Facial Emotion Perception in Patients with Epilepsy:
A Systematic Review with Meta-Analysis

Michelle Edwards\textsuperscript{ac}
Elizabeth Stewart\textsuperscript{ac}
Associate Professor Romina Palermo\textsuperscript{bc}
Associate Professor Suncica Lah\textsuperscript{ac}*  

\textsuperscript{a} School of Psychology, The University of Sydney, Sydney, NSW, Australia
\textsuperscript{b} School of Psychological Science, The University of Western Australia, Perth, WA, Australia
\textsuperscript{c} ARC Centre of Excellence in Cognition and its Disorders, Sydney, NSW, Australia

* Corresponding author
(e) suncica.lah@sydney.edu.au
(f) +61 2 9351 7328
Postal Addresses:

Michelle Edwards

The Brain and Mind Centre, Level 2 (M02F)
94 Mallett St
Camperdown, NSW 2050, Australia
medw7458@uni.sydney.edu.au

Elizabeth Stewart

The Brain and Mind Centre, Level 2 (M02F)
94 Mallett St
Camperdown, NSW 2050, Australia
este7735@uni.sydney.edu.au

Associate Professor Romina Palermo

School of Psychological Science
The University of Western Australia (M304)
35 Stirling Highway
Crawley, WA 6009, Australia
romina.palermo@uwa.edu.au

Associate Professor Suncica Lah *

School of Psychology, Brennan McCallum (A18)
The University of Sydney, NSW, 2006, Australia
suncica.lah@sydney.edu.au
Abstract

Facial emotion perception is a fundamental social competency relying on a specialised, yet distributed, neural network. This review aimed to determine whether patients with epilepsy have facial emotion perception accuracy impairments overall, or for a subset of emotions (anger, disgust, happiness, sadness, fear, and surprise), and the relationship to epilepsy type, demographic/treatment variables, and brain organisation. Database searches used PRISMA guidelines with strict inclusion/exclusion criteria. Thirty included studies assessed patients with temporal lobe (TLE; $n = 709$), frontocentral (FCE; $n = 22$), and genetic generalised (GGE; $n = 48$) epilepsy. Large deficits emerged in patients with epilepsy compared to controls ($n = 746$; Hedges’ $g = 0.908 – 1.076$). Patients with TLE were significantly impaired on all emotions except surprise; patients with GGE were significantly impaired in anger, disgust, and fear perception. Meta-regression of patients with TLE revealed younger age at testing was associated with lower accuracy. This review provides evidence for marked global deficits of emotion perception in epilepsy, with differential emotion-specific impairment patterns in patients with TLE and GGE.

Keywords: facial emotion recognition, facial expression recognition, face, emotion, perception, epilepsy, seizure.
Statement of Funding Sources

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
1 Introduction

Facial emotion perception refers to the ability to discriminate an emotion expressed on a face (Palermo, O’Connor, Davis, Irons, & McKone, 2013), but has also been used to refer to the ability to ascribe an emotion label to a facial expression (Stewart, Catroppa, & Lah, 2016). In traditional models of face processing, it is distinct from the perception of other facial attributes, such as identity or gender (e.g., Bruce & Young, 1986; Haxby, Hoffman, & Gobbini, 2000) (although there may be some common processing; see Calder & Young, 2005; Rhodes et al., 2015). Perceiving the emotion in a face is critical for social cognition as it allows information about the emotional states of others to be captured and further processed via mirror neurons, which trigger physiological changes and impact social behaviour (Van Overwalle, 2009). Hence, facial emotion perception is foundational for social integration and influence (Szaflarski et al., 2014). Deficits in facial emotion perception are associated with increased psychological and interpersonal stress through a reduction in, or avoidance of, interactions with peers (Frith & Frith, 1999; Yeates et al., 2007).

