REVIEW

Defining pleasant touch stimuli: a systematic review and meta-analysis

Pankaj Taneja[1,2] · Håkan Olausson[3] · Mats Trulsson[2,4] · Peter Svensson[1,2,5] · Lene Baad-Hansen[1,2]

Abstract

Pleasantness is generally overlooked when investigating tactile functions. Addition of a pleasant stimulus could allow for a more complete characterisation of somatosensory function. The aims of this review were to systematically assess the methodologies used to elicit a pleasant sensation, measured via psychophysical techniques, and to perform a meta-analysis to measure the effect of brush stroking velocity on touch pleasantness. Eighteen studies were included in the systematic review, with five studies included in the meta-analysis. The review found that factors such as texture, velocity, force, and the duration of continuous stroking influence tactile evoked pleasantness. Specifically, using a soft material and stroking at a velocity of 3 cm/s with light force is generally considered as particularly pleasant. The meta-analysis showed that a brush stroking velocity of 30 cm/s was rated as less pleasant than 3 cm/s, on the forearm. The present study collates the factors that are most likely to provide a stimulus to elicit a pleasant sensation. The results should be important for studies requiring a well-defined pleasant stimulus including neurosensory assessment protocols, allowing for a more complete multimodality assessment of somatosensory function.

Introduction

The perception of a stimulus is complex and constitutes a multidimensional construct (Taneja et al., 2019). What may elicit a perception to one individual may not necessarily evoke the same to another. Therefore, it can be challenging when trying to establish a standardised stimulus to produce a targeted perception, a factor important when investigating conditions such as nerve damage (Rolke et al., 2006). It is not surprising that there are a number of somatosensory investigations that exist to assess such a problem (Pigg,

Baad-Hansen, Svensson, Drangsholt, & List, 2010; Rolke et al., 2006). However, within these investigatory protocols the main focus tends to be on altered mechanical and thermal sensitivity, paying particular interest in stimuli considered as painful or unpleasant (Svensson et al., 2011; Poort, van Neck, & van der Wal, 2009; Rolke et al., 2006). A percept that is regularly overlooked is that of pleasantness.

A recent discovery is that the coding of a pleasant sensation seems to be via a neuronal network known as the C-Tactile (CT) afferent system (Löken, Wessberg, Morrison, McGlone, & Olausson, 2009). These are unmyelinated, low-threshold afferents found in the hairy skin of humans (Vallbo, Olausson, Wessberg, & Norrsell, 1993, 1999). Since their discovery there has been a plethora of research into the contribution of the CT afferents to social touch and pleasant sensations (Walker, Trotter, Woods & McGlone, 2017; Morrison, Björnsdotter & Olausson, 2011).

It has been established that the optimal activation of CT afferents is a slowly stroking stimulation (1–10 cm/s) (Nordin, 1990). This gentle stroking is considered as pleasant and often seen in affiliative interactions (Olausson, Wessberg, McGlone, & Vallbo, 2010; Gallace & Spence, 2010). Furthermore, the psychophysical measure of pleasantness correlates with the firing of CT afferents, linking this system to such a percept (Löken et al., 2009).

✉ Pankaj Taneja
  pantaneja@dent.au.dk

1 Section of Orofacial Pain and Jaw Function, Department of Dentistry and Oral Health, Aarhus University, Aarhus, Denmark

2 Scandinavian Centre for Orofacial Neurosciences (SCON), Aarhus, Denmark

3 Department of Clinical and Experimental Medicine, Linköping University, Linköping, Sweden

4 Department of Dental Medicine, Karolinska Institute, Solna, Sweden

5 Department of Orofacial Pain and Jaw Function, Faculty of Odontology, Malmö University, Malmö, Sweden

However, other peripheral afferents will be contributing, as a stroking touch can be pleasant in regions lacking CT afferents, i.e. glabrous skin (Löken, Evert, & Wessberg, 2011; McGlone et al., 2012).

The slow conduction of CT afferents renders them suboptimal for sensory discrimination (Morrison, Loken & Olausson, 2010). Nonetheless, a slow conduction velocity does not preclude a role in affective touch (McGlone, Wessberg, & Olausson, 2014). Consistently, functional magnetic resonance imaging studies identify CT signals to be processed in emotion-related areas of the brain, such as the insular cortex, and less in the primary somatosensory cortex (Olausson et al., 2002, 2008; Morrison et al., 2011) which is known to play a key role in discriminative touch (Olausson et al., 2008; McGlone et al., 2014; Case, Laubacher et al., 2016).

It is clear that stimuli optimal for activation of CT afferents is considered as pleasant, but the stimulus' properties associated with pleasant sensations are poorly defined. Hence, the aim of this review was to conduct a systematic review of the stimuli and methods used to generate a pleasant sensation, with particular focus on CT-psychophysical studies. This would aid in identifying the stimulus factors that are most relevant in producing a pleasant sensation. An additional aim was to undertake a meta-analysis to provide a pooled estimate of the difference in pleasantness ratings for CT-optimal (3 cm/s) and CT-suboptimal (30 cm/s) stroking touch. (Löken et al., 2009; Vallbo et al., 1993, Vallbo, Olausson, & Wessberg, 1999). It is anticipated that the findings can guide studies that require a pleasant tactile sensation, or in neurosensory assessment protocols, allowing for a more complete examination of somatosensory function.

Materials and methods

The systematic review was performed in accordance with the PRISMA guidelines (Moher, Liberati, Tetzlaff & Altman, 2009) and registered with the PROSPERO database (Number: CRD42017058867).

Search strategy

A literature search of articles published in English from 1974 to June 2018 was conducted through the Cochrane Library, Embase, Medline, PubMed, Scopus and Web of Science, as well as a web search (via Google Scholar). The search was performed on the same day and had been tailored to the relevant database by author PT and an experienced librarian. English search terms were used, and there were no limits placed (Table 1).

Eligibility assessment

The eligibility criteria consisted of articles of any study type, involving healthy adults (≥ 18 years), and thus excluding studies with conditions/diseases that may affect the perception of a pleasant sensation. If studies consisted of patients and a control (healthy) group, inclusion of the control group was only performed, if independent results on pleasantness were present. In addition, if a study consisted of multiple experiments, then only the relevant experiment/part was included. Eligibility also required the use of a defined stimulus (with a description) and a numerical rating for pleasantness. Exclusion of articles consisted of those that were attempting to define pleasantness by descriptive terms only, or by an active form of assessment (perceived by the participant touching a material, rather than being touched). Studies assessing pleasantness via a combination of touch with other stimuli (e.g. visual), or in which the experimental set-up may potentially influence the perceived pleasantness (e.g. within an MRI scanner, etc.), were excluded. To be included in the meta-analysis, the above-mentioned criteria had to be met and individual studies needed to compare stroke velocities, with similar experimental set-ups.

Data collection and analysis

Initially, two of the authors (PT and LBH) independently screened over half of the titles and abstracts based on the inclusion criteria listed above. The basic percentage inter-rater agreement was assessed to be 94%, and Cohen's Kappa statistic was 0.87, indicating a near perfect agreement.

Table 1 Search strategy used for identifying titles in PubMed

(((((((((((((((Touch*[tiab]) OR Therapeutic touch*[tiab]) OR Touch sensation*[tiab]) OR Touch perception*[tiab]) OR Physical stimulation*[tiab]) OR Tactile*[tiab]) OR Tactile perception*[tiab]) OR Tactile sensitivity*[tiab]) OR Tactile stimulation*[tiab]) OR ("Therapeutic Touch"[Mesh] OR "Touch"[Mesh] OR "Touch Perception"[Mesh])) OR ("Agnosia"[Mesh] OR "Hypesthesia"[Mesh] OR "Hyperesthesia"[Mesh])) OR "Physical Stimulation"[Mesh]) OR Tactile sensation*[tiab])) AND ((((((hedonic*[tiab]) OR affective*[tiab]) OR Pleasant*[tiab]) OR Pleasure*[tiab]) OR Tactile pleasant*[tiab]) OR ("Pleasure"[Mesh] OR "Pleasure-Pain Principle"[Mesh])) OR "Happiness"[Mesh])) AND (((((((((Afferent*[tiab]) OR C tactile afferent*[tiab]) OR CT-afferent*[tiab]) OR unmyelinated tactile afferent*[tiab]) OR tactile afferent*[tiab]) OR "Afferent Pathways"[Mesh]) OR "Synaptic Transmission"[Mesh]) OR Synaptic Transmission*[tiab])

Hence, PT screened the remaining titles and abstracts with any concerns discussed with LBH. Full text articles were obtained for those that appeared to address the review topic, or for those abstracts that were not clear. Finally, the reference lists of all included articles were searched to identify any studies that may have been missed. Any disagreements were resolved by discussion. PT performed data extraction, and relevant items included authors, year of publication, number of participants, gender, blinding and randomisation, region assessed, type of instrument used for stimulus delivery, velocity and force of the stimuli used, the interstimulus interval, and the rating scale used.

