on d_a between the estimated perceiver $\hat{y}$ and the true latent perceiver a and the ($10\times$) bias of the estimated correlation between the true latent perceiver and the target.

EA	Metric	Pen. SRVF	$\mathbb{L}^2$ SRVF	Unpen. SRVF	Opt. fixed	No alignment
	d_a	3.33 (0.72)	**3.30 (0.69)**	3.75 (0.58)	4.39 (1.47)	3.57 (0.63)
High	Bias	**0.04 (0.34)**	0.50 (0.29)	0.83 (0.24)	−0.50 (0.70)	−0.08 (0.33)
	d_a	**5.10 (1.75)**	5.77 (3.04)	8.13 (3.11)	6.72 (3.06)	5.33 (1.61)
Med	Bias	**0.01 (0.56)**	0.58 (0.84)	1.43 (0.82)	−2.17 (2.17)	−0.22 (0.66)
	d_a	**7.80 (2.62)**	12.08 (5.55)	14.25 (5.07)	11.35 (5.11)	7.82 (2.54)
Low	Bias	0.18 (0.62)	2.06 (1.39)	2.01 (0.87)	−0.11 (2.98)	**−0.03 (0.65)**

Note: The best metrics are highlighted for each row.

We generated perceivers' observed responses following the same data generation process as in (4.1), setting the alignment warping limit to $\nu = 8$ and drawing the true sup-norm limit from a Gamma distribution, $\eta_i \sim \text{Gamma}(8,1)$. To simulate varying levels of EA, we defined three conditions: for high EA, $\alpha_{0i}, \alpha_{1i} \sim N(0, 0.05^2)$; for medium EA, $\alpha_{0i}, \alpha_{1i} \sim N(5, 2^2)$; and for low EA, $\alpha_{0i} \sim N(0, 0.5^2)$ and $\alpha_{1i} \sim N(0.1, 0.1^2)$. The resulting average correlations between the target ratings and the perceivers' latent ratings, $\rho(a_i, x)$, are approximately 0.81, 0.62, and 0.25 for the high, medium, and low EA conditions, respectively.

Table 2 summarizes two evaluation metrics for the five alignment methods across three levels of EA. The proposed penalized SRVF method demonstrates strong performance regardless of the EA levels. When EA is high, the $\mathbb{L}^2$ penalized SRVF achieves a slightly lower average d_a than the penalized SRVF, but the latter yields the lowest average bias. The no alignment method also archives comparable low average bias to the penalized SRVF method. At the medium EA level, the penalized SRVF outperforms all other methods. In the low EA condition, the penalized SRVF's d_a is comparable to the no alignment approach but the no alignment method shows the lowest bias. The benefit of adjusting the observed rating is expected to be limited given the weak association between the target rating and the perceiver's latent rating. In contrast, the other alignment methods perform significantly worse than the penalized SRVF, particularly under low and medium EA conditions.

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press

5. Data application

5.1. Study on social empathy

In the first data application, we analyzed a dataset from Devlin et al. (2014), which consists of 121 perceivers' empathy responses of four distinct videos in which the targets discuss emotional events in their lives. The four videos vary in valence (positive or negative) and intensity (high or low), resulting in four heterogeneous videos, including high-positive, low-positive, high-negative, and low-negative. Participants provided continuous 9-point scale ratings of target emotions while watching each video. These perceiver ratings were compared to the target's self-ratings. Following standard functional data analysis practices (Srivastava & Klassen, 2016), we preprocessed the data by smoothing target and perceiver ratings using cubic smoothing splines with the default setting of the *smooth.spline* function in R and interpolating the estimated functions on a 300-point equidistant grid within the observed time interval. The goal of the subsequent analysis is to measure the level of EA for each perceiver, quantified by the correlation between the perceiver's latent ratings and the target's ratings.

Figure 3 clearly illustrates the misalignment between perceiver and target ratings. Devlin et al. (2014) did not account for this misalignment, measuring EA as a monotonic transformation of the Pearson correlation between the two rating sequences. We applied both unpenalized and penalized SRVF alignments, as these methods offer more flexible time warping than fixed delay alignment. Here, we present results for the penalized alignment with a threshold of $v = 8$ seconds. Results for thresholds of 6 and 10 seconds are included in Section S3.1 of the Supplementary Material.

To quantify the degree of warping, we computed the phase distance (d_p) between the estimated inverse warping function under each alignment method and the identity warping function for each video. The summary statistics for this measure can be found in Table S2 of the Supplementary Material. Figure 6 reveals that the unpenalized SRVF alignment consistently produces warping functions farther from the identity function than the penalized SRVF alignment, indicating the latter's effectiveness in reducing excessive warping, where the p-values (Table S3 in the Supplementary Material) corresponding to the t-tests for the mean differences between penalized SRVF and other method are very close to zero.