Evolutionary (Darwin, 1872; Rolls, 1992), cross-cultural (Ekman & Friesen, 1971; Elfenbein & Ambady, 2002), and developmental (Lawrence, Campbell, & Skuse, 2015; Rolls, 1992; Tremblay, Kirouac, & Dore, 1987) evidence has indicated that there are six so-called “basic” facial emotions: anger, disgust, fear, happiness, sadness, and surprise. These emotions are said to have a largely biological origin, suggesting that expressions are innate and universally recognisable, rather than socially-influenced constructs (Izard, 1994; but see Barrett, 2006). The emotion of surprise may be an exception, as it is visually similar to fear (e.g., raised eyebrows, wide eyes) and ambiguously interpretable as either pleasantly or unpleasantly surprised (Fowler et al., 2006; Meletti et al., 2009). Each basic emotion appears to be represented by a relatively distinct facial expression; for example, happiness by an
upturned mouth and crescent-shaped eyes, or anger by tensed mouth, furrowed brow, and narrowed eyes (Ekman, 1993).

Functional imaging studies report distinct neural activation patterns for each facial emotion, where discrete brain regions are reported to be consistently activated in response to specific facial emotion expressions, such as the insula in disgust (Krolak-Salmon et al., 2003), or the amygdala in fear (Brierley, Medford, Shaw, & David, 2004; Fowler et al., 2006). However, a recent meta-analysis found limited evidence for a purely locationist hypothesis (Gur et al., 2002; Lindquist, Wager, Kober, Bliss-Moreau, & Barrett, 2012). Instead, evidence was more consistent with a psychological constructionist approach (Barrett & Bliss-Moreau, 2009; Calder & Young, 2005; Haxby et al., 2000), where a network of interacting regions involved in both emotional and non-emotional stimuli, are activated during facial emotion perception (Lindquist et al., 2012).

Developmental studies suggest divergent trajectories of facial emotion perception, which are moderated by gender and pubertal status (Lawrence et al., 2015). For example, while recognition accuracy of happy and fearful expressions demonstrate linear improvements as age increases, anger displays a non-linear trend, with a sharp improvement in accuracy from adolescence to adulthood (Thomas, De Bellis, Graham, & LaBar, 2007). Moreover, emotion perception appears to be a graduated process concurrent with brain development. Happiness, which is typically the first recognisable emotion, depends on the left amygdala and occipital lobes that mature earlier in development (Herba & Phillips, 2004). Anger, fear, sadness, and disgust, which are recognised later, rely on brain regions such as the left pulvinar, anterior insula, orbital frontal cortex, and medial frontal gyrus that continue maturing from childhood through to early adulthood (Thomas et al., 2007).

There are also small, but statistically significant sex differences in facial emotion perception from infancy, with girls better at recognising anger, disgust, happiness, and
surprise, but not fear or sadness, than boys (Lawrence et al., 2015; McClure, 2000; Thomas et al., 2007). Emotion intensity, also affects recognition, where facial emotions displayed at 100% or full-blown intensity are consistently more accurately recognised than those of lesser intensity, with the ability to recognise lower intensity displays improving with age (Ammerlaan, Hendriks, Colon, & Kessels, 2008; Gosselin, Peretz, Hasboun, Baulac, & Samson, 2011; Sedda et al., 2013). As most studies assess facial emotion perception using full-blown intensity, and morphing techniques vary between studies, then analysis of a consistent intensity allows for a direct comparison of facial emotion perception changes across development.

1.2 Facial Emotion Perception in Epilepsy

Facial emotion perception difficulties are common in patients with epilepsy. The term “epilepsy” describes a heterogeneous group of brain diseases, where normal neuronal activity is disturbed. It is characterised by recurrent, unprovoked seizures or epileptic discharges (Fisher et al., 2014), and is associated with considerable social morbidity (Kokkonen, Kokkonen, Saukkonen, & Pennanen, 1997; Morgan, Ahmed, & Kerr, 2000; Walpole, Isaac, & Reynders, 2008). Epilepsy can be focal or generalised. In focal epilepsies, such as temporal lobe epilepsy (TLE) or frontocentral epilepsy (FCE), seizures typically originate within one hemisphere and may be the result of acquired lesions, tumours, congenital structural abnormalities, or an unknown cause (Fisher et al., 2014). In patients with TLE, hippocampal sclerosis, where there is both neuronal loss and gliosis (Blümcke, Coras, Miyata, & Özkara, 2012), represents the most common pathology, at least in adults. Focal epilepsies can be further categorised by the side of seizure focus (i.e., left, right). Conversely, the seizure activity of generalised epilepsies, such as genetic generalised epilepsy (GGE), is more dispersed, with seizures occurring bilaterally in both hemispheres simultaneously, without an identifiable focus or structural abnormality (Marini, King, Archer, Newton, &
Berkovic, 2003). Overall, given that seizure focus may involve any region of the brain (Fisher et al., 2014), epilepsy provides an opportunity to study the functional architecture of facial emotion perception.

