The validated touch-video database

Sophie Smit[1,2] & Anina N. Rich[1,2,3]

1 School of Psychological Sciences, Macquarie University, Sydney, Australia
2 Perception in Action Research Centre, Macquarie University, Sydney, Australia
3 Macquarie University Centre for Performance and Expertise, Macquarie University, Sydney, Australia

Abstract

Understanding visual perception of touch is complex. Recent research has honed in on vicarious sensory perception—sensations triggered by watching touch or pain in others. Factors like arousal, threat, and touch type appear to influence these experiences. However, the absence of standardised visual touch stimuli restricts in-depth analysis. Addressing this, we introduce the Validated Touch-Video Database (VTD), which features 90 videos of touch to a single left hand from a first-person perspective that varies in arousal, threat, hedonic qualities, touch type, and touch agent (object versus human). The database was validated in 350 participants (283 women, 66 men, 1 non-binary), who categorised the touch in each video as 'neutral', 'pleasant', 'unpleasant', or 'painful', and provided arousal and threat ratings. Using different methods to check for consistency, we found high agreement between participants in their ratings, highlighting the reliability of the validation data. Our findings indicate that the videos were effective in depicting touch with different hedonic qualities, with differences in how threatening or arousing each type of touch appeared. For example, videos showing painful touch had the highest arousal and threat ratings. On the other hand, neutral touch had very low ratings for both aspects, setting a helpful baseline when comparing more emotionally charged touch types. Women were more likely than men to rate touches as 'threatening' or 'painful', highlighting the need to consider gender differences when looking at perceptions and reactions to tactile sensory cues. The VTD is relevant for research ranging from understanding our reactions when witnessing others being touched, to the interaction of visual cues with other touch-related sensory data, or decoding neural representations of different touch types. It may also be useful in examining typical and atypical self-other perception, such as that seen in depersonalisation disorder and conditions on the autism spectrum. The inclusion of pleasant touch stimuli broadens the scope of research, offering potential applications in therapy and virtual domains. The VTD, including the validation data, is freely available online for other researchers, providing a systematic resource for researchers exploring touch observation.

1. Introduction

Vision plays an important role in the processing of tactile information, shaping our perception of touch across a wide array of scenarios, from our own bodily self-awareness to interpersonal touch interactions, and extending to clinical applications. To understand the complex interactions between vision and touch in the context of body-ownership, researchers have largely relied on the rubber hand illusion, a perceptual phenomenon in which a person's brain can be tricked into perceiving a rubber or artificial hand as their own, typically through synchronous tactile stimulation of the fake hand and their hidden real hand (Botvinick and Cohen, 1998). Adaptations of this paradigm have been used to examine the impact of observing different types of touch on tactile processing, such as roughness perception (Schütz-Bosbach et al., 2009) and affective touch (Schütz-Bosbach et al., 2009). Additionally, this paradigm has been used to study tactile processing in specific populations, such as individuals with mirror-touch synaesthesia, who feel touch when they see others being touched (Davies and White, 2013). Other research has focused on neural responses to observed tactile stimuli, including the perception of roughness (Kuehn et al., 2013), the differentiation between pleasant and unpleasant touch (Lamm et al., 2015; Morrison et al., 2011; Riva et al., 2018), and social touch interactions (Lee Masson et al., 2018; Peled-Avron et al., 2016). In clinical settings, visual touch cues can be employed beneficially; for instance, the mirror-box illusion technique has been known to alleviate phantom pain via visual feedback (Ramachandran and Altschuler, 2009; Ramachandran and Rodgers-Ramachandran, 1996). This diverse line of research inquiries collectively enhances our comprehension of how visual information influences our perception of touch, both in individual and inter-personal contexts, spanning clinical and non-clinical domains.

Within the realm of visual perception of touch, there has been an increasing focus on studying a phenomenon known as vicarious sensory perception, where individuals experience a similar sensory state to what they observe in others (for reviews see Bufalari and Ionta, 2013; Gillmeister et al., 2017). To investigate vicarious touch, previous studies have often presented participants with videos depicting touch on another person, including touch with a

finger or knife to the left and right cheek, left and right hands, crossed hands, and an inverted face (e.g., using stimuli from Blakemore et al., 2005; Holle et al., 2011). These studies show that vicarious tactile experiences are reported by a large portion of the general population (Li et al., 2022; Ward et al., 2018; Ward and Li, 2022). Additionally, studies have explored vicarious responses to videos showing individuals scratching their chest or upper arm (Holle et al., 2012; Ward et al., 2018). Vicarious pain has been studied by showing participants images or videos of someone getting injured, such as rolling an ankle, or dislocating a knee, or receiving an injection (e.g., using stimuli from Giummarra et al., 2015; Grice-Jackson et al., 2017; Osborn and Derbyshire, 2010). Approximately one-third of the general population report experiencing sensations like 'pain', 'tingling,' 'aching,' or 'shooting' when observing these stimuli, suggesting that vicarious pain is similarly common (Giummarra et al., 2015; Grice-Jackson et al., 2017; Osborn and Derbyshire, 2010; Ward and Li, 2022). In response to watching touch, itch, and pain stimuli, individuals tend to report a wide range of vicarious sensations, including 'touch', 'pain', 'pressure', and 'scratching'; in addition to affective states like 'flinching', 'unease', or 'itchiness' (Grice-Jackson et al., 2017; Li et al., 2022; Smit et al., 2023b; Ward et al., 2018; Ward and Li, 2022). These findings suggest that vicarious touch and pain are widespread and offer a unique perspective into understanding how vision and touch interact.

To date, most of this research has focused on sensations reported when participants observe neutral/threatening touch or pain in others (Giummarra et al., 2015; Grice-Jackson et al., 2017; Osborn and Derbyshire, 2010; Vandenbroucke et al., 2013; Ward et al., 2018; Ward and Li, 2022). In contrast, there is relatively little research on vicarious responses to pleasant touch. Although fMRI studies have advanced our understanding of the neural underpinnings of empathy for both pleasant and unpleasant tactile sensations (Ebisch et al., 2011; Lamm et al., 2015; Morrison et al., 2011; Riva et al., 2018), their focus has been more on the neural dimensions, leaving room for further examination of the nuanced subjective experiences tied to observing touches with different hedonic qualities. A recent systematic review and meta-analysis identified key factors that tend to elicit particularly pleasurable sensations during

direct tactile experiences, which includes soft materials and brushes stroking at a velocity of 3 cm/s with gentle force (Taneja et al., 2021). Visual touch stimuli that capture such elements could prove valuable in multiple applications, from exploring the therapeutic benefits of vicarious touch (Giummarra et al., 2016; Makary et al., 2018; Ramachandran and Altschuler, 2009; Ramachandran and Rodgers-Ramachandran, 1996) to understanding its role in enhancing consumer experiences in digital marketing and virtual environments (Luangrath et al., 2022). Access to a comprehensive collection of visual touch stimuli, portraying both unpleasant and pleasant tactile interactions, would allow a more comprehensive understanding of vicarious tactile experiences.