We subsequently calculated the Pearson correlation between each perceiver's aligned ratings and the target's ratings, which were used as the EA measure. Unlike the simulation studies, the correlation coefficient $\rho(a,x)$ between the target (x) and the perceiver's latent response (a) is not observable because the perceiver's latent response is not known. Instead, we compared these EA measures to those obtained without alignment (identity warping), referred to as identity correlations. Notably, about 2% of the cases

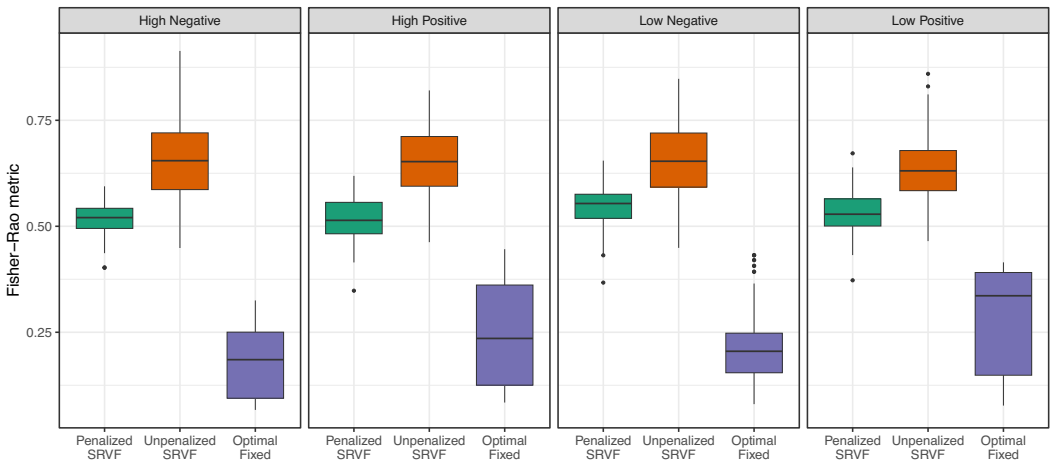
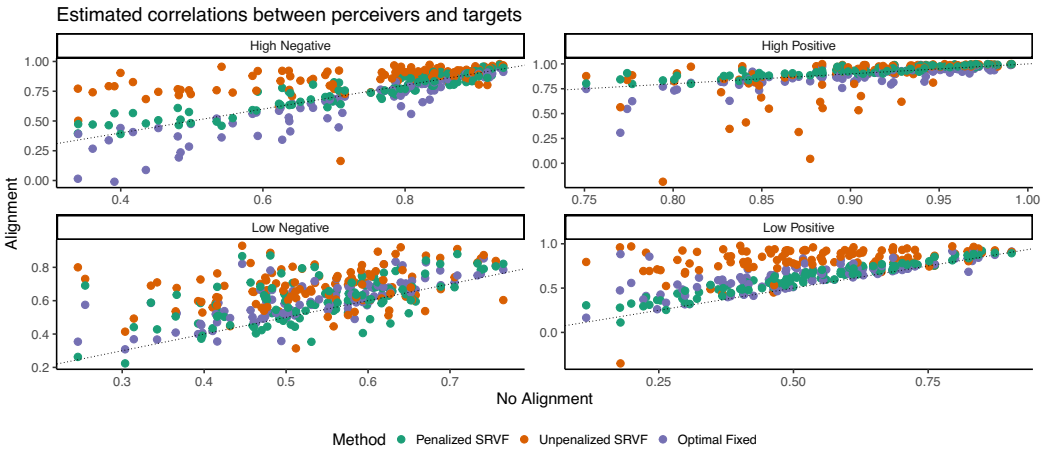
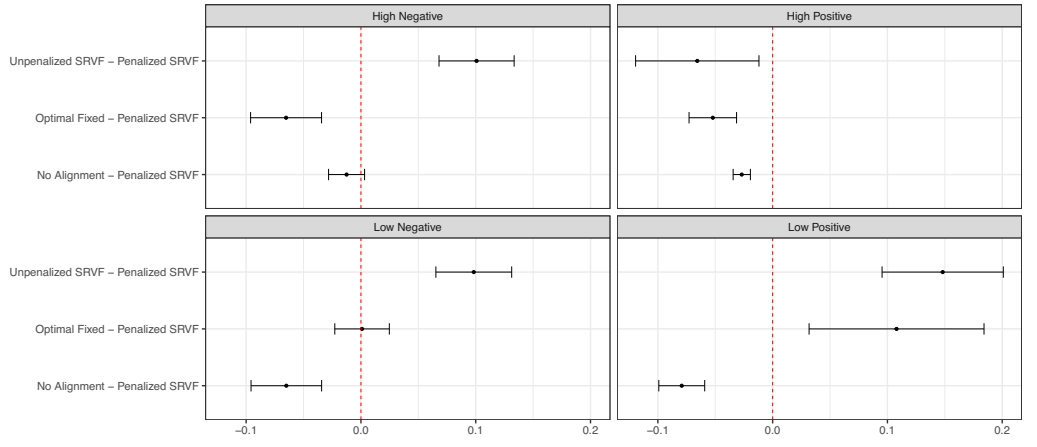


Figure 6. Boxplots for the estimated amount of warping, as measured by the Fisher–Rao metric between the identity warping γ_{id} and the estimated warping function using unpenalized SRVF and penalized SRVF method, with $v = 8$ seconds for each video.