For the meta-analysis, effect sizes in the form of standardised mean differences were computed. This method allowed for the adjustment in different pleasantness rating scales utilised across studies and allow a direct comparison of the velocities (Morris & DeShon, 2002). Hence, descriptive data (mean, standard deviations (SD) and sample sizes) were collected. When studies presented with data that allowed more than one effect size to be computed (multiplicity), the one that allowed the maximum comparability between studies was selected, for example, based on details of the experimental set-up, duration of stimulus delivery, etc. Both reviewers (PT and LBH) agreed on this approach as selection was not based on any magnitude or direction (López-López, Page, Lipsey & Higgins, 2018). In the instance that effect sizes were equally suitable, and if they were not significantly different from one another, the average was taken (López-López et al., 2018). In this way, each study would therefore contribute only 1 effect size.

Furthermore, as studies were likely to be repeated measures (investigating multiple velocities on the same participants), a correlation coefficient between the investigated velocities was needed for variance computation (Morris & DeShon, 2002; Dunlap, Cortina, Vaslow & Burke, 1996; Borenstein, Hedges, Higgins & Rothstein, 2011). If all but the correlation coefficient was available, the decision was to use the average correlation from those studies that stated or had sufficient data available for it to be computed (Morris & DeShon, 2002; Dunlap et al., 1996; Borenstein et al., 2011). Corresponding authors were contacted by email if any data were missing. If no response, the authors were contacted again after 2 weeks with an average time of 8 weeks to receive the information.

Quality assessment in individual studies

Both reviewers (LBH and PT) independently performed the quality assessment. An established quality assessment checklist was not utilised as it was the experimental technique for eliciting a pleasant sensation that was being evaluated, and not the overall clinical study. Hence, it was decided that criteria based on common checklists, as well as those factors considered as important in the investigation of CT afferents, should be devised. The assessment was at the methodological level and consisted of a 14-item checklist generally categorised into: participants selected, randomisation and blinding within the experimental set-up, consideration of confounding factors and statistical analysis used ("Appendix" section). All included studies were assessed against the checklist, and quality was categorised as a percentage with low ($\leq 33\%$), moderate (≥ 34 and $\leq 66\%$) and high ($\geq 67\%$) methodological quality.

Data synthesis and analysis

The effect sizes for each study were calculated in the form of standardised mean differences. This was achieved using the Cohen's d via the pretest–posttest design method, using the pooled within-group SD (Borenstein et al., 2011). The calculations were done in a way that a negative effect size represented increased pleasantness ratings in the 3 cm/s group relative to 30 cm/s, whereas a positive effect size represented the converse (Faraone, 2008). The magnitude of the effect size was interpreted as small $= 0.2$, medium $= 0.5$ and large $= 0.8$ (Cohen, 1992).

The meta-analytical computations were performed using Stata software (StataCorp. 2017. Stata Statistical Software: Release 15. College Station, TX: StataCorp LLC). A random effects model was fitted as this model has been identified to account for the distribution of an effect over populations (Israel & Richter, 2011). Two tests investigated heterogeneity. The Cochran Q test, in which a $p < 0.10$ (see below for justification to significance level), signified statistically significant heterogeneity between studies, and the I^2 statistic, which identifies the percentage of variation between study estimates as a result of heterogeneity rather than chance (Higgins & Thompson, 2002; Higgins, Thompson, Deeks, & Altman, 2003). An I^2 value of 0%, 50% and 75% indicated low, moderate or high heterogeneity, respectively (Higgins, Thompson, Deeks & Altman, 2003).

Sensitivity analysis was used to investigate the robustness of different factors on the pooled effect estimate. Firstly, the impact on the overall effect estimate by different statistical methods (fixed and random effects model) was undertaken. Secondly, the "leave-one-out" analysis was performed to investigate whether removing any individual study could influence the overall pooled estimate.

A hand and robot stimulation subgroup analysis was planned to assess any influence on effect size heterogeneity with a 0.10 cut-off significance level to mitigate the small number of included studies and, therefore, problems with subgroup tests for heterogeneity (Sedgwick, 2013). Finally, to assess for publication bias, inspection of funnel plots and the Egger's regression test were planned (Peters, Sutton, Jones, Abrams & Rushton, 2006).

Results

Study selection

Following title and abstract screening, 96 full text studies were reviewed (Fig. 1). Seventeen studies met the inclusion criteria, and the reference lists were analysed identifying one further study. In total 18 studies were included for the qualitative synthesis.

From these, 8 studies did not meet the inclusion criteria for the meta-analysis due to the experimental set-up or velocities tested (Bennett, Bolling, Anderson, Pelphrey & Kaiser, 2014; Essick et al., 2010; Etzi, Spence & Gallace, 2014; Gordon et al., 2013; Hua et al., 2008; Huisman,

Frederiks, van Erp & Heylen, 2016; Tsalamlal, Ouarti, Martin & Ammi, 2013, Voos, Pelphrey & Kaiser, 2013). Seven of the remaining ten studies did not provide sufficient data, and therefore corresponding authors were contacted (Ackerley, Carlsson, Wester, Olausson & Backlund Wasling, 2014; Etzi, Carta & Gallace, 2018; Jönsson et al., 2015, 2017; Kass-Iliyya et al., 2016; Luong, Bendas, Etzi, Olausson & Croy, 2017; Löken et al., 2011). Replies with data were received for 2 of the studies (Ackerley et al., 2014; Etzi et al., 2018), but not for the other 5 studies, leading to their exclusion (Jönsson et al., 2015, 2017; Kass-Iliyya et al., 2016; Luong et al., 2017; Löken et al., 2011). Therefore, a total of 5 studies were included in the meta-analysis (Ackerley et al., 2014; Etzi et al., 2018; Pawling, Cannon, McGlone & Walker, 2017;
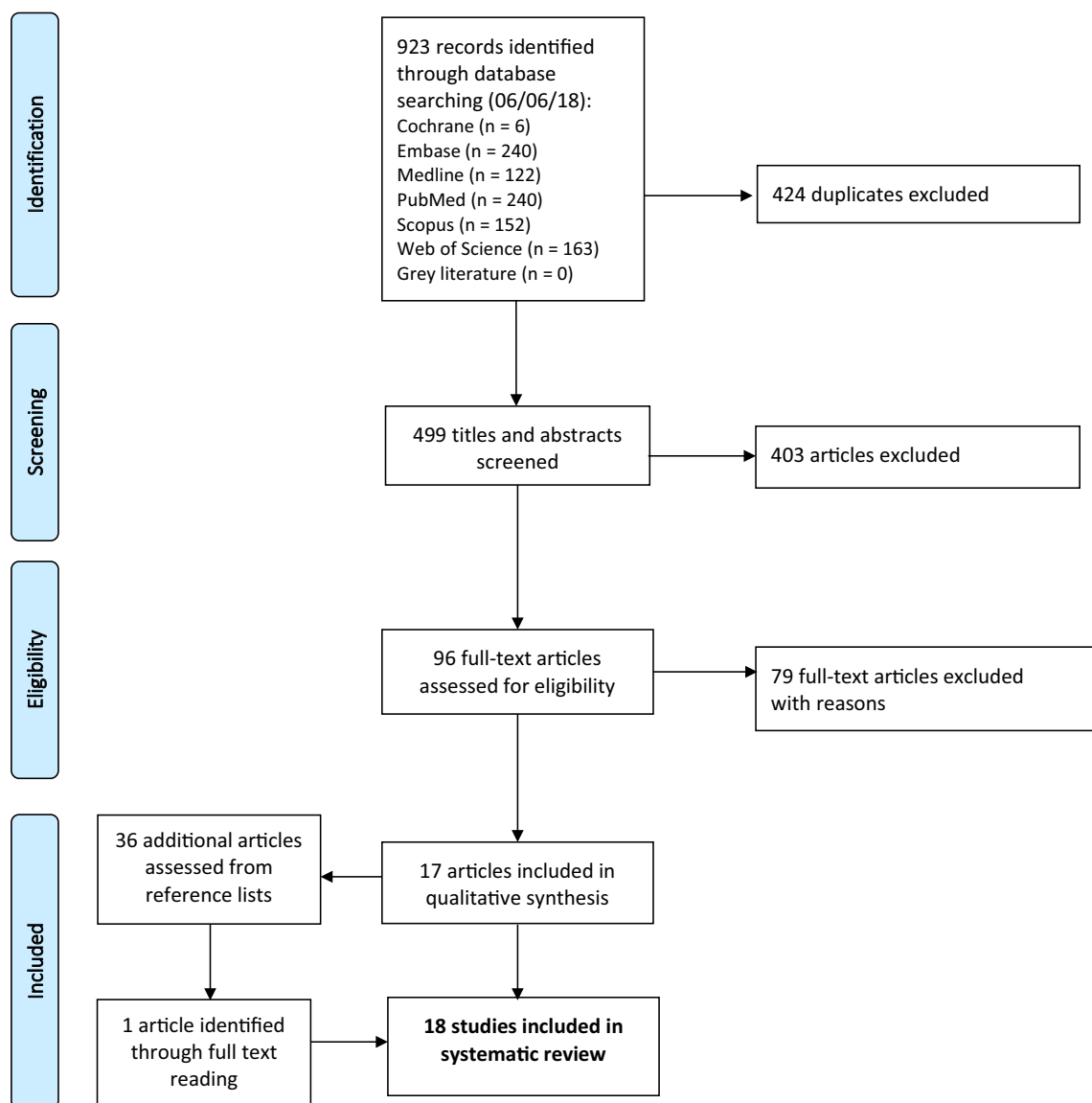

Fig. 1 PRISMA flow chart showing the process of selecting studies

Triscoli, Olausson, Sailer, Ignell & Croy, 2013; Triscoli, Ackerley & Sailer, 2014).