Previous studies provide preliminary evidence that factors such as arousal, threat, and valence may impact how visual touch stimuli are perceived and processed. For instance, stimuli depicting pain or threatening touch experiences, such as a knife tip touching a face, have been shown to elicit more frequent and intense vicarious responses compared to stimuli showing touch with non-threatening objects like a fingertip or feather (Holle et al., 2011; Ward et al., 2018; Ward and Li, 2022). Apart from emotional factors, research has also shown that physical aspects like congruent body posture and the viewer's perspective can modulate vicarious sensory perception (Holle et al., 2011; Medina and DePasquale, 2017). Studies often incorporate a wide range of visual touch stimuli that vary in perceived arousal and threat, the body part touched, hedonic qualities, viewing perspective, and the plausibility of the visual cue (e.g., touch to an upright face versus an inverted one). The diversity of stimuli across different studies can present challenges in identifying systematic patterns.

Controlled stimuli for studying responses to *interpersonal* touch have been developed by Masson and Op de Beeck (2018) in the Socio-Affective Touch Expression Database (SATED). This database comprises a collection of carefully controlled and systematically defined videos, featuring pairs of actors engaging in various social touch expressions with varying levels of valence and arousal. It provides a valuable resource for examining reactions to social touch interactions. Here, we followed a similar method to develop and validate an

extensive database of highly controlled touch videos to facilitate further research into visually perceived touch and vicarious sensory experiences.

Our Validated Touch-Video Database (VTD) contains 90 systematically controlled videos depicting different types of touch to a hand. The touch videos vary across multiple dimensions, including levels of arousal, threat, hedonic qualities, as well as the nature of the touch interaction—including touch applied by an object or another hand. Consistency is maintained across the videos by including a single female left hand, always presented from a first-person perspective in a standardised setup. To validate the effectiveness of these videos, we recruited 350 participants who were asked to categorise the touch displayed in each video as neutral, pleasant, unpleasant, or painful and to rate the strength of the indicated valence. Additionally, participants rated the perceived levels of arousal and threat for each video. Our analysis assesses the inter-subject consistency and reliability of video categorisation and ratings. Further, as existing research indicates that men and women respond differently to observed pain (Grice-Jackson et al., 2017; Singer et al., 2006a; Ward and Li, 2022; Yang et al., 2009), we examined potential gender differences for the different indices. A primary objective of this database is to offer videos that represent a range of hedonic qualities in tactile experiences. To focus on this aspect, we curated and analysed a selection of the top 10 videos in each hedonic category—neutral, pleasant, unpleasant, and painful. These selections were determined based on the most consistent categorisations across participants. To promote further research into visual perception of touch, our videos, along with their validation data, are freely accessible online.

Videos: OSF (https://osf.io/jvkqa/) and YouTube (https://bit.ly/YouTubeVTD)

Validation data: GitHub (https://sophiesmit1.github.io/VTD/)

2. Materials and Methods

2.1. Participants

Participants were recruited via the Macquarie University subject pools and consisted of undergraduate students naïve to the purpose of the study (N = 350; 283 women, 66 men; 1 non-binary, mean age = 24 years, SD = 8.66 years, range = 17-63 years). As we only had one non-binary participant, we could not include their data in our analyses of gender. Due to the large number of videos, some of which were unpleasant to watch, we set up two separate surveys with 45 videos each, each survey comprising of an equal number of neutral, pleasant, unpleasant, and painful videos. Each survey took approximately 15 minutes to complete. Participants were free to decide whether they wanted to participate in one or both surveys and received course credit accordingly. The study was approved by the Macquarie University Human Research Ethics committee and participants provided written consent.

2.2. Stimuli

The materials consisted of 90 short videos showing touch to a Caucasian female hand placed on a dark background. The hand was always presented in the first-person perspective with the palm facing down. Touch was either applied by another hand (e.g., a finger or flat hand) or by an object (e.g., a soft brush, scissors). To create pleasant videos, we showed a slow stroking all over the hand with either a brush or fingers. Some of the videos show a painful or uncomfortable touch, for instance an injection with a syringe or a pinch with tweezers. We attempted to make these videos as realistic as possible (suggesting painful touch) without inflicting any real pain or discomfort to the model.

2.3. Experimental procedure

The questionnaire was administered through the Qualtrics online platform, and participants completed it remotely on their personal devices, with the instruction to use a desktop computer or laptop instead of a phone or tablet. Videos were presented in a random order. To ensure

participant readiness and full attention throughout each video, auto-play was disabled. The initial frame of every video displayed a white screen with the instruction, “Press play to start video”. Participants were then asked to evaluate the touch depicted in the videos based on various criteria, including hedonic qualities, arousal, and perceived threat.

After watching each video, participants were asked the following four questions:

Q1: *How would you categorise the touch in this video?* [single answer out of the following checkbox options: neutral, pleasant, unpleasant, painful]

Q2: *How [pleasant/unpleasant/painful] (category from the previous answer) was the touch?* [answer: 1 = not at all, 10 = extremely]

Q3: *How threatening was the touch?* [answer: 1 = not at all, 10 = extremely]

Q4: *How arousing was this video? (Arousing in the form of a feeling, emotion, or response)* [answer: 1 = not at all, 10 = extremely]

2.4. Analyses

In the following analyses, we characterise our video database using multiple methods. Initially, we assess inter-subject consistency and dataset reliability to determine the uniformity of participants’ categorisation consistency based on hedonic attributes, as well as their ratings concerning perceived arousal and threat. We then report gender-specific variations in video ratings, based on previous existing studies that have observed gender differences in reactions to observed pain in others (Grice-Jackson et al., 2017; Singer et al., 2006a; Ward and Li, 2022; Yang et al., 2009). Our goal in these gendered analyses is to identify whether responses vary with specific video elements, such as arousal, and provide a normative dataset for future research.

Next, we analyse the different hedonic qualities—neutral, pleasant, unpleasant, and painful touches— and probe the association between the hedonic qualities and corresponding levels of perceived arousal and threat. To do this, we select the top 10 videos most consistently categorised under each hedonic descriptor and test distinctions between categories. We examine variations in the intensity of hedonic qualities across different

categories and explore gender-based differences within each respective category. Guided by prior research, we anticipate that videos with pronounced emotional valence, whether pleasant or unpleasant, will induce higher arousal levels (e.g., Masson and Op de Beeck, 2018). Lastly, we compile our findings into a user-friendly summary table, accessible through an R Markdown file. This dynamic resource enables users to sort videos based on hedonic quality, threat, and arousal, which can then be downloaded in multiple formats. It also features a search function, streamlining the process of selecting videos for subsequent research.

For all analyses, we first assessed the normality of the data using Shapiro-Wilk tests, which consistently revealed non-normal distributions. Consequently, non-parametric testing is employed throughout. Additionally, we implement Bonferroni correction to address multiple comparisons. All statistical analyses are executed using the R programming language (R Core Team, 2023).