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press



(a) Scatterplots of correlation between target's ratings and perceivers' ratings for each video. In each plot, the horizontal axis represents the correlation without alignment, and the vertical axis represents the correlations under unpenalized SRVF (orange) and penalized SRVF with $\nu = 8$ seconds (green), respectively. The dashed line represents the 45-degree line.



(b) Pairwise confidence intervals for the mean difference between the estimated correlations obtained from unpenalized SRVF, optimal fixed, and no alignment and the proposed penalized SRVF method.

Figure 7. Results for EA estimates in the social EA study.

exhibited negative correlations between perceiver and target ratings under identity warping. As a data pre-processing step, we removed those cases based on the concern that they may exhibit fundamentally different empathy patterns from the general population of perceivers.

Figure 7a presents scatterplots comparing these EA measures between target and perceiver ratings for pre- and post-aligned data across the four video conditions. The majority of points reside above the 45-degree line, indicating that accounting for misalignment generally increases EA measures compared to unaligned analyses. However, the unpenalized SRVF alignment often inflates EA considerably, as observed in the simulation results This is most pronounced in the low intensity positive video group (bottom right panel of Figure 7a), where many unpenalized EA approach one, implying near-perfect empathy for most perceivers, an unrealistic outcome given the video's low expressiveness. Conversely, for the high intensity positive video group (top right panel), some unpenalized EA fell substantially below identity EA because excessive warping distorted overall function trends.

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press

Table 3. Estimated coefficients for Trait Positive Affect as a predictor of EA as measured by different alignment methods.

	High negative	Low negative	High positive	Low positive
No alignment	−0.013 (0.005)*	−0.003 (0.002)	−0.011 (0.006)	0.012 (0.005)*
Fixed delay	−0.017 (0.006)*	−0.004 (0.003)	−0.017 (0.007)*	0.005 (0.004)
Unpenalized SRVF	−0.009 (0.005)*	−0.002 (0.003)	−0.024 (0.011)*	0.019 (0.009)*
Penalized SRVF (8s)	−0.014 (0.005)*	−0.004 (0.003)	−0.016 (0.008)*	0.013 (0.005)*

Note: Standard errors are included in parentheses and * indicates significance at the significance level 5%.

The proposed penalized SRVF alignment provides a reasonable compromise between the identity EA and the unpenalized SRVF EA. For low-expressivity videos (bottom row, Figure 7a), penalized alignment EA measures generally exceed those from no alignment, likely due to increased misalignment challenges under reduced emotional cues. It is also interesting to find that for videos under positive emotion (second column, Figure 7a), the EA with the penalized SRVF alignment has a trivial difference compared with the EA with no alignment, while for videos under negative emotion (first column, Figure 7a), the difference is much bigger. This suggests a stronger time-warping effect for negative emotions, which is consistent with psychological research indicating better recognition of positive emotions (Bandyopadhyay et al., 2021). Figure 7b shows that the penalized SRVF methods lead to estimated EA with a significantly smaller mean than those obtained from unpenalized SRVF in all four videos. In three out of four videos, the mean EA obtained from the penalized SRVF is also significantly larger than those from no alignment, and generally differs significantly from those obtained under the optimal fixed delay method. In addition, Figure S1 in the Supplementary Material shows the correlations among the EA measures obtained by different alignment methods. Although they are positively correlated with others, they are not equivalent and our proposed alignment method can lead to improved inference in a downstream analysis.

We also examined the associations between perceiver-specific trait positive emotion and their EA. Trait positive emotion reflects a perceiver's stable tendency to experience positive emotions across diverse situations and over time, and it is typically associated with greater sociability, prosocial behavior, and openness (Devlin et al., 2014). To make this analysis consistent with the approach adopted by Devlin et al. (2014), we fitted a simple linear regression model with the Fisher transformed EA measure as the outcome and trait positive emotion as the predictor. Table 3 summarizes the estimated slope of each regression. It reveals that the penalized SRVF method consistently yields larger absolute coefficient estimates than the no-alignment approach across all four videos. This pattern is not consistently observed with the unpenalized SRVF or the fixed delay methods. These findings suggest that failing to properly address misalignment may obscure important relationships between EA and perceiver characteristics.

5.2. Study on music empathy

Tabak et al. (2022) conducted an EA study investigating three primary emotions: joy/happiness, sadness, and anger ($I = 3$). For each emotion, three original solo piano pieces ($J = 3$) were composed and performed by experienced musicians. These pieces were designed to evoke the target emotions within familiar musical styles (classical, popular, and jazz). Both musicians (as targets) and 123 participants (as perceivers) rated the emotional content of each piece on a 9-point scale. As with the previous dataset, we preprocessed the data by smoothing and interpolating rating functions.

Unlike the correlation-based approach, Tabak et al. (2022) employed a linear mixed-effect model for a more nuanced analysis of EA. This model decomposed perceiver responses into three latent factors: bias, discrimination, and variance. Bias represented the systematic deviation between perceiver and

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press

target ratings, while discrimination captured a perceiver's sensitivity to changes in the target's expressed emotion. Finally, variance accounted for random noise in perceiver ratings.