Quality assessment

The quality assessment of the methods undertaken was rated as high in eight complete studies (Ackerley et al., 2014; Essick et al., 2010; Etzi et al., 2018; Jönsson et al., 2015, 2017; Triscoli et al., 2013, 2014; Tsalamlal et al., 2013) and experiments 1 and 2 in Löken et al. (2011). A satisfactory score was achieved by four complete studies (Huisman et al., 2016; Kass-Iliyya et al., 2016; Luong et al., 2017; Pawling et al., 2017), experiment 3 in Löken et al. (2011), and experiment 2 in Etzi et al. (2014). Four studies had their described methodology rated as weak (Bennett et al., 2014, Gordon et al., 2013, Hua et al., 2008, Voos et al., 2013). The quality assessment score was not a factor for inclusion, and therefore no studies were excluded on their quality rating.

Qualitative synthesis (systematic review)

Study characteristics

The participant and study method characteristics were evaluated and are summarised in Tables 2 and 3. From the included studies, seven composed of more than one experiment (Essick et al., 2010; Etzi et al., 2014; Jönsson et al., 2015, 2017; Löken et al., 2011; Luong et al., 2017; Triscoli et al., 2014). Two studies provided data prior to the main experiment that required MRI scanning (Gordon et al., 2013; Voos et al., 2013), and one study obtained data from the healthy controls that were compared with patients with Parkinson's disease (Kass-Iliyya et al., 2016).

The included studies (where k = number of studies) were published between 2014 and 2018 with the number of participants (as defined by the inclusion criteria) ranging from 8 (Löken et al., 2011) to 80 (Jönsson et al., 2017). Six studies had a greater number of female participants (k = 6) (Bennett et al., 2014; Etzi et al., 2014; Gordon et al., 2013; Kass-Iliyya et al., 2016; Pawling et al., 2017; Voos et al., 2013), three studies had more

Table 2 Study design and participant demographics from included studies

References	Study design	Number	Age (years) Mean ± SD	Gender f:m
Ackerley et al. (2014)	Ob	34	25 ± 3	17:17
Bennett et al. (2014)	Ob	22	24.1 ± 3.80 (range)	13:9
Essick et al. (2010)	Ob	Exp 1: 23	Exp 1: males: 26.1, females: 25.2	Exp 1: 12:11
		Exp 2: 20	Exp 2: males: 37.1, female 39.5	Exp 2: 10:10
Etzi et al. (2018)	Ob	30	26 ± 4	15:15
Etzi et al. (2014)	Ob	Exp 2: 12	Exp 2: 22.8	Exp 2: 10:2
Gordon et al. (2013)	Ob	19 for pre-scan behavioural ratings	24.52 ± 3.56 (main experiment)	9:13 (main experiment)
Hua et al. (2008)	Ob	12 (9 for psychophysical investigation)	22 ± 2	6:6
Huisman et al. (2016)	Ob	19	31.1 ± 10.4	4:15
Jönsson et al. (2017)	Ob	Study 1: 58	Study 1: 26.2 ± 6.3	Study 1: 34:24
		Study 2: 80	Study 2: 24.9 ± 4.1	Study 2: 51:29
Jönsson et al. (2015)	Ob	Study 1: 29	Study 1: 23.5 ± 3.2	Group 1: 10:10
		Study 2: 46	Study 2: 26.6 ± 6.8	Group 2: 27:19
Kass-Iliyya et al. (2016)	RCT	27 healthy controls	Control: 63.6	Control: 16:11
Löken et al. (2011)	Ob	Exp 1: 10	21–28 (range)	Exp 1: 5:5
		Exp 2: 10		Exp 2: 5:5
		Exp 3: 8		Exp 3: 4:4
Luong et al. (2017)	Ob	Study 1: 30	Study 1: 22 ± 2.6	Study 1: 17:13
		Study 2: 22	Study 2: 25.9 ± 5.9	Study 2: 11:11
Pawling et al. (2017)	Ob	29	30.4 ± 16.4	18:11
Triscoli et al. (2013)	Ob	31	24.5 ± 2.61	15:16
Triscoli et al. (2014)	Ob	Exp 1: 12	Exp 1: 19–28 (range)	Exp 1: 6:6,
		Exp 2: 17	Exp 2: 19–66 (range)	Exp 2: 11:6
Tsalamlal et al. (2013)	Ob	16	22–53 (range)	6:10
Voos et al. (2013)	Ob	19 (18 for pre-scan)	23.73 ± 2.51	12:7

Ob observational study, *RCT* randomised controlled trial, *Exp* experiment, *SD* standard deviation

Table 3 Characteristics of the methods utilised for pleasantness assessment within included studies

References	Blinding of pleasantness investigation	Randomisation within pleasantness method	Region (s) assessed	Mode of stimulus delivery	Stimulus utilised	Velocities cm/s	Force	Interstimulus interval (seconds)	Rating scale used	Comments
Ackerley et al. (2014)	Glasses with side covers	Pseudorandomised velocities Order of skin site	5 sites: forehead, left lower arm, left palm, left thigh and left shin.	Robot (RTS)	5 cm wide GHB	0.3, 1, 3, 10, 30	0.4 N	10	VAS	From pre-experimental component
Bennett et al. (2014)	NR	NR	Right arm and palm	Human (?)	NR	NR	NR	NR	Likert scale	
Essick et al. (2010)	Curtain between RTS and participant. Ear plugs	Exp 1 and 2: Order of stimuli conditions Exp 2: Body site	Exp 1: Left forearm Exp 2: Forehead, left forearm, middle finger of left hand (palm side), thigh and left calf	Robot (RTS)	Exp 1: Cosmetic brush, velvet fabric, burlap fabric (hessian), strip of rough Velcro. Exp 2: Cupro, cotton interlock, terry towelling, denim	Exp 1: 0.75 and 18 Exp 2: 5 and 20.	Exp 1: 0.19 N and 0.73 N Exp 2: 0.46 N and 0.92 N	NR	VAS	
Etzi et al. (2018)	NR	Pseudorandomised order of stimuli	Dominant forearm	Human (f)	Soft round brush 2 cm diameter	3, 30	Not controlled	1 for random order tapping 800 ms and 80 ms for fixed order tapping	VAS	Stimuli also random order tapping and fixed tapping
Etzi et al. (2014)	Sound proof head phones and blindfolded	NR	Right or left forearms or palms and fingertips	Human (f)	Oasis, satin, tinfoil, sandpaper and abrasive sponge	5	Not controlled	NR	VAS	Exp 2 characteristics
Gordon et al. (2013)	NR	NR	Right forearm and palm	Human (?)	7 cm wide watercolour brush	8 cm/s	NR	12	Likert scale	From pre-scan component

Table 3 (continued)

References	Blinding of pleasantness investigation	Randomisation within pleasantness method	Region (s) assessed	Mode of stimulus delivery	Stimulus utilised	Velocities cm/s	Force	Interstimulus interval (seconds)	Rating scale used	Comments
Hua et al. (2008)	NR	NR	Lower leg	Human (?)	Velvet, silk, cotton, terylene, soft plastic mesh and hard plastic mesh wrapped on a palmate resilient foam pad.	~10	400 g	NR	VAS	From preliminary psychophysical investigation
Huisman et al. (2016)	NR	Velocities and intensities	Non dominant lower arm	Robot (4 microdrives cylindrical vibration motors)	Fabric straps	0.5, 1, 3, 10, 30	Not force but intensity measured	NR	VAS	High intensity: 50% of motors maximum voltage, ±1.8 g peak amplitude, 200 Hz peak frequency Low intensity: 35% of motors maximum voltage, ±0.9 g peak amplitude, 140 Hz peak frequency
Jönsson et al. (2017)	Visual shielding goggles and headphones	Pseudorandomised velocities	Study 1: Left and right forearm Study 2: Left forearm	Study 1 and 2: Robot (RTS)	Study 1 and 2: 6 cm wide GHB	Study 1: 0.3, 1, 3, 10, 30 Study 2: 0.1, 0.3, 1, 3, 10, 30	Study 1 and 2: 0.4 N	NR	VAS	
Jönsson et al. (2015)	Goggles to inhibit peripheral vision Pink noise through headphones	Skin site and velocities	Study 1: Left thigh and left forearm Study 2: Left forearm	Robot (RTS)	50 mm wide GHB	Study 1 and 2: 0.3, 1, 3, 10, 30	0.4 N	NR	VAS	
Kass-Iliyya et al. (2016)	NR	Randomisation of velocities	Both forearms and shins	Human (?)	70 mm wide GHB	0.3, 3, 30	NR	10	VAS	