2.4.1. Inter-subject consistency & dataset reliability

We evaluate the consistency and reliability of participants' video ratings, focusing on perceived levels of arousal and threat associated with touch across the 90 videos in our database. We use the Intraclass Correlation Coefficient (ICC) as this considers both correlation and agreement among measurements, thereby offering a measure of reliability (Koo and Li, 2016). Further, we leverage the Spearman-Brown split-half reliability test (Brown, 1910; Spearman, 1910) to investigate the internal consistency of the data related to arousal and threat perception. In using this method, we draw from and adapt approaches used in a prior study that validated stimuli in social-affective touch research (Masson and Op de Beeck, 2018). Finally, we use Fleiss' Kappa (Landis and Koch, 1977) to ascertain the degree of agreement among participants in categorising the videos into various emotional categories: neutral, pleasant, unpleasant, or painful.

<u>The Intraclass Correlation Coefficient</u>

The ICC serves as our principal metric for assessing the reliability of video ratings, capturing the variance across individual videos as well as among raters. We employ a two-way mixed-effects model (ICC3k; see Koo and Li, 2016 for model details) for mean-rating consistency to compute ICC estimates along with their 95% confidence intervals, using the R package 'Psych' (Revelle and Revelle, 2015). A high ICC value signifies strong agreement among participants, indicating reliable and consistent assessments of perceived arousal and threat in the videos. However, it is important to acknowledge that the ICC can be susceptible to biases, such as systematic rating differences among participants or constrained variability in the ratings, which can skew its value. To address these challenges, our approach combines means and medians: we primarily use means for ratings to provide an overall measure of central tendency, but in cases where a participant's rating is considered an outlier, we replace that specific outlier with the participant's median rating to mitigate the impact of extreme values. This combined approach allows us to strike a balance between robustness against outliers and retaining the central tendency of the ratings, essential for ICC calculations. Additionally, we apply Min-Max normalisation to rescale all ratings to a 0-to-1 range, ensuring that each participant's ratings contribute equally to the ICC calculation and meeting the ICC's assumption of normally-distributed data. This approach enhances the reliability of our findings by addressing potential distribution issues while maintaining robustness in our analysis.

The Spearman-Brown split-half reliability test

The Spearman-Brown split-half reliability test is used to evaluate the internal consistency of our arousal and threat data, thereby further validating the video set (see also Masson and Op de Beeck, 2018). First, participants are randomly partitioned into two equally sized groups. For each group, we compute the medians separately for arousal and threat ratings. A Spearman rank-order correlation is then calculated between the median ratings of these two subgroups to examine the relationship between the divided datasets. To account for the potential underestimation of the correlation coefficient—due to splitting the data—we apply the Spearman-Brown prophecy formula (Brown, 1910; Spearman, 1910). This adjustment

provides a more accurate estimate of what the reliability would be for the full-length dataset based on its half-length subsets. To bolster the robustness of our estimates, this entire process is iteratively conducted 100 times. In each iteration, an adjusted correlation value is derived via the Spearman-Brown formula. We summarise the findings by determining the median of these adjusted correlation coefficients. A higher median value signifies better internal consistency and, by extension, the greater reliability of the arousal and threat data.

Fleiss' Kappa

Fleiss' Kappa is used to assess the consistency of participants' categorisation of videos into four distinct hedonic touch types—neutral, pleasant, unpleasant, and painful. This statistical measure is well-suited for evaluating inter-rater agreement on categorical data from multiple raters (Landis and Koch, 1977). Initially, for each of the 90 videos in our study, we calculated the proportion of participants who classified the video into one of these four categories. We then selected the 10 videos that received the highest level of participant agreement, as evidenced by the highest percentages of consistent classifications for the four categories. Our analysis for consistency extends to both the complete set of 90 videos as well as this refined subset of 40 videos. Fleiss' Kappa values can range from -1 to 1, serving as an interpretative scale for the strength of agreement among participants. Negative Kappa values suggest less agreement than would be expected by chance, 0 indicates agreement equivalent to chance, and positive values signify better agreement than chance. Positive values are typically interpreted as indicating poor (0 to 0.2), slight (0.21 to 0.4), fair (0.41 to 0.6), moderate (0.61 to 0.8), substantial (0.81 to 1.0), or almost perfect agreement (Landis and Koch, 1977).

2.4.2. Hedonic qualities, threat & arousal

To investigate potential gender-specific differences in the perception of our video stimuli, we examine the hedonic categorisations along with arousal and threat ratings across all 90 videos, stratified by the gender of the participant. For arousal and threat metrics, we calculate the mean and standard deviation for each group and employ Mann-Whitney U tests to

statistically assess the differences. In addition, we use two-sample tests for equality of proportions to contrast how men and women categorise the videos across the four hedonic types: neutral, pleasant, unpleasant, and painful. Through these analyses, we aim to gain deeper insights into gender-based variations in hedonic categorisation as well as in the perception of arousal and threat.

We then look at differences in the strength of hedonic quality across categories and investigate gender-specific differences within each category. Additionally, we examine the relationship between these hedonic categories and participants' ratings for arousal and threat. To ensure a robust analysis, we limit our analysis to the subset of 40 videos that showed the highest degree of participant agreement in their initial hedonic categorisation. For assessing the strength of hedonic qualities, we use participants' responses to a question that asked about the level of pleasantness, unpleasantness, or painfulness of the touch depicted in each video. This question is contingent on their initial categorisation and excludes videos categorised as neutral. We statistically compare mean ratings across different hedonic categories and genders. Additionally, we compare mean arousal and threat ratings for the top 10 videos in the different hedonic categories.

2.4.3. Video validation & accessibility

We synthesised the results into an easily navigable summary table that features both a touch description for each video and the corresponding YouTube link. To optimise user accessibility and data organisation, this table is available as an R Markdown file (see https://osf.io/jvkqa/ and Fig. 6). Within this file, users can sort the collection of 90 videos based on key metrics—such as hedonic quality, threat, and arousal—either ascending or descending. Moreover, the R Markdown file incorporates a search functionality, allowing users to find videos that align with specific keywords or criteria. This feature streamlines the process for users who wish to curate videos based on targeted characteristics. For added convenience, the table is also downloadable in multiple formats, including CSV, Excel, and PDF.