Epilepsy also allows examination of the effect of interruption to developing neural networks important for the maturation of facial emotion perception, as epilepsy often begins in childhood and persists into adulthood (Byars et al., 2014). It is possible that the onset of seizures at varying developmental stages differentially impact maturation of facial emotion perception, as speed and accuracy of emotion perception improve progressively throughout development and across brain regions (Herba & Phillips, 2004). However, relatively few studies have reliably examined the facial emotion perception processing speed of patients with epilepsy. Those studies typically use unlimited-time tasks, and suggest that lengthy or repeated presentations may encourage participants to use compensatory strategies that reduce power to detect more subtle deficits from generally reduced speed of processing and attention (e.g., Graham, Devinsky, & LaBar, 2007; Hennion et al., 2015).

Epilepsy also enables neural plasticity of emotion perception to be examined through studies of patients who undergo surgical treatment for intractable epilepsy. Assuming emotion perception is mediated by distributed brain regions, and the brain holds potential for neural plasticity, seizure freedom following surgery may result in improved facial emotion perception (see Lindquist et al., 2012; Monti & Meletti, 2015 for similar suggestions). In contrast, assuming that discrete brain regions are critical for the perception of specific facial emotions, then resection of particular brain regions may induce selective deficits following surgery, such as deficits in fear recognition following amygdala resection moderated by expression intensity (Palermo, Schmalzl, Mohamed, Bleasel, & Miller, 2010). Finally, some antiepileptic drugs (e.g., phenobarbital) are known to hinder facial emotion perception (Meletti et al., 2009).
Tasks used to assess facial emotion perception abilities typically employ one of two designs: (i) identification, where participants are required to assign a qualitative label to an image of a facial expression, usually from a small list of descriptors (e.g., Hennion et al., 2015), or (ii) differentiation, where participants judge differences of expression without requiring explicit emotion identification (e.g., Hlobil, Rathore, Alexander, Sarma, & Radhakrishnan, 2008).

The impact of epilepsy variables on facial emotion perception globally or for specific emotions remains unclear despite research. Evidence is mixed regarding the impact of site of epilepsy focus. While most research focuses on facial emotion perception deficits arising from TLE, the impact of side of focus is still contended, with researchers arguing the distinct roles of left and right hemispheres. The functional architecture of facial emotion perception is also debated as being underlain by specialised brain regions versus a dispersed network. This has implications for counselling patients in relation to a risk of a decline or a likelihood of an improvement in facial emotion perception following surgery, where evidence of its effects on facial emotion perception remains unclear. Clarification of the impact of epilepsy factors on facial emotion perception will assist in the early detection of patients at risk, and management of deficits to reduce the considerable social morbidity typically found in patients with epilepsy.

The primary aim of this review and meta-analysis is to determine whether facial emotion perception is impaired in epilepsy, and if so, whether the impairment is global or emotion specific. The secondary aim is to establish whether the severity and pattern of impairment is moderated by epilepsy variables (i.e., age of onset, duration of epilepsy disorder, side of seizure focus, and seizure frequency), treatment factors (i.e., pharmacological, surgical), and demographic factors (i.e., age at testing, gender). In addition,
we wanted to examine neural correlates of facial emotion perception in patients with epilepsy.

2 Method

2.1 Registration

The review was registered with PROSPERO on 5 November 2015. Initially, the review planned to address facial emotion perception in children with epilepsy only. However, due to the limited number of studies, the review was extended to include adults. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement and guidelines (Moher, Liberati, Tetzlaff, & Altman, 2009) were used to develop and conduct searches, summarise evidence, and report results.