Table 3 (continued)

References	Blinding of pleasantness investigation	Randomisation within pleasantness method	Region (s) assessed	Mode of stimulus delivery	Stimulus utilised	Velocities cm/s	Force	Interstimulus interval (seconds)	Rating scale used	Comments
Löken et al. (2011)	Participants wore earplugs and stimulation site covered by curtain	Exp 1 and 2: Order of stimuli delivery Exp 3: Pseudorandomised velocities	Left palm or forearm	Exp 1 and 2: Robot (RTS) Exp 3: Human (f)	20 mm wide GHB	Exp 1 and 2: 0.1, 0.3, 1, 3, 10 and 30 Exp 3: 0.3, 1, 3, 10 and 30	Exp 1 and 2: 0.4 N	30	VAS	
Luong et al. (2017)	NR	Study 1: Pairs of velocities Study 2: Velocities	Study 1: Left forearm Study 2: forearm and palm	Study 1 and 2: Robot (RTS)	Study 1 and 2: 7 cm wide GHB	Study 1: 1 – 10.5 Study 2: 0.3 – 30	Study 1 and 2: 0.4 N	Study 1 and 2: 15	VAS	
Pawling et al. (2017)	Eyes closed to remove any visual distractions	Location and velocity	Palm and forearm	Human (m)	44 mm wide GHB	3, 30	0.3 N	NR	VAS	Multiple (x10) strokes of 30 cm/s also delivered
Triscoli et al. (2013)	First 2 conditions participants not informed about source of stimulation. All conditions headphones worn with pink noise and occluding glasses that blocked peripheral vision	Source: hand/robot and velocity	Left forearm	Human (f) and robot (RTS)	5 cm wide GHB	0.3, 3, 30	Robot 0.4 N Handheld: Visual controlling bending of bristles	15	VAS	
Triscoli et al. (2014)	Headphones to minimise noise and occluding glasses that blocked peripheral vision	Exp 1: Pseudorandomised velocities	Left forearm	Robot (RTS)	5 cm wide GHB	Exp 1: 0.3, 3, 30 Exp 2: 3,30	0.4 N	NR	VAS	

Table 3 (continued)

References	Blinding of pleasantness investigation	Randomisation within pleasantness method	Region(s) assessed	Mode of stimulus delivery	Stimulus utilised	Velocities cm/s	Force	Interstimulus interval (seconds)	Rating scale used	Comments
Tsalamlal et al. (2013)	Headphones playing white noise	Stimuli	Left forearm	Air jet system		Low movement velocity (0.6 rad/s) High movement velocity (12 rad/s)	Flow rate of 7.5 l/min and 50 l/min	NR	Likert scale	4 jet flow profiles (continuous low, discontinuous low, continuous high and discontinuous high)
Voos et al. (2013)	NR	NR	Right forearm	Human (?)	Brush	8 and 32	NR	NR	Likert scale	From pre-scan component

?: gender not stated, *Exp* experiment, *f* female, *m* male, *g* grams, *GHB* goat hair brush, *RTS* rotary tactile stimulator, *Hz* hertz, *mN* millinewton, *N* newton, *NR* not recorded

male participants ($k=3$) (Huisman et al., 2016; Triscoli et al., 2013; Tsalamlal et al., 2013), and three studies had an equal gender distribution (Ackerley et al., 2014; Etzi et al., 2018; Hua et al., 2008). Those that contained multiple experiments had either a greater proportion of females, or an equal gender distribution (Essick et al., 2010; Jönsson et al., 2015, 2017; Luong et al., 2017; Löken et al., 2011; Triscoli et al., 2014). In all but four studies, a visual analogue scale (VAS) was used to capture pleasantness ratings ($k=14$) (Ackerley et al., 2014; Essick et al., 2010; Etzi et al., 2014, 2018; Hua et al., 2008; Huisman et al., 2016; Jönsson et al., 2015, 2017; Kass-Iliyya et al., 2016; Luong et al., 2017; Löken et al., 2011; Pawling et al., 2017; Triscoli et al., 2013, 2014).

Application

There was an equal distribution in the mode of application of the tactile stimulus. Eight studies utilised a robot (Ackerley et al., 2014; Essick et al., 2010; Huisman et al., 2016; Jönsson et al., 2015, 2017; Luong et al., 2017; Tsalamlal et al., 2013; Triscoli et al., 2014), most often a rotary tactile stimulator (RTS) ($k=6$) (Ackerley et al., 2014; Essick et al., 2010; Jönsson et al., 2015; Jönsson et al., 2017; Luong et al., 2017; Triscoli et al., 2014), and eight studies had a human deliver the stimulus (Bennett et al., 2014; Etzi et al., 2014, 2018; Gordon et al., 2013; Hua et al., 2008; Kass-Iliyya et al., 2016; Pawling et al., 2017; Voos et al., 2013). In two studies, both an experimenter (female) and a robot (RTS) delivered the stimuli (Löken et al., 2011; Triscoli et al., 2013).

A number of different materials were used to contact the skin in the desired region. Most often a goat hair brush was utilised ($k=9$) (Ackerley et al., 2014; Jönsson et al., 2015, 2017; Kass-Iliyya et al., 2016; Luong et al., 2017; Löken et al., 2011; Pawling et al., 2017; Triscoli et al., 2013, 2014) with a width that ranged from 20 to 70 mm.

The time take between stimulations (interstimulus interval) ranged from 80 to 30 s (Etzi et al., 2018; Löken et al., 2011), with 10 and 15 s most frequently utilised ($k=4$) (Ackerley et al., 2014; Kass-Iliyya et al., 2016; Luong et al., 2017; Triscoli et al., 2013).

Various approaches were employed to conceal the stimulation site from the participant. For example, visual input was prevented by using glasses blocking peripheral vision ($k=5$) (Ackerley et al., 2014; Jönsson et al., 2015, 2017; Triscoli et al., 2013, 2014) or a curtain ($k=2$) (Essick et al., 2010; Löken et al., 2009). Eight studies did not state if any measures were taken to blind the participant from testing. Furthermore, audible noises were reduced with earplugs ($k=2$) (Essick et al., 2010; Löken et al., 2009) or headphones ($k=6$) (Etzi et al., 2014, Jönsson et al., 2015, 2017; Triscoli et al., 2013, 2014; Tsalamlal et al., 2013).

Velocity

The velocity of stimulus application was the most common factor randomised and fell within (and including) 0.3–32 cm/s ($k = 10$) (Ackerley et al., 2014; Huisman et al., 2016; Jönsson et al., 2015, 2017; Kass-Iliyya et al., 2016; Luong et al., 2017; Löken et al., 2011; Pawling et al., 2017; Triscoli et al., 2013, 2014). Tsalamlal et al. (2013) utilised air as the stimulus and described velocity as low movement (0.6 rad/s) and high movement (12 rad/s). Fourteen studies investigated more than one velocity for stimulus application with velocities of 0.3, 1, 3, 10 and 30 cm/s most often used (Ackerley et al., 2014; Essick et al., 2010; Etzi et al., 2018; Huisman et al., 2016; Jönsson et al., 2015, 2017; Kass-Iliyya et al., 2016; Luong et al., 2017; Löken et al., 2011; Pawling et al., 2017; Tsalamlal et al., 2013; Triscoli et al., 2013, 2014; Voos et al., 2013). In 9 of these studies, a significant main effect of pleasantness ratings (all $p < 0.05$) was found (Ackerley et al., 2014; Essick et al., 2010; Huisman et al., 2016; Jönsson et al., 2015, 2017; Luong et al., 2017; Triscoli et al., 2013, 2014; Tsalamlal et al., 2013). This comprised utilising a robot (RTS) with a goat hair brush ($k = 5$) (Ackerley et al., 2014; Jönsson et al., 2015, 2017; Luong et al., 2017; Triscoli et al., 2014), and a robot applying multiple fabrics ($k = 2$) (Essick et al., 2010; Huisman et al., 2016). Triscoli et al. (2013) had both a human and robot administer a goat hair brush, and Tsalamlal et al., 2013 utilised air in their study. The remaining studies did not report if velocity had a main effect (Etzi et al., 2018; Kass-Iliyya et al., 2016; Löken et al., 2011; Pawling et al., 2017; Voos et al., 2013).