3. Results

3.1. Inter-subject consistency & dataset reliability

3.1.1. High inter-subject agreement for arousal and threat ratings for all 90 videos

Assessing the reliability of video ratings for the 90 videos from our entire sample demonstrates that there is sufficient consistency to proceed with further analyses. The ICC analysis yielded high values for both arousal and threat ratings, suggesting a strong level of agreement among participants' ratings (values greater than 0.90 indicate excellent reliability; Koo and Li, 2016). Across all participants, the ICC values were very high for both arousal (ICC = 0.98, 95% CI: 0.98 - 0.99) and threat (ICC = 0.99, 95% CI: 0.99 - 1.00), with very similar values after normalisation (arousal ICC = 0.97, 95% CI: 0.96 - 0.98; threat ICC = 0.99, 95% CI: 0.99 - 1.00). The findings suggest that the participants' rating consistency for arousal and threat levels in the video stimuli is robust and minimally affected by systematic bias or distribution-related issues. This pattern is largely consistent across genders (see Supplementary Materials), suggesting that both men and women rated the videos in a highly consistent manner. These finding suggest that the validation dataset gives reliable estimates of the arousal and threat ratings of the videos.

The Spearman-Brown split-half reliability test on data from the two separate questionnaires (45 videos each) for both arousal and threat variables assessed the internal consistency of the ratings. We subjected each dataset to 100 iterations of the split-half method and then computed the reliability coefficients. The arithmetic mean of the reliability coefficients for each questionnaire gives an overall measure of reliability. We found high reliability coefficients for arousal (rSB = 0.96) and threat (rSB = 0.99), demonstrating a very high level of internal consistency for both variables. Participant ratings of the videos are therefore highly reliable.

These tests provide complementary insights into the reliability of participants' video ratings. The ICC showed high levels of agreement among participants, confirming the robustness of the ratings and minimising concerns about bias or distribution issues. The

Spearman-Brown test corroborated these findings by revealing high internal consistency. Together, these tests confirm that the participants' ratings are both internally consistent and mutually agreed upon, suggesting that the validation data is a good indicator for how the videos will be perceived in terms of arousal and threat.

3.1.2. Moderate inter-subject agreement for hedonic categorisation

The inter-rater reliability analysis to evaluate agreement among the participants in categorising videos as neutral, pleasant, unpleasant, or painful was first performed on all 90 videos, which produced a Fleiss' Kappa value of 0.274 (z-value = 534, p-value < 0.001). This value is typically considered to indicate 'fair' agreement among the raters (Landis and Koch, 1977). This somewhat low indication of agreement for categorisation seems likely due to some videos being less clearly categorised than others. Therefore, we selected the top 10 from each category (neutral, pleasant, unpleasant, painful) for inclusion in a 40-video subset of the database with reliable categorisation of hedonic qualities.

We selected the top 10 videos with the highest percentage of participants categorising the touch in the video as neutral, pleasant, unpleasant, or painful. As would be expected, when considering this subset of 40 videos, the Fleiss' Kappa value was higher with 0.387 (z-value = 516, p-value < 0.001), indicating 'moderate' agreement among the raters (Landis and Koch, 1977). When diving by genders, we ended up with very similar results; women = 0.389 (z-value = 422, p < 0.001), and men = 0.348 (z-value = 79.6, p < 0.001). After identifying the most clearly defined hedonic quality category for each video, we calculated the average percentage range of participants who categorised the video as such across the top 10 videos (detailed in Supplementary Materials Table 1). The distribution across categories was as follows: neutral (ranging from 71% to 84% agreement), pleasant (ranging from 60% to 89% agreement), unpleasant (ranging from 57% to 66% agreement), and painful (ranging from 41% to 84% agreement). As an example, these results show that in the neutral category, the most clearly categorised video was labelled as 'neutral' by 84% of participants, indicating strong agreement, whereas the 10th most clearly categorised video in this category received

agreement from 71% of participants, still showing a consistent but slightly lower level of agreement. Overall, this shows substantial agreement across videos in the top 10 for the neutral category. Further inspection shows that some videos, especially in the painful category, were categorised as both ‘painful’ and ‘unpleasant’ by similar proportions of the sample (the overlap was mostly painful/unpleasant and pleasant/neutral). However, all videos in the top 10 for the neutral, pleasant, and unpleasant category, and the top 6 for the painful category, are categorised by the majority of participants, indicating a clear distinction for the hedonic qualities.

3.2. Hedonic qualities, threat & arousal

3.2.1. Gender-specific differences in video ratings for threat but not arousal

We compared arousal and threat ratings for all 90 videos by gender, which showed a significant difference for perceived threat but not arousal (Fig. 1).

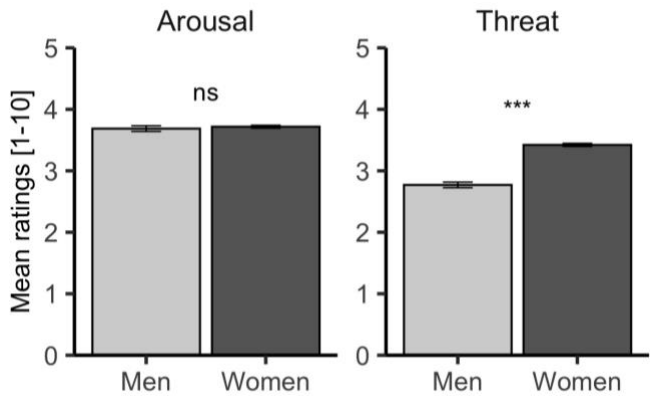

Figure 1. Arousal and threat ratings by gender. The bar graphs display the average ratings for arousal and threat across all 90 videos, divided by male and female participants. Arousal and threat are scored on a scale from 1 to 10, with higher values indicating stronger responses. Statistical significance of the contrast between men and women is indicated by p-values: ns = non-significant; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, all Bonferroni corrected.

Mann-Whitney U tests showed there was no significant difference between women and men in arousal ratings (mean Women = 3.72 vs. mean Men = 3.69, W = 17900000, $p = 0.736$). For threat ratings, women gave significantly higher ratings than men (mean Women = 3.42 vs.

mean Men = 2.77, W = 15710000, p < 0.0001), showing that women perceived the videos as more threatening overall.

3.2.2. Gender-specific differences in video categorisation according to hedonic qualities

For the categorisation of the hedonics of the 90 videos there were also some gender differences (Fig. 2).

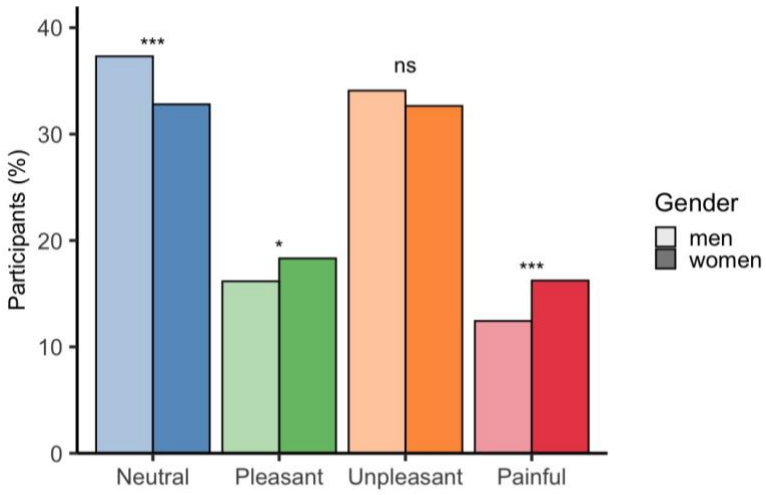

Figure 2. Hedonic categorisation of touch videos by gender. Categorisation of the hedonics of the full 90 video set separated by gender (light bars: men; dark bars: women). Statistical significance of the contrast between men and women is indicated by p-values: ns = non-significant; *** p < 0.001, ** p < 0.01, * p < 0.05, all Bonferroni corrected.