2.2 Identification of Studies

2.2.1 Search strategy.

Databases including PsycINFO, PubMed/MEDLINE, Embase, Scopus, Web of Science, and the Cochrane Library were searched on 8 December 2016. The following terms were used: [(face OR facial) AND (affect* OR emotion* OR expression*) AND (perce* OR identif* OR recogni* OR process*)] AND [(epilepsy OR epilep* OR seizure* OR (epileptic AND seizure) OR convulsion)]. All Medical Subject Heading (MeSH) terms were exploded. Studies were limited to peer-reviewed journal articles. No date, language, or age limits were set in any of the database searches. Using the Ancestry Method, the reference lists of appropriate reviews and empirical studies were also searched to identify papers that were not indexed in these databases.

2.2.2 Study selection.

2.2.2.1 Inclusion criteria.

The review included studies that: (i) reported original, empirical research published in peer-reviewed journals; (ii) included people with an epilepsy diagnosis; (iii) reported data
separately from other patient groups included in the study; (iv) included a control group, or compared results to normative data; (v) used a discrimination or identification task to assess accuracy of facial emotion perception; (vi) involved human participants; and (vii) were published in the English language.

2.2.2.2 Exclusion criteria.

Studies were excluded if they: (i) aggregated results of people with epilepsy and other patient groups; (ii) did not include a control group or normative data; (iii) were single case studies, reviews, meta-analyses, editorials, or conference proceedings (iv) included facial stimuli, but did not examine facial emotion perception accuracy (e.g., eye-gaze or other facial attributes, familiar faces, famous faces, recently present faces, etc.), and (v) employed facial emotion perception tasks, but did not report data needed to calculate a mean weighted effect for facial emotion perception accuracy, such as means, standard deviations, $t$-values or $p$-values. Studies were not excluded on the basis of other premorbid conditions, origin of epilepsy (e.g., acquired lesions, tumours, congenital structural abnormalities), or medication use.

2.3 Selection

All titles and abstracts were screened by one reviewer (ME). A second independent reviewer (ES) screened a random selection of 25% of the titles and abstracts to ensure inter-rater reliability, with agreement on 96% of this subset. Discrepancies were resolved with discussion, until a concordance rate of 100% was reached. Full-text manuscripts of papers that either met selection criteria, or could not be determined to meet full selection criteria through title and abstract screening, were obtained. The reference lists of articles that met the inclusion criteria were examined for studies not identified in the main search. One author (ME) reviewed all full-text papers.

2.4 Data Extraction
Prior to conducting the meta-analyses, the following were extracted from each paper (by ME):

1. Name of the first author and the year of publication.
2. Number of participants.
3. Means of background variables (age at assessment, sex, education), epilepsy variables (type, age of onset, duration, seizure frequency), and treatment (surgery; number and type of antiepileptic drugs).
4. Facial emotion perception task type (identification, differentiation)
5. Means and standard deviations (SDs) of each group on facial emotion perception tasks. Where means and SDs were not available, $t$-values were obtained, followed by $p$-values.

Authors were contacted via email if papers did not report necessary statistical or demographic information required for inclusion in the meta-analysis. Those who did not respond within 4 weeks were excluded from the meta-analysis ($n = 6$), but included in the systematic review.

2.5 Data Analysis

Data was analysed using the Comprehensive Meta Analyses Program, Version 3 (CMA3; Borenstein, Hedges, Higgins, & Rothstein, 2014). Standardised differences between epilepsy and control groups were calculated using Hedges’ $g$, which also corrected for effect size overestimation due to small sample sizes (Borenstein, Hedges, Higgins, & Rothstein, 2009). Hedges’ $g$ was interpreted similarly to Cohen’s $d$, where 0.2 indicated a small effect, 0.5 a medium effect, and 0.8 a large effect (Cohen, 1988). Positive effect sizes indicated that patients with epilepsy performed more poorly than controls. Facial emotion perception effect sizes by task type (differentiation and identification) were compared using the $Q$-test. Two-tailed tests with $p < 0.05$ were used in all analyses.
The primary meta-analysis examined (i) overall facial emotion perception accuracy between and within epilepsy groups (TLE, FCE, GGE, and unspecified epilepsy) across all emotions, and (ii) for specific emotions within epilepsy groups, separately. Because some studies included more than one epilepsy group (e.g., TLE and FCE) and compared patients to a single control group, epilepsy groups were not directly compared to one another. Similarly, studies with more than one epilepsy group included in the overall analysis used the primary epilepsy group, rather than the clinical comparison epilepsy group to avoid a unit-of-analysis error that arises with multiple correlated comparisons (Higgins & Green, 2008). The secondary analysis examined whether effect sizes differed with respect to side of seizure focus and surgical status. A random effects model was used, and the homogeneity of mean weighted effect sizes across analyses were tested with the $Q$-test. Meta-regression for continuous moderator variables assessed whether age of seizure onset, duration of illness, seizure frequency, antiepileptic drugs, and age at testing were related to facial emotion perception in each epilepsy group.