A stroking velocity of 3 cm/s compared to slower [0.3 cm/s, $P \leq 0.001$ (Triscoli et al., 2013; 2014); 0.5 cm/s, $p < 0.05$ (Huisman et al., 2016)] and faster velocities (30 cm/s, $P \leq 0.035$) (Ackerley et al., 2014; Etzi et al., 2018; Pawling et al., 2017; Triscoli et al., 2013, 2014) was rated on average as most pleasant (Kass-Iliyya et al., 2016; Löken et al., 2011). Triscoli et al. found that there was no significant difference in mean pleasantness scores when stroked at velocities of 0.3 or 30 cm/s [$P > 0.05$ (Triscoli et al., 2013) and $P = 0.632$ (Triscoli et al., 2014)]. Another study found that 8 cm/s was more pleasant than 32 cm/s ($P = 0.003$) (Voos et al., 2013) as was low air jet movement velocity (0.6 rad/s) versus static and high velocity (12 rad/s) ($p < 0.05$) (Triscoli et al. 2013).

The velocity-pleasantness profile was continuously identified to be best fit by a quadratic model (Ackerley et al., 2014; Huisman et al., 2016; Jönsson et al., 2015; 2017; Luong et al., 2017; Löken et al., 2011). When velocity against pleasantness ratings was plotted, the resultant profile was an inverse U-shape with peak ratings in the range of 1–3 cm/s.

The duration of stroking (multiple strokes over time) at the same velocity was also found to contribute to the pleasantness experienced. Triscoli et al. (2014) found that for stroking at 3 cm/s there was a significant decrease in participant pleasantness ratings over time (40 trials of 5 back and forth strokes, $P = 0.042$). This was not observed when stroking at 0.3 cm/s ($P = 0.290$) or 30 cm/s ($P = 0.617$). In addition, the pleasantness ratings at the end of the session, when further trials at 3 cm/s were performed, showed a decline between two groups of healthy participants (40 vs. 120 trials, $P = 0.003$). Such a decline over time was not seen at a velocity of 30 cm/s (40 vs. 267 trials, $P = 0.349$). In contrast, (Etzi et al., 2018) identified that only fast stroking (30 cm/s) was rated as less pleasant over time (9 s vs. 60 s, $P = 0.02$) compared to slower stroking (3 cm/s).

Force

The force of stimulus delivery ranged from 0.19 N to 400 grams (corresponding to 3.92 N) ($k = 10$) (Ackerley et al., 2014; Essick et al., 2010; Hua et al., 2008; Jönsson et al., 2015, 2017; Luong et al., 2017; Löken et al., 2011; Pawling et al., 2017; Triscoli et al., 2013, 2014) with a force of 0.4 N most often used ($k = 7$) (Ackerley et al., 2014; Jönsson et al., 2015; 2017; Luong et al., 2017; Löken et al., 2011; Triscoli et al., 2013, 2014). Seven studies did not control or did not state the force applied (Bennett et al., 2014; Etzi et al., 2014, 2018; Gordon et al., 2013; Huisman et al., 2016; Kass-Iliyya et al., 2016; Voos et al., 2013). Use of an air jet with a flow rate of 7.5 and 50 l/min was described by Triscoli et al. (2013). However, they did not identify what force levels these flow rates exerted onto the desired region.

Essick et al. (2010) was the only included study that investigated the effect of different forces. They found that as force increased the pleasantness scores decreased (experiment 1 and 2, $p < 0.001$). Tsalamlal et al. (2013) measured the flow rate of air onto the skin surface. The study found that low intensity flow (7.5 l/min) was rated significantly more pleasant than the high intensity flow (50 l/min) (continuous and discontinuous flows, $p < 0.0001$). A similar observation was also identified by Huisman et al. (2016), with low intensity stimuli, delivered by a robot rated more pleasant than high ($p < 0.05$).

Site

The most common site for stimulus application was the forearm ($k = 16$) (Ackerley et al., 2014; Essick et al., 2010; Etzi et al., 2014, 2018; Gordon et al., 2013; Huisman et al., 2016; Jönsson et al., 2015; 2017; Kass-Iliyya et al., 2016; Luong et al., 2017; Löken et al., 2011; Pawling et al., 2017; Triscoli et al., 2013, 2014; Tsalamlal et al., 2013; Voos et al., 2013). Eleven studies applied a stimulus to more than one site (Ackerley et al., 2014; Bennett et al., 2014; Essick et al., 2010; Etzi et al., 2014; Gordon et al., 2013; Jönsson et al., 2015; 2017; Kass-Iliyya et al., 2016; Luong et al., 2017; Löken et al., 2011; Pawling et al., 2017) (see Table 2), with

an interaction with perceived pleasantness investigated in 8 studies (Ackerley et al., 2014; Bennett et al., 2014; Essick et al., 2010; Etzi et al., 2014; Gordon et al., 2013; Luong et al., 2017; Löken et al., 2011; Pawling et al., 2017). The most common comparisons included the forearm and the palm ($k = 8$) (Ackerley et al., 2014; Bennett et al., 2014; Essick et al., 2010; Etzi et al., 2014; Gordon et al., 2013; Luong et al., 2017; Löken et al., 2011; Pawling et al., 2017), with 5 studies concluding that applying the stimulus to the arm was rated as significantly more pleasant ($p < 0.05$, irrespective of velocity, material or mode of application) (Bennett et al., 2014; Essick et al., 2010; Etzi et al., 2014; Gordon et al., 2013; Löken et al., 2011). Two studies found no significant interaction between sites (Luong et al., 2017; Pawling et al., 2017) and 1 study found that pleasantness ratings were significantly lower when stoking the arm compared to palm (velocity of 30 cm/s, $P = 0.028$) (Ackerley et al., 2014).

Gender

Seven studies considered the effect of gender on perceived pleasantness (Ackerley et al., 2014; Essick et al., 2010; Etzi et al., 2018; Jönsson et al., 2015, 2017; Löken et al., 2011; Triscoli et al., 2013). Jönsson et al. (2017) found that women rated pleasantness to touch significantly higher than men (combined study 1 and 2: $P > 0.001$). Essick et al. (2010) identified significant interactions of gender with body site (experiment 2: $p < 0.001$), material (experiment 1 and 2, $p < 0.002$), body site by material (experiment 2: $p < 0.001$)

and velocity (experiment 1: $p < 0.001$ but not for experiment 2: $P = 0.26$). Furthermore, they also identified that pleasantness ratings differed between males and females for increasing force (experiment 2: $p < 0.001$). Within this experiment, both genders had a significant decrease in pleasantness ratings with increasing force ($p < 0.05$); however, the decrease was greater for males ($p < 0.05$). The remaining five studies found no significant effect of gender on pleasantness ratings (Ackerley et al., 2014; Etzi et al., 2018; Löken et al., 2011; Jönsson et al., 2015; Triscoli et al., 2013).

Quantitative synthesis (meta-analysis)

To reduce heterogeneity in the meta-analysis, studies that utilised the same material for stimulus delivery on the same site were included. Hence, quantitative synthesis was possible for velocities of 3 and 30 cm/s in the region of the forearm using a brush stimulus. A velocity of 3 cm/s (pleasant and CT favourable) was treated as the pretest score and 30 cm/s (less pleasant and non-CT favourable) as the posttest score. One study provided the raw data of a single stroke at 3 cm/s and 30 cm/s on the forearm for which a pretest–posttest correlation of 0.48 was calculated (Pawling et al., 2017). This estimate was generalised to the remaining studies to allow the computation of the variance of the effect estimate (Morris & DeShon, 2002; Dunlap et al., 1996; Borenstein et al., 2011).

Five studies undertook comparisons of stroking at 3 cm/s and 30 cm/s with similar experimental set-ups for which
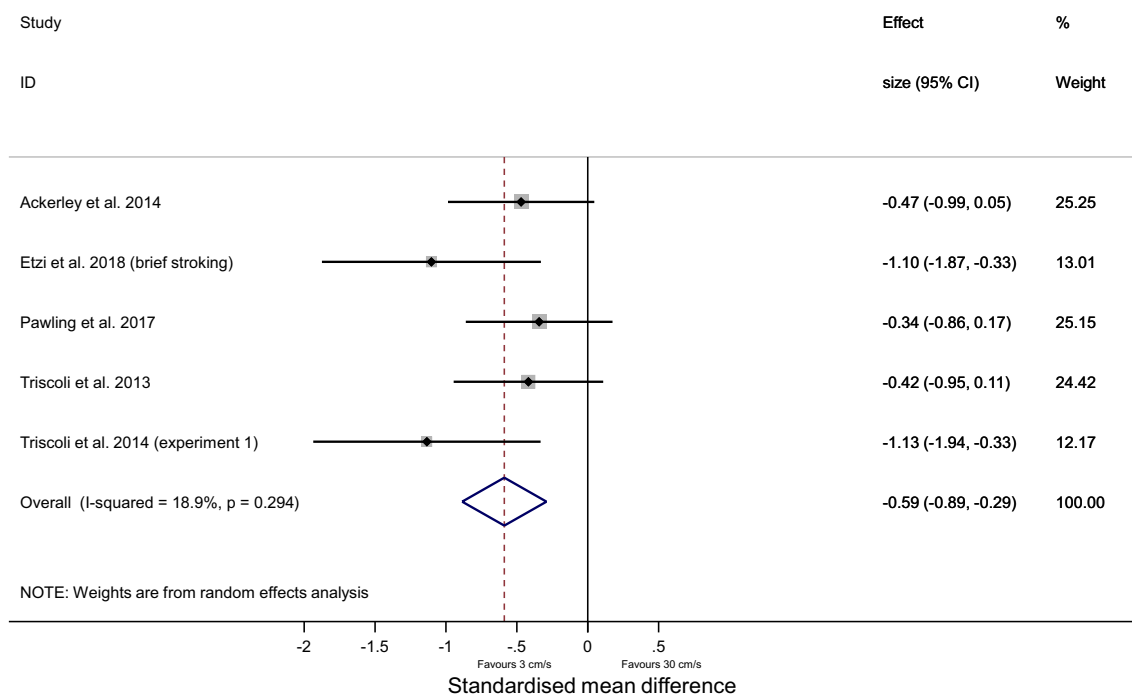