A chi-squared test showed there was a significant association between gender and category ($X^2(3) = 42.475$, $p < 0.0001$). We conducted subsequent separate two-sample tests for equality of proportions to compare the categorisation of videos between men and women across four categories. This suggests a difference between men and women categorising the videos as 'neutral' ($X^2 = 20.68$, $p < 0.001$), 'pleasant' ($X^2 = 7.07$, $p = 0.031$), and 'painful' ($X^2 = 24.87$, $p < 0.001$), but not unpleasant ($X^2 = 2.09$, $p = 0.593$). Men were more likely than women to categorise a video as 'neutral' whereas women showed a stronger tendency than men to categorise videos as 'pleasant' or 'painful'.

3.2.3. Hedonic qualities, threat & arousal

In the following section, we delineate findings concerning the strength of hedonic qualities, including a focus on gender-related differences. Additionally, we explore the relationship between these hedonic categories and levels of arousal and threat. To ensure clear classification of the videos into their respective hedonic types, our analysis from this point forward is confined to the subset of 40 videos selected based on the highest percentage of participant agreement (across gender) in categorising the videos as neutral, pleasant, unpleasant, and painful.

3.2.4. The strength of hedonic qualities

Where participants categorised the touch as valanced (i.e., not those categorised as neutral), we considered participant ratings for the question: “How pleasant/unpleasant/painful was the touch?”. We first look across the whole sample (Fig. 3) and then divide by gender (Fig. 4), given the results above suggest some differences in categorisation between men and women.

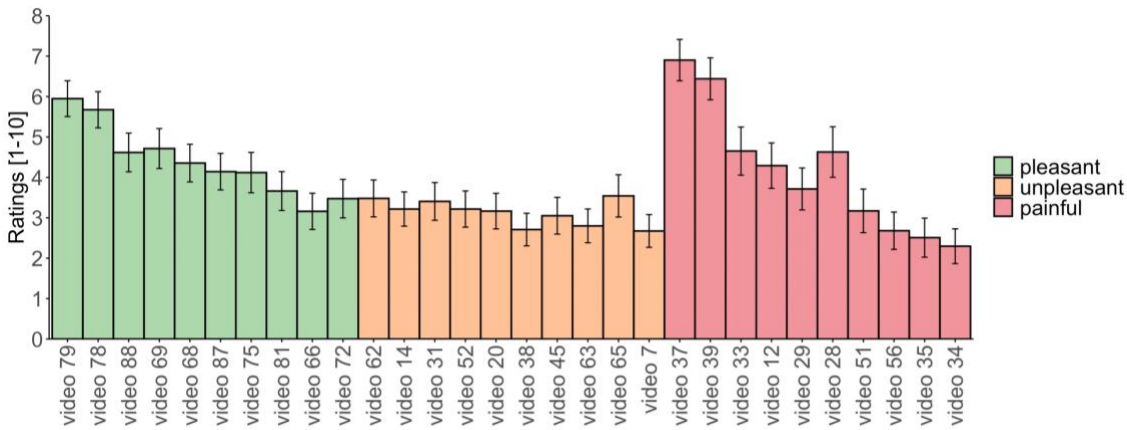

Figure 3. Hedonic quality strength ratings. Each bar graph displays the average hedonic strength ratings for the entire sample, accompanied by 95% CI error bars. Strength ratings are scored on a scale from 1 to 10, with higher values indicating stronger responses. The top 10 most distinctly categorised videos are organised from left to right, with pleasant (green), unpleasant (orange), and painful (red). For a detailed description of each touch video, see Supplementary Materials Table 1. Gender-specific plots can also be found in the Supplementary Materials.

A Kruskal-Wallis rank sum test for the whole sample showed a difference across the categories ($X^2(2) = 134.71$, $p < 0.0001$). Subsequent Mann-Whitney U tests revealed a notable disparity between the 'pleasant' and 'unpleasant' hedonic ratings (mean pleasant = 4.38 vs. mean unpleasant = 3.10, W = 1830100, $p < 0.0001$). There was also a significant difference in ratings between the 'painful' and 'unpleasant' categories (mean painful = 4.20 vs. mean unpleasant = 3.10, W = 1261900, $p < 0.0001$). However, the 'pleasant' and 'painful' categories showed no significant distinction (mean pleasant = 4.38 vs. mean painful = 4.20, W = 1524000, $p = 0.338$). These results suggest that the top 10 videos from both the 'pleasant' and 'painful' categories elicited comparable hedonic responses, whereas videos from the 'unpleasant' category were perceived as having a less intense negative hedonic quality.

3.2.5. Gender differences for painful touch

We conducted an analysis to assess potential gender-based differences in the average hedonic strength ratings across the three categories (Fig. 4).

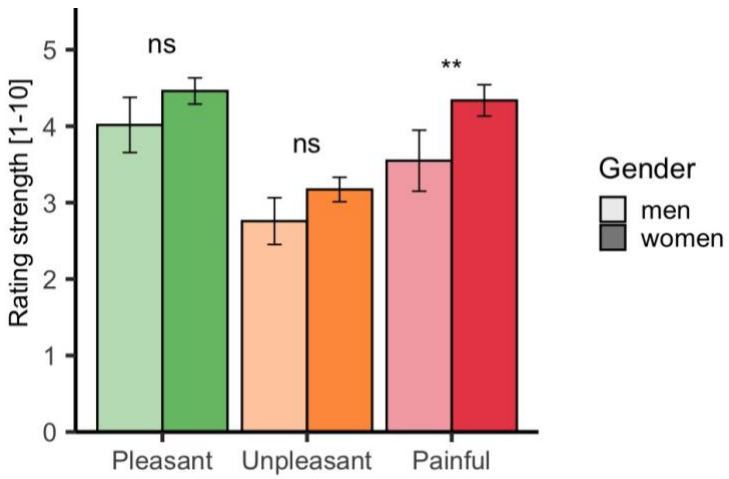

Figure 4. Hedonic rating strength for each category divided by gender. Each bar graphs shows the average hedonic strength ratings for a subset of videos (top 10 per category), differentiated by gender (light bars for men and dark bars for women) with 95% CI error bars. Videos are grouped as pleasant (green), unpleasant (orange), and painful (red). Statistical significance of the contrast between men and women is indicated by p-values: ns = non-significant; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, all Bonferroni corrected.