Within each group of emotion ($i = 1, \ldots, I$), let $x_j(\cdot)$ and $y_j(\cdot)$ be the target and a perceiver's ratings, respectively, for the jth stimulus. Tabak et al. (2022) proposed the following linear mixed-effect model to describe the relation between $x_j(\cdot)$ and $y_j(\cdot)$:

$$y_{jk} = \beta_0 + \beta_1 x_{jk} + b_{0j} + b_{1j} x_{jk} + \varepsilon_{jk},\ j = 1, \ldots, J,\ k = 1, \ldots, T_j, \tag{5.1}$$

where $y_{jk} = y_j(t_k)$ and $x_{jk} = x_j(t_k)$ are the perceiver and target's respective ratings at the kth time point, and T_j is the number of points for the jth stimuli. The (fixed) intercept β_0 and slope β_1 represent a perceiver's mean bias and discrimination ability across all the J stimuli, respectively. The random intercept b_{0j}, random slope b_{1j}, and the random noise ε_{jk} are assumed to follow a normal distribution with zero mean and respective variance component $\sigma^2_{b_0}, \sigma^2_{b_1}$, and σ^2, which represents the variability of bias, discrimination, and random noise across different stimuli. This model treats ratings as discrete points and does not account for potential misalignments between perceiver and target responses.

To address this limitation, we integrated an alignment step into the model framework. Treating the observed ratings as sampled points from corresponding functions, we applied and compared penalized and unpenalized time-warping SRVF alignments to account for potential misalignments. Let $\tilde{y}_j(t) = y_j \circ \hat{\gamma}_j(t)$ be the estimated aligned function with $\hat{\gamma}_j(t)$ being an estimated inverse warping function from aligning $y_j(\cdot)$ with $x_j(\cdot)$, we then modeled

$$\tilde{y}_j(t) = \beta_0 + \beta_1 x_j(t) + b_{0j} + b_{1j} x_j(t) + \varepsilon_j(t),\ b_{0j} \sim N(0, \sigma^2_{b_0}),\ b_{1j} \sim N(0, \sigma^2_{b_1}), \tag{5.2}$$

where β_0, β_1, $\sigma^2_{b_0}$, $\sigma^2_{b_1}$, and σ^2 in Model (5.2) maintain the same interpretations as in Model (5.1).

We fitted Model (5.2) for each perceiver and primary emotion. Using the `lme4` package in R (Bates et al., 2015), we employed restricted maximum likelihood estimation to obtain parameter estimates $\hat{\Psi} = (\hat{\alpha}, \hat{\beta}, \hat{\sigma}^2_{b_0}, \hat{\sigma}^2_{b_1}, \hat{\sigma}^2)$ and best linear unbiased predictions (BLUPs) of random effects $\hat{b}_{0j}$ and $\hat{b}_{1j}$ for $j = 1, \ldots, J$. To assess the impact of time warping, we compared parameter estimates $\hat{\Psi}$ under no alignment (i.e., γ_{id}), the unpenalized SRVF ($\hat{\gamma}_u$), and the penalized SRVF alignment ($\hat{\gamma}_p$), setting the penalty threshold at $v = 8$ seconds for the penalized alignment. Results for 6- and 10-second thresholds are provided in Section S3.2 of the Supplementary Material.

To assess model fit, we computed two metrics: average warping and average goodness of fit across all J tasks. The first metric, average Fisher–Rao distance, quantifies the mean warping magnitude relative to the identity warping: $\bar{d}_p = J^{-1} \sum_{j=1}^{J} \cos^{-1}\left(\int_0^1 \sqrt{\hat{\gamma}_j'(t)}dt\right)$. A higher $\bar{d}_p$ indicates greater warping. The second metric measures the vertical distance between the estimated aligned response function and the fitted value function. Specifically, letting $\hat{y}_j(\cdot) = \hat{\beta}_0 + \hat{\beta}_1 x_j(\cdot) + \hat{b}_{0j} + \hat{b}_{1j} x_j(\cdot)$, this vertical distance is calculated to be the $\mathbb{L}^2$ distance between the estimated aligned response $\tilde{y}_j$ and the fitted value function $\hat{y}_j$, i.e., $\sum_{j=1}^{J} \|\tilde{y}_j - \hat{y}_j\|_2^2$, where a lower value signifies a better model fit.

Figure 8 reveals that the no alignment model exhibits significantly inferior fit compared to the other two approaches (all the p-values for comparing pairwise mean differences are close to zero, see Table S4 in the Supplementary Material). This underscores the importance of addressing misalignment between perceiver and target ratings to prevent model underfitting. Similar to our simulation findings, the unpenalized SRVF alignment demonstrates overfitting, sacrificing model fit for excessive warping. In contrast, the penalized alignment method provides a reasonable compromise between these extremes, enhancing model fit while mitigating overfitting through judicious penalty application.