Study ID	Effect size (95% CI)	% Weight
Ackerley et al. 2014	-0.47 (-0.99, 0.05)	25.25
Etzi et al. 2018 (brief stroking)	-1.10 (-1.87, -0.33)	13.01
Pawling et al. 2017	-0.34 (-0.86, 0.17)	25.15
Triscoli et al. 2013	-0.42 (-0.95, 0.11)	24.42
Triscoli et al. 2014 (experiment 1)	-1.13 (-1.94, -0.33)	12.17
Overall (I-squared = 18.9%, p = 0.294)	-0.59 (-0.89, -0.29)	100.00

NOTE: Weights are from random effects analysis

Favours 3 cm/s — Favours 30 cm/s

Standardised mean difference

Fig. 2 Meta-analysis forest plot providing pooled estimate of perceived pleasantness

effect sizes were calculated (Ackerley et al., 2014; Etzi et al., 2018; Pawling et al., 2017; Triscoli et al., 2013, 2014). The overall standardised effect size estimate for velocity was -0.59 (95% confidence interval [CI] -0.89, -0.29; $p < 0.001$) which is considered a medium effect size (Fig. 2) (Cohen, 1992). This supports that a stroke velocity of 30 cm/s is rated as less pleasant than a stroke at 3 cm/s using a brush stimulus on the forearm, furthermore complementing the findings from the systematic review. The I^2 test indicated little variability between studies (18.9%), and the Cochrane Q test determined that there was no statistically significant heterogeneity present (Q [df = 4] = 4.94, $P = 0.29$). Consequently, a subgroup analysis was not performed.

Sensitivity analysis showed little change in the overall effect when using a fixed effects model (-0.57, CI -0.83, -0.57; $p < 0.001$), and the leave-one-out analysis had minimal effect (min = 0.50, max = 0.68, all P values < 0.001).

An assessment of publication bias was not performed due to inadequate numbers of included studies to properly assess a funnel plot or more advanced regression-based assessments (Higgins & Green, 2011).

Discussion

To the authors' knowledge, this is the first systematic review performed in order to identify the optimum psychophysical stimulus to elicit a pleasant sensation. Furthermore, it is the first meta-analysis that has been conducted to establish an estimate of effect between a pleasant (CT favourable) and less pleasant (non-CT favourable) velocity.

The qualitative synthesis found that stroking at a velocity of 3 cm/s was repeatedly found to be significantly more pleasant than slower [0.3 cm/s (Triscoli et al., 2013, 2014)] or faster [30 cm/s (Ackerley et al., 2014; Etzi et al., 2018; Pawling et al., 2017; Triscoli et al., 2013, 2014)] velocities, which in turn were rated as pleasant as each other. This is not surprising as studies have established that a velocity of 3 cm/s is CT favourable, in turn, giving rise to the characteristic inverse U-shape profile seen within several studies in this review (Ackerley et al., 2014; Huisman et al., 2016; Jönsson et al., 2015; Luong et al., 2017; Löken et al., 2011). The meta-analysis reinforced these findings as 30 cm/s was found to be less pleasant relative to 3 cm/s from the pooled data. However, the available data were limited restricting the conclusions that could be drawn. Nonetheless, velocity plays a fundamental role in the perception of a pleasant sensation and provides for an ideal variable to randomise and investigate within the methodology of experimental studies involving tactile pleasantness.

The ability to maintain a constant velocity when delivered by a human experimenter may be a technical limitation. In

addition, the force of application presents the same problem. Lighter forces are rated as most pleasant (Essick et al., 2010) and maintaining a light force, particularly through a dynamic movement, may inevitably create fluctuations. Methods incorporated to overcome these potential issues included using auditory signals (Etzi et al., 2018) visual aids (e.g. moving bar on a screen) (Etzi et al., 2014; Kass-Iliyya et al., 2016; Löken et al., 2011; Pawling et al., 2017); copying the stroke from an adjacent robot (Triscoli et al., 2013) or visual control in the bend of the brush bristles (Triscoli et al., 2013). A need for stimulus control may justify the use of a robot as the mode of stimulus delivery. Nonetheless, previous research has found that there is no difference in perceived pleasantness when a stimulus is delivered by a human or robot (Triscoli et al., 2013). However, too few included studies in the meta-analysis did not allow this factor to be investigated by a subgroup analysis.

The associated disadvantages with a robotic device would be the costs associated, as well as limitations on the ease of use for the robot to access and deliver a stimulus on certain regions of the body. Inevitably, the use of a robot could also reduce any potential confounding factor caused by the gender of the experimenter (Essick et al., 2010).

Within pain sensitivity studies, an interaction between gender and psychophysical responses have been demonstrated, whereby males behave differently in response to a painful stimulation in the presence of a female, compared to male, experimenter (Heslin et al., 1983). The studies included in this review did not investigate the interaction between the experimenters' and participants' gender on perceived pleasantness ratings. Instead, the gender of the participant was investigated with the majority of studies finding there was no significant difference between gender and pleasantness ratings (Ackerley et al., 2014; Etzi et al., 2018; Löken et al., 2011; Jönsson et al., 2015; Triscoli et al., 2013). However, the effect of gender interactions may lie within more complex patterns, e.g. interactions with body site, material, or velocity (Essick et al., 2010).

Whether handheld or via a robot, methodologies varied with which material was used to apply the stimuli. Three studies assessed pleasantness ratings from different materials (Essick et al., 2010; Etzi et al., 2014; Hua et al., 2008) and were consistent with each other, whereby the softest materials evaluated were rated as most pleasant. This further justifies why a goat hair brush was most often used (Ackerley et al., 2014; Case, Čeko et al., 2016, Case, Laubacher et al., 2016; Croy et al., 2016; Ellingsen et al., 2013; Jönsson et al., 2015; Kass-Iliyya et al., 2016; Liljencrantz et al., 2013, 2014; Löken et al., 2011; Sailer et al., 2016; Triscoli et al., 2013, 2014; Trotter et al., 2016), as it allows the uniform application of a soft material with ideal characteristics of a light force and velocity (Vallbo et al., 1999). Therefore, a brush would provide as an optimum "prototypical pleasant

stimulus" as it may "remain pleasant irrespective of how it is moved across the skin, even if not delivered with the optimal stimulus parameters" (Guest & Essick, 2016). A brush was also a suitable stimulus for studies on the palm. Although the glabrous skin does not contain CT afferents, a brush stroke still elicits a pleasant sensation (Ackerley et al., 2014; Gordon et al., 2013; Luong et al., 2017; Löken et al., 2011; Pawling et al., 2017). This is hypothesised to be from learnt behaviour from top-down mechanisms as well as emotional memory circuits (Löken et al., 2011; McGlone et al., 2012). It should also not be overlooked that there is likely an important contributory effect from the Aβ pathway (McGlone et al., 2014) and may also justify why studies presented contrasting findings, whereby optimal and suboptimal CT velocities were equally pleasant in some (3 cm/s vs. 30 cm/s) (Triscoli et al., 2013, 2014), yet significantly different in others (8 cm/s vs. 32 cm/s) (Voos et al., 2013).

The number of strokes delivered also played a role in the perceived pleasantness. Long-lasting stroking at a pleasant and CT favourable velocity caused a small reduction in the pleasantness ratings (Triscoli et al., 2014). This, in part, may have resulted from a property of CT afferents described as fatigue, where the firing of the afferents is reduced upon repeated stimulation (Vallbo et al., 1999; Bessou et al., 1971; Morrison, 2012). Post-activation depression of CT afferents has been identified, in which they reduce their responsiveness to succeeding stimuli following an initial touch (Vallbo et al., 1999; Iggo, 1960). It was found that a resting period of greater duration allowed for better recovery of CT afferents, but full recovery could take several minutes (Iggo, 1960; Iggo & Kornhuber, 1977). Within the methodologies reviewed, the interstimulus interval ranged from 80 to 30 s. It would be expected that repeated stimulations with short intervals could thus reduce the pleasantness response, as the CT afferents are fatigued and have not had sufficient time to recover. Studies varied in the number of strokes delivered, and the fatigue effect may present a source of heterogeneity as it cannot be assumed to be equal across studies.