A Kruskal-Wallis rank sum test showed that there is a statistically significant interaction between category and gender ($X^2(5) = 154.45$, $p < 0.0001$). Subsequent Mann-Whitney U tests indicated that only the 'painful' category showed significant gender differences, with women rating these videos higher than men ($W = 245640$, $p < 0.01$) These findings indicate a gender-specific sensitivity to painful stimuli, but do not suggest any gender-related variations for ratings in the pleasant and unpleasant categories.

3.2.6. Clear association between hedonic qualities and perceived arousal and threat

We explored the interplay between the hedonic qualities of touch depicted in videos and participant perceptions of arousal and threat (Fig. 5). This confirmed a significant association between these variables, demonstrating variations across hedonic categories.

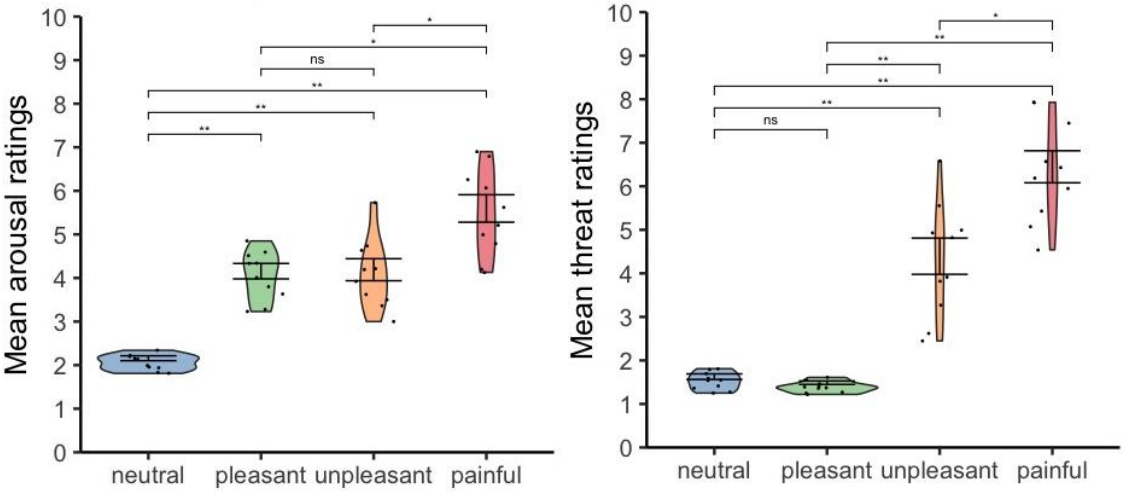

Figure 5. Average ratings for arousal and threat across the top 10 videos in each category. On the left, violin plots depict the distribution of arousal ratings for the touch shown in the videos. On the right, the distribution for threat ratings is presented. The scatter plot within each violin display represents the mean arousal and threat scores for each of the 10 individual videos across the four categories. Statistical significance is indicated by p-values: ns = non-significant; *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, all Bonferroni corrected.

Kruskal-Wallis rank sum test showed that both arousal ($X^2(3) = 28.292$, $p < 0.0001$) and threat ($X^2(3) = 31.869$, $p < 0.0001$) were different across the categories. Subsequent Mann-Whitney U showed that painful videos were rated as the most arousing on average (mean = 5.50), followed by unpleasant videos (mean = 4.09), pleasant videos (mean = 4.06), and neutral

videos (mean = 2.06). Only the pleasant and unpleasant category were not statistically different. Additionally, painful videos received the highest threat rating on average (mean = 6.35), followed by unpleasant videos (mean = 4.29), neutral videos (mean = 1.53), and pleasant videos (mean = 1.39). Only the neutral and pleasant category were not statistically different. These findings are in line with the typically perceived relationship between valence and arousal, where very positive and negative stimuli evoke the strongest arousal and neutral images the least strong arousal (e.g., Bradley and Lang, 1994; Masson and Op de Beeck, 2018).

3.3. Video validation & accessibility

A comprehensive summary of the video data for the 90 videos, which includes a description of the touch, the YouTube link, and validation results, can be accessed on OSF (https://osf.io/jvkqa/) as an R Markdown file (see Fig. 6 for a preview).

Copy CSV Excel PDF Search:

YouTube video	Description of touch	Link to YouTube	Neutral % pts	Pleasant % pts	Unpleasant % pts	Painful % pts	Mean Threat	Mean Arousal
1	repeated touch with finger (fast)	mvZ-J93FhB8	70	8	22	0	1.8	2.0
2	single touch with finger	DGrxS_mWwDg	85	8	8	0	1.4	1.7
3	single push with finger	X7TYV3JlDGg	61	4	34	1	2.2	2.4
4	repeated tap with finger (slow)	4LCl3lFTI5A	76	12	11	0	1.5	1.9
5	repeated tap with finger moving from top to bottom	VGk7DAEMUqY	61	31	6	1	1.4	2.5
6	single touch with flat hand	hZUtrE-RKm4	76	14	10	0	1.6	2.2
7	repeated touch with hairbrush	jj9cYYLwz3c	34	9	54	4	2.6	2.7
8	single touch with hairbrush	lK5uo4vyk4o	51	8	40	1	2.0	2.3
9	single touch with hammer	UihTe8un04c	34	2	50	14	5.1	3.8

Figure 6. Overview of the videos along with their validation data. Preview of the R Markdown file which can be found on OSF (https://osf.io/jvkqa/). The 90 videos in this table can be arranged based on the highest or lowest values for each index, such as hedonic quality (percentage of participants), threat, and arousal (1 = not at all, 10 = extremely). Additionally, it allows you to search for keywords associated with the touch. These functions can be helpful if you are selecting videos based on certain characteristics. Further, there is the option to download the video list as a CSV, Excel, or PDF file.

4. Discussion

Our aim was to create a systematically defined database of videos depicting diverse types of touch to a hand in a controlled environment. The use of multiple methods to assess inter-subject agreement shows that the videos were interpreted with high consistency across participants. Our findings suggest that the videos were successful in eliciting varying levels of arousal, threat, and hedonic qualities. A subset of 40 videos was reliably categorised into neutral, pleasant, unpleasant, and painful touch, providing a wide range of stimuli for future research. The validation of these videos through participant ratings showed significant differences in perceived arousal and threat levels across the different categories of touch. Videos depicting painful touch were associated with the highest arousal and threat levels, and our neutral touch videos were an effective baseline, eliciting low arousal and threat levels. This highlights the benefit of having a neutral category in the database, allowing researchers to compare responses to emotionally charged touch (pleasant, unpleasant, painful) with responses to emotionally neutral touch.