Figure 9a compares parameter estimates (aligned versus unaligned) for the fixed effect discrimination ($\hat{\beta}_1$) and random noise standard deviation ($\hat{\sigma}$) across the three emotion groups, and Figure 9b shows the pairwise confidence intervals in the mean estimates of the penalized SRVF method against other alignment methods. In the top row of both figures, while both unpenalized SRVF and penalized SRVF alignment methods tend to increase the discrimination estimates $\hat{\beta}_1$, the optimal fixed delay tends to decrease it compared to the no alignment. However, similar to the social empathy study in Section 5.1, the unpenalized SRVF alignment increases this discrimination estimate much more than

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press

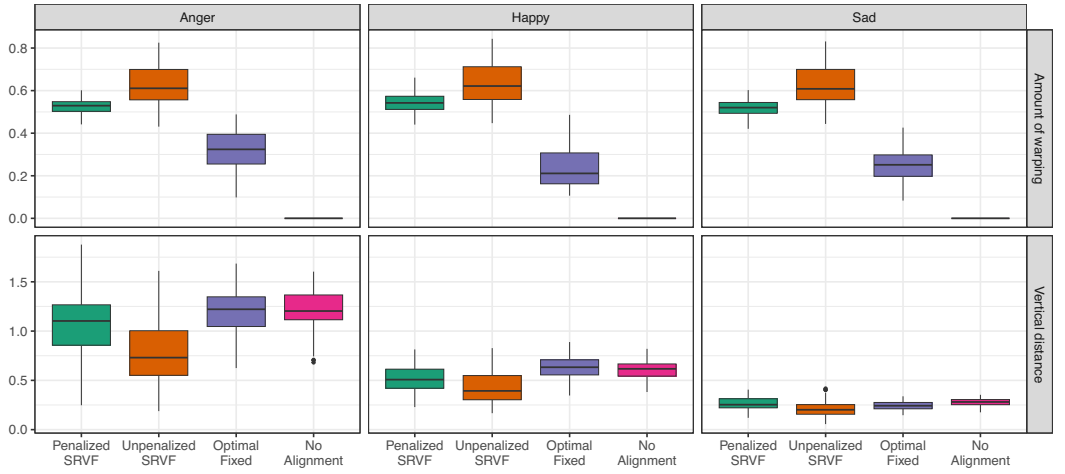


Figure 8. Boxplots of the metrics for the average amount of warping (top row) and goodness of fit (bottom row) of the estimated models for the three sets of music recordings. The penalized SRVF alignment was conducted using the threshold $\nu = 8s$.

the penalized SRVF, making the unpenalized SRVF more prone to overfit. This conclusion is further evidenced in the bottom row, where the estimated standard deviation $\hat{\sigma}$ from the unpenalized SRVF method is substantially smaller than that from the other two alignment methods.

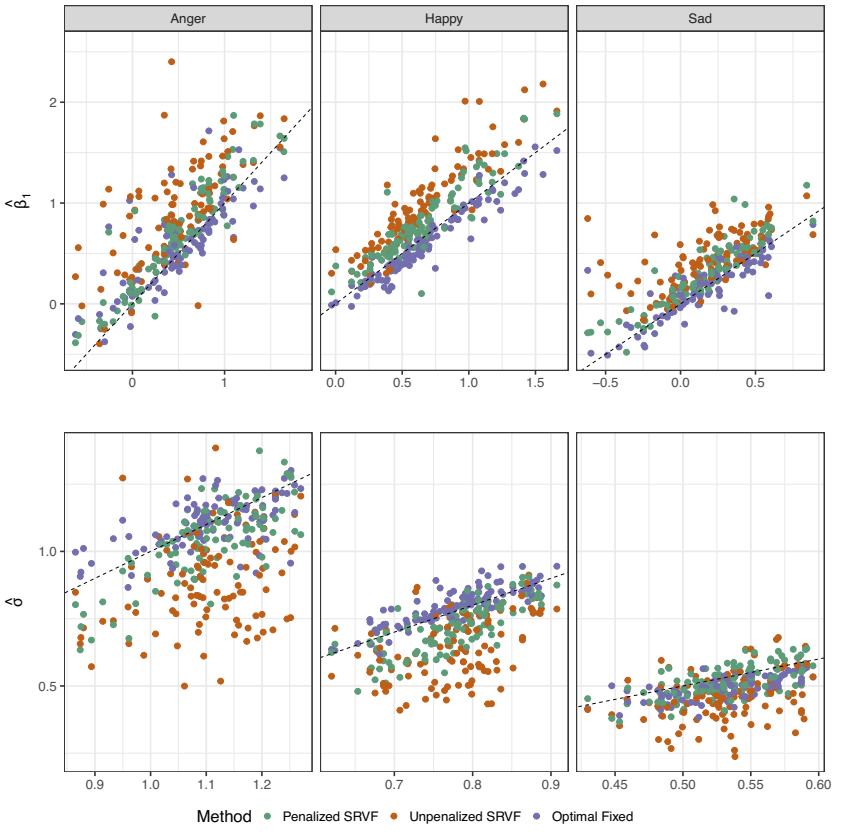
6. Discussion

In emotional perception research, misalignment caused by complex cognitive decoding processes and the time needed to enact a behavioral response is a well-established phenomenon. In EA studies, the discrepancy between the perceiver's observed rating and the target's rating is influenced by both the misalignment not due to lack of EA and the psychologically meaningful disagreement resulting from lack of EA. Yet, most of the conventional EA studies either ignore this kind of misalignment or apply an oversimplified fixed delay for adjustment, where both options can lead to biased results. This study introduces a novel, flexible approach using a new constrained optimization problem based on the SRVF representation of the functions to reduce the misalignment, which varies from individual to individual. Considering realistic conditions of the warping process, our simulation studies demonstrate that the proposed penalized SRVF alignment method provides improved estimates of the true EA measure compared to existing approaches. In two case studies on social and music empathy, this method yields plausible EA measures, which subsequently reveals more potential associations between EA and perceivers' characteristics.