The VAS was most often utilised within the included studies. Known to allow evaluation of individuals' experience of a phenomenon of interest (Wewers & Lowe, 1990), as seen in pain studies, it provides a method to transfer a sensation to a measurable dimension and allows a reliable way to assess what the patient actually feels (Ohnhaus & Adler, 1975). In addition, the VAS is one of the most common forms of rating scales utilised in pain assessment, with proven validity and reliability (Ferreira-Valente et al., 2011; Price et al., 1983; Hawker et al., 2011). However, the VAS is not without disadvantages. The data obtained are not always normally distributed, and the entirety of the scale is not always utilised (Williamson & Hoggart, 2005).

The systematic review and meta-analysis are associated with limitations. The search strategy was restricted to English due to resource limitations. Furthermore, the inclusion criteria were strict and led to the exclusion of studies that may be considered fundamental in CT afferent/pleasantness research (Ackerley et al., 2014; Croy et al., 2016; Moher et al., 2009; Morrison et al., 2011). The quality assessment designed was also unique with no weight for each item.

Within the meta-analysis, to compute the effect estimate, the pretest–posttest correlation was derived from 1 study and generalised to the others. This could result in a less precise pooled estimate as the correlation may change because of study characteristics such as duration of stroking and number of trials. However, with limited data this method was considered the most appropriate (Morris & DeShon, 2002; Dunlap et al., 1996).

Effect size multiplicity was dealt with by a reductionist approach, in order to resolve multiple effect estimates in a single study. As a result, the exploration of heterogeneity within studies was restricted (López-López et al. 2018).

Although the result of the Cohen's Q test was not statistically significant, this does not necessarily indicate homogeneity. The efficacy in detecting true heterogeneity is reduced when only few studies are included due to a lack of power (Higgins et al., 2003). This also precluded a subgroup analysis, as well as evaluation of publication bias.

Conclusion

This review focused on passively received stimuli that were assessed for pleasantness and showed that careful standardisation of methodological factors such as texture, velocity, and force, as well as the duration of continuous stroking are most likely to play key roles, namely using a brush and stroking at a velocity of approximately 3 cm/s, with a light force, which is also the optimal type of stimulus in activating CT afferents. Having taken some of these factors into account, the meta-analysis confirmed that a stroking velocity of 30 cm/s was rated as significantly less pleasant than 3 cm/s.

The standardised stimulations identified within the included studies showed that the assessment of pleasantness bares many parallels to the assessment of pain and other somatosensory modalities.

Compliance with ethical standards

Appendix

See Table 4

Table 4 14 item quality assessment checklist

Item:	Yes (1)	No (0)	Unclear (0)
1. Were eligibility criteria for study entry specified?			
2. Was the sample size justified?			
3. Is the method adequately described (sufficient to repeat method)?			
4. Was the stimulation adequately controlled (i.e. measurement device to control force/velocity)?			
5. Was a form of randomisation used for stimulation e.g. velocity/force?			
6. Were participants blinded to the stimulus characteristics e.g. velocity/force			
7. Were participants blinded to the site of stimulus delivery?			
8. Were those delivering the stimuli blinded to the stimulation characteristics e.g. velocity/force			
9. Was the time interval between stimuli stated? (NOT between blocks if used)			
10. Were confounding factors identified?			
11. Were strategies to deal with confounding factors stated?			
12. Were outcomes measured using a reliable method?			
13. Was appropriate statistical analysis used?			
14. Were there any conflicts of interest related to the methodology e.g. instrumentation for stimulation delivery/recording (reverse scored)	(0)	(1)	

Score for each question contained within brackets. Total score = total score/total number of items × 100. Appraisal score: Weak: ≤ 33%, Satisfactory: ≥ 34% ≤ 66%, Good: ≥ 67%

References

Ackerley, R., Carlsson, I., Wester, H., Olausson, H., & Backlund Wasling, H. (2014). Touch perceptions across skin sites: Differences between sensitivity, direction discrimination and pleasantness. *Frontiers in Behavioral Neuroscience*, 8, 54. https://doi.org/10.3389/fnbeh.2014.00054.

Bennett, R. H., Bolling, D. Z., Anderson, L. C., Pelphrey, K. A., & Kaiser, M. D. (2014). fNIRS detects temporal lobe response to affective touch. *Social Cognitive and Affective Neuroscience, 9,* 470–476.

Bessou, P., Burgess, P., Perl, E., & Taylor, C. (1971). Dynamic properties of mechanoreceptors with unmyelinated (C) fibers. *Journal of Neurophysiology*, 34, 116–131.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2011). *Introduction to meta-analysis*. Hoboken: Wiley.

Case, L. K., Čeko, M., Gracely, J. L., Richards, E. A., Olausson, H., & Bushnell, M. C. (2016a). Touch perception altered by chronic pain and by opioid blockade. *eNeuro, 3,* 0138.

Case, L. K., Laubacher, C. M., Olausson, H., Wang, B. Q., Spagnolo, P. A., & Bushnell, M. C. (2016b). Encoding of touch intensity but not pleasantness in human primary somatosensory cortex. *Journal of Neuroscience, 36,* 5850–5860.

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112,* 155.

Croy, I., Geide, H., Paulus, M., Weidner, K., & Olausson, H. (2016). Affective touch awareness in mental health and disease relates to autistic traits—An explorative neurophysiological investigation. *Psychiatry Research, 245,* 491–496.

Dunlap, W. P., Cortina, J. M., Vaslow, J. B., & Burke, M. J. (1996). Meta-analysis of experiments with matched groups or repeated measures designs. *Psychological Methods, 1,* 170.

Ellingsen, D.-M., Wessberg, J., Eikemo, M., Liljencrantz, J., Endestad, T., Olausson, H., & Leknes, S. (2013). Placebo improves pleasure and pain through opposite modulation of sensory processing. *Proceedings of the National Academy of Sciences of the United States of America, 110,* 17993–17998.

Essick, G. K., McGlone, F., Dancer, C., Fabricant, D., Ragin, Y., Phillips, N., … Guest, S. (2010). Quantitative assessment of pleasant touch. *Neuroscience and Biobehavioral Reviews, 34,* 192–203.

Etzi, R., Carta, C., & Gallace, A. (2018). Stroking and tapping the skin: behavioral and electrodermal effects. *Experimental Brain Research, 236,* 453–461.

Etzi, R., Spence, C., & Gallace, A. (2014). Textures that we like to touch: An experimental study of aesthetic preferences for tactile stimuli. *Consciousness and Cognition, 29,* 178–188.

Faraone, S. V. (2008). Interpreting estimates of treatment effects: Implications for managed care. *Pharmacy and Therapeutics, 33,* 700.

Ferreira-Valente, M. A., Pais-Ribeiro, J. L., & Jensen, M. P. (2011). Validity of four pain intensity rating scales. *Pain, 152,* 2399–2404.

Gallace, A., & Spence, C. (2010). The science of interpersonal touch: an overview. *Neuroscience and Biobehavioral Reviews, 34,* 246–259.

Gordon, I., Voos, A. C., Bennett, R. H., Bolling, D. Z., Pelphrey, K. A., & Kaiser, M. D. (2013). Brain mechanisms for processing affective touch. *Human Brain Mapping, 34,* 914–922.

Guest, S., & Essick, G. K. (2016). Psychophysical assessment of the sensory and affective components of touch. *Affective Touch and the Neurophysiology of CT Afferents*. New York: Springer.

Hawker, G. A., Mian, S., Kendzerska, T., & French, M. (2011). Measures of adult pain: Visual analog scale for pain (VAS pain), numeric rating scale for pain (NRS pain), McGill pain questionnaire (MPQ), short-form McGill pain questionnaire (SF-MPQ), chronic pain grade scale (CPGS), short form-36 bodily pain scale (SF-36 BPS), and measure of intermittent and constant osteoarthritis pain (ICOAP). *Arthritis Care & Research, 63,* S240–S252.

Heslin, R., Nguyen, T. D., & Nguyen, M. L. (1983). Meaning of touch: The case of touch from a stranger or same sex person. *Journal of Nonverbal Behavior, 7,* 147–157.

Higgins, J. P. T., & Green, S. E. (2011). *Cochrane handbook for systematic reviews of interventions version 5.1.0 [updated March 2011]* (Online). The Cochrane Collaboration. www.handbook.cochrane.org. Accessed September 2018.

Higgins, J. P., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21,* 1539–1558.

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *British Medical Journal, 327,* 557.

Hua, Q. P., Zeng, X. Z., Liu, J. Y., Wang, J. Y., Guo, J. Y., & Luo, F. (2008). Dynamic changes in brain activations and functional connectivity during affectively different tactile stimuli. *Cellular and Molecular Neurobiology, 28,* 57–70.

Huisman, G., Frederiks, A. D., van Erp, J. B. F., & Heylen, D. K. J. (2016). Simulating affective touch: Using a vibrotactile array to generate pleasant stroking sensations. In F. Bello, H. Kajimoto, & Y. Visell (Eds.), *Haptics: Perception, devices, control, and applications*. Cham: Springer.

Iggo, A. (1960). Cutaneous mechanoreceptors with afferent C fibres. *Journal of Physiology, 152,* 337–353.