This database will allow further investigation of vicarious touch perception, a phenomenon where individuals report feeling tactile sensations when they witness others being touched. The inclusion of a wide array of visual touch stimuli enables nuanced explorations into how observing different types of touch affects individuals who are prone to experiencing vicarious tactile sensations. For instance, researchers can investigate whether observers report heightened vicarious sensations when exposed to painful touch as opposed to pleasurable touch, building on prior research findings (Holle et al., 2011; Ward et al., 2018; Ward and Li, 2022). Further, researchers can examine how the intensity levels of these vicarious experiences relate to the arousal and perceived threat associated with the observed

touch. The inclusion of validation data for each video, encompassing hedonic qualities, perceived arousal levels, and threat assessments, equips researchers with a tool for delving deeper into these inquiries.

Our study reveals several gender-related differences in the categorisation and perception of videos depicting tactile interactions. For instance, women exhibited a higher likelihood of classifying videos as 'painful' or 'pleasant' compared to men, while men tended to categorise videos as 'neutral' relatively more often. Further, women assigned higher valence intensity ratings to videos in the painful category, implying a heightened perception of pain in these videos among women. Additionally, women provided higher threat ratings than men when considering the entire database of 90 videos, although their arousal ratings did not differ significantly. These findings align with prior research indicating that women tend to display stronger vicarious responses to observed *pain* (Grice-Jackson et al., 2017; Singer et al., 2006a; Ward and Li, 2022; Yang et al., 2009), with a lack of such a bias for reported (neutral) vicarious touch (Ward et al., 2018; Ward and Li, 2022). Future research could use this video database to delve into the relationship between perceived pain or touch in others and mirrored pain or touch in the observer (see for example Ioumpa et al., 2019), with a specific focus on examining potential gender differences. Nevertheless, it is crucial to exercise caution when interpreting these gender-related findings. For example, women generally self-report higher levels of empathy than men (Baez et al., 2017; Baron-Cohen and Wheelwright, 2004), a tendency that may be influenced by gender-role stereotypes (Baez et al., 2017; Eisenberg and Lennon, 1983; Löffler and Greitemeyer, 2023; Simner et al., 2006). Such stereotypes could potentially also impact video assessments and introduce bias into the data. Overall, our findings underscore the importance of considering gender as a factor when evaluating perceptions and responses to visually perceived touch and pain stimuli.

One strength of our study lies in the extensive sample size used for video validation, which enhances the robustness of the validation data. Moreover, our participant pool showed a rich cultural diversity, with representation from various backgrounds, including Asian (20%), Australian (19%), Chinese (5%), Indian (3%), Lebanese (2%), European (2%), and Filipino

(2%). It is worth acknowledging, however, that our participants consisted primarily of undergraduate students, which might somewhat restrict the generalisability of our validation data. Additionally, there was a higher proportion of women than men in our sample, although the number of male participants remained substantial.

5. Conclusion

In conclusion, the development of this comprehensive video database, featuring a diverse range of touch stimuli along with its accompanying validation data, serves as a resource for advancing research in the domain of visual touch perception. The videos can be used in a variety of experimental paradigms, and the validation data provides a useful reference for researchers to select videos based on specific criteria. Touch is a complex sense that can convey a multitude of sensations and emotions. By providing this resource, we aim to facilitate future investigations in the field, thereby enriching our understanding of the nuanced ways in which different forms of touch are visually perceived and processed.

6. References

Baez, S., Flichtentrei, D., Prats, M., Mastandueno, R., García, A.M., Cetkovich, M., Ibáñez, A., 2017. Men, women… who cares? A population-based study on sex differences and gender roles in empathy and moral cognition. PLoS One 12, e0179336.

Baron-Cohen, S., Wheelwright, S., 2004. The empathy quotient: An investigation of adults with asperger syndrome or high functioning autism, and normal sex differences. J Autism Dev Disord 34, 163–175. https://doi.org/10.1023/B:JADD.0000022607.19833.00/METRICS

Blakemore, S.-J., Bristow, D., Bird, G., Frith, C., Ward, J., 2005. Somatosensory activations during the observation of touch and a case of vision–touch synaesthesia. Brain 128, 1571–1583. https://doi.org/10.1093/brain/awh500

Botvinick, M., Cohen, J., 1998. Rubber hands “feel” touch that eyes see. Nature 391, 756.

Bradley, M.M., Lang, P.J., 1994. Measuring emotion: the self-assessment manikin and the semantic differential. J Behav Ther Exp Psychiatry 25, 49–59.

Bufalari, I., Ionta, S., 2013. The social and personality neuroscience of empathy for pain and touch. Front Hum Neurosci 0, 393. https://doi.org/10.3389/fnhum.2013.00393

Davies, A.M.A., White, R.C., 2013. A sensational illusion: vision-touch synaesthesia and the rubber hand paradigm. Cortex 49, 806–818.

Ebisch, S.J.H., Ferri, F., Salone, A., Perrucci, M.G., D’Amico, L., Ferro, F.M., Romani, G.L., Gallese, V., 2011. Differential involvement of somatosensory and interoceptive cortices during the observation of affective touch. J Cogn Neurosci 23, 1808–1822.

Eisenberg, N., Lennon, R., 1983. Sex differences in empathy and related capacities. Psychol Bull 94, 100–131. https://doi.org/10.1037/0033-2909.94.1.100

Gillmeister, H., Bowling, N., Rigato, S., Banissy, M.J., 2017. Inter-individual differences in vicarious tactile perception: a view across the lifespan in typical and atypical populations. Multisens Res 30, 485–508. https://doi.org/10.1163/22134808-00002543

Giummarra, M.J., Fitzgibbon, B.M., Georgiou-Karistianis, N., Beukelman, M., Verdejo-Garcia, A., Blumberg, Z., Chou, M., Gibson, S.J., 2015. Affective, sensory and empathic sharing of another's pain: The Empathy for Pain Scale. European Journal of Pain 19, 807–816.

Giummarra, M.J., Tracy, L.M., Young, K.A., Fitzgibbon, B.M., 2016. The social side of pain: What does it mean to feel another's pain? Meanings of Pain 355–373. https://doi.org/10.1007/978-3-319-49022-9_21/FIGURES/2

Grice-Jackson, T., Critchley, H.D., Banissy, M.J., Ward, J., 2017. Common and distinct neural mechanisms associated with the conscious experience of vicarious pain. Cortex 94, 152–163. https://doi.org/10.1016/j.cortex.2017.06.015

Holle, H., Banissy, M., Wright, T., Bowling, N., Ward, J., 2011. "That's not a real body": Identifying stimulus qualities that modulate synaesthetic experiences of touch. Conscious Cogn 20, 720–726.