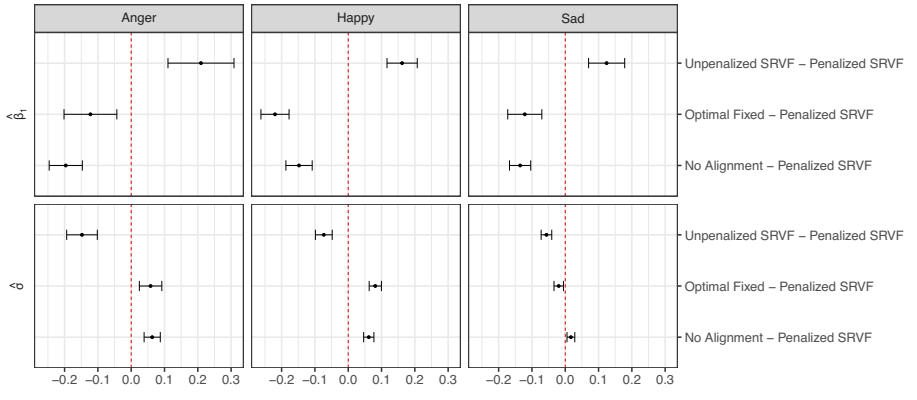
The proposed penalized alignment approach offers several advantages. 1) Individualized adjustments: It tailors alignment to unique patterns of misalignment for each perceiver. 2) Prevention of over-alignment: It incorporates a natural constraint on the extent of allowable warping. 3) Simplicity and interpretability: The penalty term can be easily set up by using the maximum delay in perceivers' responses, where it is straightforward to use empirical evidence and expert opinion. Moreover, EA studies often vary in context and focus. In situations, where reaction time is less critical, such as listening to a friend's long story, a broader penalization window may be appropriate. Conversely, in high-stakes scenarios like heated arguments, a narrower window can better reflect the urgency of responses. This flexibility enables researchers to tailor penalization parameters to the specific demands of each study. By integrating these features, our approach enhances the accuracy of downstream EA analyses, including correlational studies and complex linear mixed models.

The core component in our proposed method, the warping functions, has been widely employed to correct misalignment in fields like physics and biology, where objective benchmarks exist. To the

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press



(a) Scatterplots of estimates for the fixed effect discrimination β_1 (top) and random noise standard deviation σ (bottom). In each plot, the horizontal axis represents the estimate when no alignment is conducted, and the vertical axis represents the estimates under unpenalized SRVF (orange) and penalized SRVF with $\nu = 8s$ (green). The dashed line represents the 45-degree line.



(b) Pairwise confidence intervals for the mean difference between the estimates obtained from unpenalized SRVF, optimal fixed, and no alignment and the proposed penalized SRVF method.

Figure 9. Results for parameter estimates in the music EA study.

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press

best of our knowledge, the application of warping functions to the abstract and subjective domain of human perception is unexplored. In this study, we have demonstrated their effectiveness and flexibility in adjusting individually varying misalignment across different types of emotional stimuli (visual and audio). It further expands the potential for using warping functions in new research areas.

Future research could focus on several key areas. One potential direction is to model the similarity of warping functions of the same individual across different stimuli by introducing random effects. Another area of interest could be to develop a new EA alignment method by incorporating additional data, such as the functional magnetic resonance imaging (fMRI) blood-oxygen-level-dependent (BOLD) signals of targets and perceivers during rating assessments, which could help detect true emotional changes. Although our method accurately identified simulated noise and showed favorable psychometric characteristics, we caution against interpreting the corrected scores from our method as definitive indicators of EA devoid of all measurement error. Future work is needed to explicitly test the extent to which the penalized alignment approach can distinguish measurement noise from meaningful differences in EA, for example, by experimentally manipulating whether participants can pause the video to make ratings or by varying the cognitive load placed on participants. Finally, we note that these analyses were exploratory in nature. Given the methodological flexibility of the proposed method and the number of analytic decisions involved (e.g., the upper limit of warping functions), future work can aim to replicate these findings using pre-registered designs to increase confidence in the robustness of the results provided by the penalized SRVF alignment method.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/psy.2025.10040.

Data availability statement. The data used in the two case studies, along with the R codes that implement the penalized SRVF alignment method and produce numerical results in the article, are available at the GitHub repository: https://github.com/chulmoon/EA_Alignment.

Funding statement. This research received no specific grant funding from any funding agency, commercial or not-for-profit sectors.