Meta-regression could not be used to analyse non-continuous moderator variables, including site and side of pathology, surgical status, type of medication, or gender of participants, due to an overlap of control groups within studies. These variables were instead reviewed systematically.

Publication bias was tested with funnel plots and Egger’s regression test (Egger, Smith, Schneider, & Minder, 1997). Tasks with a significant asymmetry (Egger’s test, $p < 0.05$) were further analysed to reduce the chance of Type I error. Individual study characteristics were investigated, and Rosenthal’s (1979) failsafe number was calculated to identify the number of negative studies required to nullify a significant between-group difference.

2.5.1 Quality ratings of selected papers.
The methodological quality of included studies was evaluated using the Downs and Black Checklist (1998). The checklist demonstrated sound psychometric properties for both randomised controlled trials and non-randomised studies, including high internal consistency (Kuder-Richardson formula-20 = .89), high test-retest reliability (Spearman’s r = .88), and good inter-rater reliability (Spearman’s r = .75; Downs & Black, 1998). An adapted version of the Down’s and Black Checklist was used, which excluded 10 items from the original checklist that related to interventional trials, as no intervention studies were included in this review. This modified checklist included five subscales assessing reporting quality (items 1 to 8), external validity (item 9), internal validity (statistical and methodological bias, items 10 to 12; selection bias, items 13 to 16), and power (item 17). All items were scored as 0 (no, or unable to determine), or 1 (yes), except for item 4, which was scored 0 (no, or unable to determine), 1 (partially), or 2 (yes). Items 7 and 16 were specific to participant attrition in longitudinal studies. Therefore, longitudinal studies received a quality rating of 0 to 18 points, while cross-sectional studies received a rating of 0 to 16 points. Papers were categorised as having (i) a high (0 to 5 points), average (6 to 11 points), or low (12 to 18 points) risk of bias for longitudinal studies, and (ii) a high (0 to 5 points), average (6 to 10 points), or low (11 to 16 points) risk of bias for cross-sectional studies. Two reviewers (ME, ES) rated each paper using the modified checklist. Inter-rater reliability between items ranged from 90 to 100%. Rating discrepancies were resolved through discussion.

3 Results

Please see Figure 1 for details of study selection process. The initial search retrieved 1804 articles, with 993 articles remaining after duplicates were removed. Of the 77 full text papers that were reviewed, 47 were excluded as these papers: (i) did not involve a task of facial emotion perception (n = 19); (ii) included complex emotions or social scenes that did not focus specifically on facial emotion perception (n = 17); (iii) were case studies (n = 7);
(iv) involved samples with no formal diagnosis of epilepsy \((n = 1)\); or (v) were conference proceedings \((n = 3)\). Thirty studies remained. Six studies were only included in the systematic review as insufficient information was available within the paper, or from the author for inclusion in the meta-analysis.

3.1 Study Characteristics and Patient Demographics

Of the 24 studies included in the meta-analysis, 21 were cross-sectional, and three were longitudinal. Twenty-two studies included adults with epilepsy \((n = 771)\), and two
studies included children with epilepsy \((n = 49)\). All studies included a group of healthy controls with no history of seizures, epilepsy, or neurological or psychiatric disorders.