Holle, H., Warne, K., Seth, A.K., Critchley, H.D., Ward, J., 2012. Neural basis of contagious itch and why some people are more prone to it. Proc Natl Acad Sci U S A 109, 19816–19821. https://doi.org/10.1073/PNAS.1216160109/SUPPL_FILE/PNAS.201216160SI.PDF

Ioumpa, K., Graham, S.A., Clausner, T., Fisher, S.E., Van Lier, R., Van Leeuwen, T.M., 2019. Enhanced self-reported affect and prosocial behaviour without differential physiological responses in mirror-sensory synaesthesia. https://doi.org/10.1098/rstb.2019.0395

Koo, T.K., Li, M.Y., 2016. A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. J Chiropr Med 15, 155. https://doi.org/10.1016/J.JCM.2016.02.012

Kuehn, E., Trampel, R., Mueller, K., Turner, R., Schütz-Bosbach, S., 2013. Judging roughness by sight—A 7-tesla fMRI study on responsivity of the primary somatosensory cortex during observed touch of self and others. Hum Brain Mapp 34, 1882–1895. https://doi.org/10.1002/hbm.22031

Lamm, C., Silani, G., Singer, T., 2015. Distinct neural networks underlying empathy for pleasant and unpleasant touch. Cortex 70, 79–89. https://doi.org/10.1016/J.CORTEX.2015.01.021

Landis, J.R., Koch, G.G., 1977. An Application of Hierarchical Kappa-type Statistics in the Assessment of Majority Agreement among Multiple Observers. Biometrics 33, 363. https://doi.org/10.2307/2529786

Lee Masson, H., Van De Plas, S., Daniels, N., Op de Beeck, H., 2018. The multidimensional representational space of observed socio-affective touch experiences. Neuroimage 175, 297–314. https://doi.org/10.1016/j.neuroimage.2018.04.007

Li, M., Hao, L., Ren, Z., Qiu, J., Ward, J., 2022. Vicarious experiences of touch (mirror touch) in a Chinese sample: Cross-cultural and individual differences. PLoS One 17, e0266246. https://doi.org/10.1371/JOURNAL.PONE.0266246

Löffler, C.S., Greitemeyer, T., 2023. Are women the more empathetic gender? The effects of gender role expectations. Current Psychology 42, 220–231. https://doi.org/10.1007/S12144-020-01260-8/FIGURES/2

Luangrath, A.W., Peck, J., Hedgcock, W., Xu, Y., 2022. Observing Product Touch: The Vicarious Haptic Effect in Digital Marketing and Virtual Reality. https://doi.org/10.1177/00222437211059540

Makary, M.M., Lee, J., Lee, E., Eun, S., Kim, J., Jahng, G.-H., Kim, K., Youn, Y.-S., Lee, J.-H., Park, K., 2018. Phantom Acupuncture Induces Placebo Credibility and Vicarious Sensations: A Parallel fMRI Study of Low Back Pain Patients OPEN. Scientific RePORtS | 8, 930. https://doi.org/10.1038/s41598-017-18870-1

Masson, H.L., Op de Beeck, H.P., 2018. Socio-affective touch expression database. PLoS One 13, e0190921. https://doi.org/10.1371/JOURNAL.PONE.0190921

Medina, J., DePasquale, C., 2017. Influence of the body schema on mirror-touch synesthesia. Cortex 88, 53–65.

Morrison, I., Björnsdotter, M., Olausson, H., 2011. Vicarious Responses to Social Touch in Posterior Insular Cortex Are Tuned to Pleasant Caressing Speeds. Journal of Neuroscience 31, 9554–9562. https://doi.org/10.1523/JNEUROSCI.0397-11.2011

Osborn, J., Derbyshire, S.W.G., 2010. Pain sensation evoked by observing injury in others. Pain 148, 268–274.

Peled-Avron, L., Levy-Gigi, E., Richter-Levin, G., Korem, N., Shamay-Tsoory, S.G., 2016. The role of empathy in the neural responses to observed human social touch. Cogn Affect Behav Neurosci 16, 802–813. https://doi.org/10.3758/S13415-016-0432-5

R Core Team, 2023. R: A Language and Environment for Statistical Computing. URL: https://www.R-project.org/.

Ramachandran, V.S., Altschuler, E.L., 2009. The use of visual feedback, in particular mirror visual feedback, in restoring brain function. Brain 132, 1693–1710. https://doi.org/10.1093/BRAIN/AWP135

Ramachandran, V.S., Rodgers-Ramachandran, D., 1996. Synaesthesia in phantom limbs induced with mirrors. Proc R Soc Lond B Biol Sci 263, 377–386. https://doi.org/10.1098/RSPB.1996.0058

Revelle, W., Revelle, M.W., 2015. Package ‘psych.’ The comprehensive R archive network 337–338.

Riva, F., Tschernegg, M., Chiesa, P.A., Wagner, I.C., Kronbichler, M., Lamm, C., Silani, G., 2018. Age-related differences in the neural correlates of empathy for pleasant and unpleasant touch in a female sample. Neurobiol Aging 65, 7–17. https://doi.org/10.1016/J.NEUROBIOLAGING.2017.12.028

Schütz-Bosbach, S., Tausche, P., Weiss, C., 2009. Roughness perception during the rubber hand illusion. Brain Cogn 70, 136–144. https://doi.org/10.1016/J.BANDC.2009.01.006

Simner, J., Mulvenna, C., Sagiv, N., Tsakanikos, E., Witherby, S.A., Fraser, C., Scott, K., Ward, J., 2006. Synaesthesia: The prevalence of atypical cross-modal experiences. Perception 35, 1024–1033.

Singer, T., Seymour, B., O'Doherty, J.P., Stephan, K.E., Dolan, R.J., Frith, C.D., 2006. Empathic neural responses are modulated by the perceived fairness of others. Nature 2005 439:7075 439, 466–469. https://doi.org/10.1038/nature04271

Smit, S., Moerel, D., Zopf, R., Rich, A.N., 2022. Vicarious touch: overlapping neural patterns between seeing and feeling touch. bioRxiv. https://doi.org/https://doi.org/10.1101/2022.06.21.497107

Taneja, P., Olausson, H., Trulsson, M., Svensson, P., Baad-Hansen, L., 2021. Defining pleasant touch stimuli: a systematic review and meta-analysis. Psychol Res 85, 20–35. https://doi.org/10.1007/S00426-019-01253-8/TABLES/4

Vandenbroucke, S., Crombez, G., Van Ryckeghem, D., Brass, M., Van Damme, S., Goubert, L., 2013. Vicarious pain while observing another in pain: an experimental approach. Front Hum Neurosci 7, 265.

Ward, J., Li, M., 2022. The Vicarious Experiences Questionnaire(s): Online Tools for Measuring Mirror-Touch and Vicarious Pain. Preprint. https://doi.org/10.31234/OSF.IO/AE6X4

Ward, J., Schnakenberg, P., Banissy, M.J., 2018. The relationship between mirror-touch synaesthesia and empathy: New evidence and a new screening tool. Cogn Neuropsychol 35, 314–332. https://doi.org/https://doi.org/10.1080/02643294.2018.1457017

Yang, C.Y., Decety, J., Lee, S., Chen, C., Cheng, Y., 2009. Gender differences in the mu rhythm during empathy for pain: An electroencephalographic study. Brain Res 1251, 176–184. https://doi.org/10.1016/j.brainres.2008.11.062