Competing interests. The authors declare none.

References

Bandyopadhyay, A., Sarkar, S., Mukherjee, A., Bhattacherjee, S., & Basu, S. (2021). Identifying emotional facial expressions in practice: A study on medical students. *Indian Journal of Psychological Medicine*, *43*(1), 51–57.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

Berndt, D. J., & Clifford, J. (1994). Using dynamic time warping to find patterns in time series. In *Proceedings of the 3rd international conference on knowledge discovery and data mining* (pp. 359–370). AAAI Press.

Bharath, K., Kurtek, S., Rao, A., & Baladandayuthapani, V. (2018). Radiologic image-based statistical shape analysis of brain tumours. *Journal of the Royal Statistical Society. Series C: Applied statistics*, *67*(5), 1357.

De Waal, F. B. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology*, *59*, 279–300.

Devlin, H. C., Zaki, J., Ong, D. C., & Gruber, J. (2014). Not as good as you think? trait positive emotion is associated with increased self-reported empathy but decreased empathic performance. *PloS One*, *9*(10), e110470.

Ekman, P. (1992). An argument for basic emotions. *Cognition & Emotion*, *6*(3–4), 169–200.

Ghosal, R., Maity, A., Clark, T., & Longo, S. B. (2020). Variable selection in functional linear concurrent regression. *Journal of the Royal Statistical Society Series C: Applied Statistics*, *69*(3), 565–587.

Gunes, H., & Pantic, M. (2010). Automatic, dimensional and continuous emotion recognition. *International Journal of Synthetic Emotions*, *1*(1), 68–99.

Gunes, H., & Schuller, B. (2013). Categorical and dimensional affect analysis in continuous input: Current trends and future directions. *Image and Vision Computing*, *31*(2), 120–136.

Guo, X., Wu, W., & Srivastava, A. (2022). Data-driven, soft alignment of functional data using shapes and landmarks. Preprint, arXiv:2203.14810.

Huang, Z., Dang, T., Cummins, N., Stasak, B., Le, P., Sethu, V., & Epps, J. (2015). An investigation of annotation delay compensation and output-associative fusion for multimodal continuous emotion prediction. In *Proceedings of the 5th international workshop on audio/visual emotion challenge* (pp. 41–48). Association for Computing Machinery.