Overall, the studies included 820 patients with epilepsy and 746 controls. Eighteen studies involved patients with TLE \((n = 709)\), two studies included patients with FCE \((n = 22)\), three studies examined patients with GGE \((n = 48)\), and in one study, epilepsy type was unspecified \((n = 41)\).

The mean age of epilepsy onset ranged from 5.80 to 30.81 years (TLE), 5.80 to 11.80 years (FCE), 11.80 to 15.14 years (GGE), and 5.90 years (unspecified epilepsy). The mean duration of epilepsy ranged from 13.00 to 32.38 years (TLE), 4.70 years (FCE; mean duration of epilepsy not reported by Farrant et al., 2005)), 13.50 to 21.10 years (GGE), and 5.60 years (unspecified epilepsy). The mean age of patients at testing ranged from 30.51 to 45.31 years (TLE), 12.60 to 34.36 years (FCE), 26.30 to 32.70 years (GGE), and 13.50 years (unspecified epilepsy).

A summary of characteristics and patient variables of studies included in the meta-analysis can be found in Table 1.

### 3.2 Tasks Used

Facial emotion perception accuracy was assessed with either an identification \((n = 22)\) or differentiation \((n = 2)\) task. Across studies, meta-analysis revealed large effect sizes for both identification \((g = 1.008, 95\% \text{ CI } 0.862 - 1.153, z = 13.547, p < 0.001)\) and differentiation \((g = 0.1925, 95\% \text{ CI } 0.282 - 1.768, z = 2.703, p = 0.007)\) tasks across epilepsies, which did not differ significantly from one another \((Q = 1.358, df = 1, p = 0.244)\). Thus, identification and differentiation tasks were collapsed into a single variable, referred to as facial emotion perception tasks, for the rest of this study.
Table 1. Characteristics and patient demographics of studies included in the meta-analysis.

<table>
<thead>
<tr>
<th>Study Variables</th>
<th>Background Variables</th>
<th>Epilepsy Variables</th>
<th>Treatment Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Study Design</td>
<td>Age Group</td>
<td>Type</td>
</tr>
<tr>
<td>Amlerova et al., (2014)</td>
<td>Long</td>
<td>Adult</td>
<td>TLE</td>
</tr>
<tr>
<td>Ammerlaan et al., (2008)</td>
<td>CS</td>
<td>Adult</td>
<td>TLE</td>
</tr>
<tr>
<td>Boucher et al., (2015)</td>
<td>CS</td>
<td>Adult</td>
<td>TLE</td>
</tr>
<tr>
<td>Braams et al., (2015)</td>
<td>Long</td>
<td>Child</td>
<td>Unspec.</td>
</tr>
<tr>
<td>Brierley et al., (2004)</td>
<td>CS</td>
<td>Adult</td>
<td>TLE</td>
</tr>
<tr>
<td>Broicher et al., (2012)</td>
<td>CS</td>
<td>Adult</td>
<td>TLE</td>
</tr>
<tr>
<td>Carvajal et al., (2009)</td>
<td>CS</td>
<td>Adult</td>
<td>TLE</td>
</tr>
<tr>
<td>Farrant et al., (2005)</td>
<td>CS</td>
<td>Adult</td>
<td>FCE</td>
</tr>
<tr>
<td>Golouboff et al., (2008)</td>
<td>CS</td>
<td>Child</td>
<td>FCE</td>
</tr>
<tr>
<td>Gomez-Ibanez et al., (2014)</td>
<td>CS</td>
<td>Adult</td>
<td>GGE</td>
</tr>
<tr>
<td>Hennion et al., (2015)</td>
<td>CS</td>
<td>Adult</td>
<td>TLE</td>
</tr>
<tr>
<td>Study</td>
<td>Design</td>
<td>Age</td>
<td>Type</td>
</tr>
<tr>
<td>------------------------------</td>
<td>--------</td>
<td>-----</td>
<td>------</td>
</tr>
<tr>
<td>Hlobil et al., (2008)</td>
<td>CS</td>
<td>Adult</td>
<td>TLE</td>
</tr>
<tr>
<td>McClelland et al., (2006)</td>
<td>CS</td>
<td>Adult</td>
<td>TLE</td>
</tr>
<tr>
<td>Meletti et al., (2003b)</td>
<td>CS</td>
<td>Adult</td>
<td>TLE</td>
</tr>
<tr>
<td>Meletti et al., (2009)</td>
<td>CS</td>
<td>Adult</td>
<td>TLE</td>
</tr>
<tr>
<td>Meletti et al., (2014)</td>
<td>CS</td>
<td>Adult</td>
<td>TLE</td>
</tr>
<tr>
<td>Reyners et al., (2005)</td>
<td>CS</td>
<td>Adult</td>
<td>GGE</td>
</tr>
<tr>
<td>Sedda et al., (2013)</td>
<td>CS</td>
<td>Adult</td>
<td>TLE</td>
</tr>
<tr>
<td>Shaw et al., (2007)</td>
<td>Long</td>
<td>Adult</td>
<td>TLE</td>
</tr>
<tr>
<td>Szafarski et al., (2014)</td>
<td>CS</td>
<td>Adult</td>
<td>TLE</td>
</tr>
<tr>
<td>Tanaka et al., (2013)</td>
<td>CS</td>
<td>Adult</td>
<td>TLE</td>
</tr>
<tr>
<td>Walpole et al., (2008)</td>
<td>CS</td>
<td>Adult</td>
<td>TLE</td>
</tr>
</tbody>
</table>