Iggo, A., & Kornhuber, H. (1977). A quantitative study of C-mechanoreceptors in hairy skin of the cat. *Journal of Physiology, 271,* 549–565.

Israel, H., & Richter, R. R. (2011). A guide to understanding meta-analysis. *Journal of Orthopaedic & Sports Physical Therapy, 41,* 496–504.

Jönsson, E. H., Backlund Wasling, H., Wagnbeck, V., Dimitriadis, M., Georgiadis, J. R., Olausson, H., & Croy, I. (2015). Unmyelinated tactile cutaneous nerves signal erotic sensations. *Journal of Sexual Medicine, 12,* 1338–1345.

Jönsson, E. H., Bendas, J., Weidner, K., Wessberg, J., Olausson, H., Wasling, H. B., & Croy, I. (2017). The relation between human hair follicle density and touch perception. *Scientific Reports, 7,* 2499.

Kass-Iliyya, L., Leung, M., Marshall, A., Trotter, P., Kobylecki, C., Walker, S., et al. (2016). The perception of affective touch in Parkinson's disease and its relation to small fibre neuropathy. *European Journal of Neuroscience*, *45,* 232–237.

Liljencrantz, J., Björnsdotter, M., Morrison, I., Bergstrand, S., Ceko, M., Seminowicz, D. A., … Olausson, H. (2013). Altered C-tactile processing in human dynamic tactile allodynia. *Pain, 154,* 227–234.

Liljencrantz, J., Marshall, A., Ackerley, R., & Olausson, H. (2014). Discriminative and affective touch in human experimental tactile allodynia. *Neuroscience Letters, 563,* 75–79.

Löken, L. S., Evert, M., & Wessberg, J. (2011). Pleasantness of touch in human glabrous and hairy skin: Order effects on affective ratings. *Brain Research, 1417,* 9–15.

Löken, L. S., Wessberg, J., Morrison, I., McGlone, F., & Olausson, H. (2009). Coding of pleasant touch by unmyelinated afferents in humans. *Nature Neuroscience, 12,* 547–548.

López-López, J. A., Page, M. J., Lipsey, M. W., & Higgins, J. P. (2018). Dealing with effect size multiplicity in systematic reviews and meta-analyses. *Research Synthesis Methods, 9,* 336–351.

Luong, A., Bendas, J., Etzi, R., Olausson, H., & Croy, I. (2017). The individual preferred velocity of stroking touch as a stable measurement. *Physiology & Behavior, 177,* 129–134.

McGlone, F., Olausson, H., Boyle, J. A., Jones-Gotman, M., Dancer, C., Guest, S., & Essick, G. (2012). Touching and feeling: Differences in pleasant touch processing between glabrous and hairy skin in humans. *European Journal of Neuroscience, 35,* 1782–1788.

McGlone, F., Wessberg, J., & Olausson, H. (2014). Discriminative and affective touch: Sensing and feeling. *Neuron, 82,* 737–755.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6,* e1000097.

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods, 7,* 105.

Morrison, I. (2012). CT afferents. *Current Biology, 22,* 77–78.

Morrison, I., Björnsdotter, M., & Olausson, H. (2011). Vicarious responses to social touch in posterior insular cortex are tuned to pleasant caressing speeds. *Jouranl of Neuroscience, 31,* 9554–9562.

Morrison, I., Loken, L. S., & Olausson, H. (2010). The skin as a social organ. *Experimental Brain Research, 204,* 305–314.

Nordin, M. (1990). Low-threshold mechanoreceptive and nociceptive units with unmyelinated (C) fibres in the human supraorbital nerve. *Journal of Physiology, 426,* 229.

Ohnhaus, E. E., & Adler, R. (1975). Methodological problems in the measurement of pain: a comparison between the verbal rating scale and the visual analogue scale. *Pain, 1,* 379–384.

Olausson, H. W., Cole, J., Vallbo, Å., McGlone, F., Elam, M., Krämer, H. H., … Bushnell, M. C. (2008). Unmyelinated tactile afferents have opposite effects on insular and somatosensory cortical processing. *Neuroscience Letters, 436,* 128–132.

Olausson, H., Lamarre, Y., Backlund, H., Morin, C., Wallin, B., Starck, G., … Vallbo, Å. (2002). Unmyelinated tactile afferents signal touch and project to insular cortex. *Nature Neuroscience, 5,* 900–904.

Olausson, H., Wessberg, J., McGlone, F., & Vallbo, Å. (2010). The neurophysiology of unmyelinated tactile afferents. *Neuroscience and Biobehavioral Reviews, 34,* 185–191.

Pawling, R., Cannon, P. R., McGlone, F. P., & Walker, S. C. (2017). C-tactile afferent stimulating touch carries a positive affective value. *PLoS ONE, 12,* e0173457.

Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2006). Comparison of two methods to detect publication bias in meta-analysis. *JAMA, 295,* 676–680.

Pigg, M., Baad-Hansen, L., Svensson, P., Drangsholt, M., & List, T. (2010). Reliability of intraoral quantitative sensory testing (QST). *Pain, 148,* 220–226.

Poort, L. J., van Neck, J. W., & van der Wal, K. G. (2009). Sensory testing of inferior alveolar nerve injuries: A review of methods used in prospective studies. *Journal of Oral and Maxillofacial Surgery, 67,* 292–300.

Price, D. D., McGrath, P. A., Rafii, A., & Buckingham, B. (1983). The validation of visual analogue scales as ratio scale measures for chronic and experimental pain. *Pain, 17,* 45–56.

Rolke, R., Magerl, W., Campbell, K. A., Schalber, C., Caspari, S., Birklein, F., & Treede, R. D. (2006). Quantitative sensory testing: A comprehensive protocol for clinical trials. *European Journal of Pain, 10,* 77.

Sailer, U., Triscoli, C., Haggblad, G., Hamilton, P., Olausson, H., & Croy, I. (2016). Temporal dynamics of brain activation during 40 min of pleasant touch. *NeuroImage, 139,* 360–367.

Sedgwick, P. (2013). Meta-analyses: heterogeneity and subgroup analysis. *British Medical Journal, 346,* f4040.

Svensson, P., Baad-Hansen, L., Pigg, M., List, T., Eliav, E., Ettlin, D., … Jääskeläinen, S. (2011). Guidelines and recommendations for assessment of somatosensory function in oro-facial pain conditions—A taskforce report. *Journal of Oral Rehabilitation, 38,* 366–394.

Taneja, P., Olausson, H., Trulsson, M., Vase, L., Svensson, P., & Baad-Hansen, L. (2019). Assessment of experimental orofacial pain, pleasantness and unpleasantness via standardised psychophysical testing. *European Journal of Pain*, *23,* 1297–1308.

Triscoli, C., Ackerley, R., & Sailer, U. (2014). Touch satiety: Differential effects of stroking velocity on liking and wanting touch over repetitions. *PLoS ONE, 9*(11), e113425.

Triscoli, C., Olausson, H., Sailer, U., Ignell, H., & Croy, I. (2013). CT-optimized skin stroking delivered by hand or robot is comparable. *Frontiers in Behavioral Neuroscience*, 7, 208. https://doi.org/10.3389/fnbeh.2013.00208.

Trotter, P. D., McGlone, F., McKie, S., McFarquhar, M., Elliott, R., Walker, S. C., … Foxe, J. (2016). Effects of acute tryptophan depletion on central processing of CT-targeted and discriminatory touch in humans. *European Journal of Neuroscience, 44,* 2072–2083.

Tsalamlal, M. Y., Ouarti, N., Martin, J.-C. & Ammi, M. (2013). EmotionAir: Perception of emotions from air jet based tactile stimulation. In *2013 Humaine association conference on affective computing and intelligent interaction (ACII)* (pp. 215–220). IEEE.

Vallbo, Å., Olausson, H., & Wessberg, J. (1999). Unmyelinated afferents constitute a second system coding tactile stimuli of the human hairy skin. *Journal of Neurophysiology, 81,* 2753–2763.

Vallbo, A. B., Olausson, H. K., Wessberg, J., & Norrsell, U. (1993). A system of unmyelinated afferents for innocuous mechanoreception in the human skin. *Brain Research, 628,* 301–304.

Voos, A. C., Pelphrey, K. A., & Kaiser, M. D. (2013). Autistic traits are associated with diminished neural response to affective touch. *Social Cognitive and Affective Neuroscience, 8,* 378–386.

Walker, S. C., Trotter, P. D., Woods, A., & McGlone, F. (2017). Vicarious ratings of social touch reflect the anatomical distribution & velocity tuning of C-tactile afferents: A hedonic homunculus? *Behavioural Brain Research, 320,* 91–96.

Wewers, M. E., & Lowe, N. K. (1990). A critical review of visual analogue scales in the measurement of clinical phenomena. *Research in Nursing & Health, 13,* 227–236.

Williamson, A., & Hoggart, B. (2005). Pain: A review of three commonly used pain rating scales. *Journal of Clinical Nursing, 14,* 798–804.