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press

Ickes, W. J. (1997). *Empathic accuracy*. Guildford Press.
Jospe, K., Genzer, S., Klein Selle, N., Ong, D., Zaki, J., & Perry, A. (2020). The contribution of linguistic and visual cues to physiological synchrony and empathic accuracy. *Cortex*, *132*, 296–308.
Khorram, S., McInnis, M. G., & Provost, E. M. (2019). Jointly aligning and predicting continuous emotion annotations. *IEEE Transactions on Affective Computing*, *12*(4), 1069–1083.
Kneip, A., Li, X., MacGibbon, K., & Ramsay, J. (2000). Curve registration by local regression. *Canadian Journal of Statistics*, *28*(1), 19–29.
Kokoszka, P., & Reimherr, M. (2017). *Introduction to functional data analysis*. Chapman and Hall/CRC.
Koster-Hale, J., & Saxe, R. (2013). Theory of mind: A neural prediction problem. *Neuron*, *79*(5), 836–848.
Laga, H., Kurtek, S., Srivastava, A., & Miklavcic, S. J. (2014). Landmark-free statistical analysis of the shape of plant leaves. *Journal of Theoretical Biology*, *363*, 41–52.
Lee, J., Zaki, J., Harvey, P.-O., Ochsner, K., & Green, M. (2011). Schizophrenia patients are impaired in empathic accuracy. *Psychological Medicine*, *41*(11), 2297–2304.
Levenson, R. W. (1988). Emotion and the autonomic nervous system: A prospectus for research on autonomic specificity. In *Social psychophysiology and emotion: Theory and clinical applications* (pp. 17–42). John Wiley & Sons.
Mackes, N. K., Golm, D., O'Daly, O. G., Sarkar, S., Sonuga-Barke, E. J., Fairchild, G., & Mehta, M. A. (2018). Tracking emotions in the brain—Revisiting the empathic accuracy task. *NeuroImage*, *178*, 677–686.
Mariooryad, S., & Busso, C. (2014). Correcting time-continuous emotional labels by modeling the reaction lag of evaluators. *IEEE Transactions on Affective Computing*, *6*(2), 97–108.
Marron, J. S., Ramsay, J. O., Sangalli, L. M., & Srivastava, A. (2015). Functional data analysis of amplitude and phase variation. *Statistical Science*, *30*(4), 468–484.
Matuk, J., Bharath, K., Chkrebtii, O., & Kurtek, S. (2022). Bayesian framework for simultaneous registration and estimation of noisy, sparse, and fragmented functional data. *Journal of the American Statistical Association*, *117*(540), 1964–1980.
Mitchell, E. G., Dryden, I. L., Fallaize, C. J., Andersen, R., Bradley, A. V., Large, D. J., & Sowter, A. (2025). Object oriented data analysis of surface motion time series in peatland landscapes. *Journal of the Royal Statistical Society Series C: Applied Statistics*, *74*, 406–428.
Mörters, P., & Peres, Y. (2010). *Brownian motion* (vol. 30). Cambridge University Press.
Nicolle, J., Rapp, V., Bailly, K., Prevost, L., & Chetouani, M. (2012). Robust continuous prediction of human emotions using multiscale dynamic cues. In *Proceedings of the 14th ACM international conference on Multimodal interaction* (pp. 501–508). Association for Computing Machinery.
Olofsson, J. K., Nordin, S., Sequeira, H., & Polich, J. (2008). Affective picture processing: An integrative review of ERP findings. *Biological Psychology*, *77*(3), 247–265.
Ramsay, J. O., & Li, X. (1998). Curve registration. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *60*(2), 351–363.
Ramsay, J. O., & Silverman, B. W. (2005). *Functional data analysis*. Springer.
Ringeval, F., Eyben, F., Kroupi, E., Yuce, A., Thiran, J.-P., Ebrahimi, T., Lalanne, D., & Schuller, B. (2015). Prediction of asynchronous dimensional emotion ratings from audiovisual and physiological data. *Pattern Recognition Letters*, *66*, 22–30.
Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, *26*(1), 43–49.
Scherer, K. R. (2003). Vocal communication of emotion: A review of research paradigms. *Speech Communication*, *40*(1–2), 227–256.
Schweinle, W. E., Ickes, W., & Bernstein, I. H. (2002). Emphatic inaccuracy in husband to wife aggression: The overattribution bias. *Personal Relationships*, *9*(2), 141–158.
Sened, H., Lavidor, M., Lazarus, G., Bar-Kalifa, E., Rafaeli, E., & Ickes, W. (2017). Empathic accuracy and relationship satisfaction: A meta-analytic review. *Journal of Family Psychology*, *31*(6), 742.
Srivastava, A., & Klassen, E. P. (2016). *Functional and shape data analysis*. Springer.
Srivastava, A., Wu, W., Kurtek, S., Klassen, E., & Marron, J. S. (2011). Registration of functional data using Fisher-Rao metric. Preprint, arXiv:1103.3817.
Su, J., Kurtek, S., Klassen, E., & Srivastava, A. (2014). Statistical analysis of trajectories on Riemannian manifolds: Bird migration, hurricane tracking and video surveillance. *The Annals of Applied Statistics*, *8*(1), 530–552.
Tabak, B. A., Wallmark, Z., Nghiem, L. H., Alvi, T., Sunahara, C. S., Lee, J., & Cao, J. (2022). Initial evidence for a relation between behaviorally assessed empathic accuracy and affect sharing for people and music, *23*(2), 437–449. *Emotion.*
Thornton, M. A., & Tamir, D. I. (2017). Mental models accurately predict emotion transitions. *The Proceedings of the National Academy of Sciences*, *114*(23), 5982–5987.
Tucker, J. D. (2025). fdasrvf: Elastic functional data analysis. R package version 2.3.6.
Tucker, J. D., Wu, W., & Srivastava, A. (2014). Analysis of proteomics data: Phase amplitude separation using an extended Fisher-Rao metric. *Electronic Journal of Statistics*, *8*, 1724–1733.
Wallace, W. E., Srivastava, A., Telu, K. H., & Simón-Manso, Y. (2014). Pairwise alignment of chromatograms using an extended Fisher–Rao metric. *Analytica Chimica Acta*, *841*, 10–16.

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press

Wang, K., & Gasser, T. (1997). Alignment of curves by dynamic time warping. *The Annals of Statistics*, *25*(3), 1251–1276.
Wu, W., & Srivastava, A. (2011). An information-geometric framework for statistical inferences in the neural spike train space. *Journal of Computational Neuroscience*, *31*, 725–748.
Wu, W., & Srivastava, A. (2014). Analysis of spike train data: Alignment and comparisons using the extended Fisher-Rao metric. *Electronic Journal of Statistics*, *8*, 1776–1785.
Xie, W., Kurtek, S., Bharath, K., & Sun, Y. (2017). A geometric approach to visualization of variability in functional data. *Journal of the American Statistical Association*, *112*(519), 979–993.
Zaki, J., Bolger, N., & Ochsner, K. (2008). It takes two: The interpersonal nature of empathic accuracy. *Psychological Science*, *19*(4), 399–404.
Zaki, J., Bolger, N., & Ochsner, K. (2009). Unpacking the informational bases of empathic accuracy. *Emotion*, *9*(4), 478.
Zhao, W., Xu, Z., Li, W., & Wu, W. (2020). Modeling and analyzing neural signals with phase variability using Fisher-Rao registration. *Journal of Neuroscience Methods*, *346*, 108954.

Cite this article: Nghiem, L. H., Cao, J., Kouros, C. D. and Moon, C., (2025). Enhancing Empathic Accuracy: Penalized Functional Alignment Method to Correct Temporal Misalignment in Real-Time Emotional Perception. *Psychometrika*, 1536–1557. https://doi.org/10.1017/psy.2025.10040

https://doi.org/10.1017/psy.2025.10040 Published online by Cambridge University Press