AEDs: anti-epileptic drugs; CS: cross-sectional study; FCE: frontocentral epilepsy; GGE: genetic generalised epilepsy; Long: longitudinal study; n: number of participants with epilepsy; n<sub>c</sub>: number of control participants; TLE: temporal lobe epilepsy; Unspec.: unspecified (mixed sample).
Fourteen of the 24 studies included in the meta-analysis also included a non-diagnostic face identity perception control task to screen for prosopagnosia (difficulty recognising face identity). Deficits in face perception were not indicated in any of these 14 studies.

3.3 Meta-Analysis

Mixed effects meta-analyses (see Figures 2 and 3, and Table 2) demonstrated that compared to controls, patients with epilepsy (collapsed across groups) were less accurate in face emotion perception (collapsed across emotions) \((n = 20, g = 0.990, 95\% \text{ CI } 0.846 - 1.133, z = 13.540, p < 0.001; \text{ see Figure 2})\). A \(Q\)-test analysis revealed no evidence of heterogeneity between epilepsy groups (TLE, FCE, GGE, and unspecified epilepsy; \(Q = 0.352, df = 3, p = 0.950\)). Moreover, people with epilepsy obtained significantly lower accuracy scores relative to controls for each and every facial emotion, namely: anger, disgust, fear, happiness, sadness, and surprise.
Figure 2. Individual and mean weighted effect sizes (Hedges’ $g$ and 95% CIs) for facial emotion perception studies based on accuracy of responding (collapsed across emotions) for patients with temporal lobe epilepsy (TLE), frontocentral epilepsy (FCE), genetic generalised epilepsy (GGE), and unspecified epilepsy.
Figure 3. Individual and mean weighted effect sizes (Hedges’ $g$ and 95% CIs) for facial emotion perception accuracy collapsed across epilepsy groups by emotion.
<table>
<thead>
<tr>
<th>Study</th>
<th>n</th>
<th>Anger</th>
<th>Disgust</th>
<th>Fear</th>
<th>Happiness</th>
<th>Sadness</th>
<th>Surprise</th>
<th>Neutral</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TLE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ammerlaan et al., (2008)</td>
<td>9</td>
<td>0.179 (0.413)</td>
<td>0.944 (0.435)*</td>
<td>1.364 (0.550)*</td>
<td>0.564 (0.420)</td>
<td>0.028 (0.449)</td>
<td>0.578 (0.435)</td>
<td>NM</td>
</tr>
<tr>
<td>Banks et al., (2014)</td>
<td>8</td>
<td>#</td>
<td>#</td>
<td>#</td>
<td>#</td>
<td>#</td>
<td>#</td>
<td>1.008 (0.425)*</td>
</tr>
<tr>
<td>Bonora et al., (2011)</td>
<td>41</td>
<td>7.569 (0.599)*</td>
<td>4.645 (0.403)*</td>
<td>5.860 (0.482)*</td>
<td>3.308 (0.322)*</td>
<td>4.845 (0.415)*</td>
<td>NM</td>
<td>NM</td>
</tr>
<tr>
<td>Boucher et al., (2015)</td>
<td>15</td>
<td>0.457 (0.338)</td>
<td>0.318 (0.336)</td>
<td>0.749 (0.346)*</td>
<td>0.617 (0.342)</td>
<td>0.145 (0.334)</td>
<td>0.583 (0.341)</td>
<td>NM</td>
</tr>
<tr>
<td>Brierley et al., (2004)</td>
<td>25</td>
<td>1.080 (0.282)*</td>
<td>0.498 (0.267)</td>
<td>0.516 (0.268)</td>
<td>0.377 (0.266)</td>
<td>0.411 (0.266)</td>
<td>NM</td>
<td>NM</td>
</tr>
<tr>
<td>Broicher et al., (2012)</td>
<td>28</td>
<td>0.207 (0.262)</td>
<td>0.733 (0.270)*</td>
<td>0.902 (0.275)*</td>
<td>#</td>
<td>0.630 (0.268)*</td>
<td>0.099 (0.261)</td>
<td>#</td>
</tr>
<tr>
<td>Gosselin et al., (2011)</td>
<td>16</td>
<td>0.183 (0.345)</td>
<td>0.000 (0.345)</td>
<td>0.813 (0.359)*</td>
<td>0.183 (0.345)</td>
<td>0.000 (0.345)</td>
<td>0.000 (0.345)</td>
<td>NM</td>
</tr>
<tr>
<td>Hennion et al., (2015)</td>
<td>50</td>
<td>0.155 (0.199)</td>
<td>0.569 (0.203)*</td>
<td>1.116 (0.214)*</td>
<td>0.080 (0.199)</td>
<td>0.457 (0.201)*</td>
<td>NM</td>
<td>0.851 (0.207)*</td>
</tr>
<tr>
<td>Hlobil et al., (2008)</td>
<td>76</td>
<td>0.130 (0.244)</td>
<td>NM</td>
<td>0.976 (0.230)*</td>
<td>0.003 (0.244)</td>
<td>NM</td>
<td>NM</td>
<td>NM</td>
</tr>
<tr>
<td>Meletti et al., (2009)</td>
<td>140</td>
<td>0.642 (0.167)*</td>
<td>0.608 (0.167)*</td>
<td>0.589 (0.167)*</td>
<td>0.356 (0.165)*</td>
<td>0.636 (0.167)*</td>
<td>NM</td>
<td>NM</td>
</tr>
<tr>
<td>Meletti et al., (2014)</td>
<td>42</td>
<td>1.453 (0.248)*</td>
<td>0.702 (0.227)</td>
<td>0.736 (0.228)*</td>
<td>0.458 (0.223)*</td>
<td>0.643 (0.226)*</td>
<td>NM</td>
<td>NM</td>
</tr>
<tr>
<td>Sedda et al., (2013)</td>
<td>57</td>
<td>0.321 (0.190)</td>
<td>1.285 (0.207)*</td>
<td>0.904 (0.198)*</td>
<td>0.620 (0.193)*</td>
<td>1.209 (0.205)*</td>
<td>NM</td>
<td>NM</td>
</tr>
<tr>
<td>Shaw et al., (2007)</td>
<td>19</td>
<td>1.132 (0.343)*</td>
<td>1.118 (0.343)*</td>
<td>0.351 (0.320)</td>
<td>0.136 (0.318)</td>
<td>2.234 (0.408)*</td>
<td>2.120 (0.400)*</td>
<td>NM</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td></td>
<td>526</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TLE Overall</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.012 (0.298)*</td>
<td>1.008 (0.263)*</td>
<td>1.166 (0.250)*</td>
<td>0.595 (0.227)*</td>
<td>0.991 (0.290)*</td>
<td>0.650 (0.357)</td>
<td>0.881 (0.186)*</td>
<td></td>
</tr>
<tr>
<td><strong>FCE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Golouboff et al., (2008)</td>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.119 (0.383)</td>
<td>0.524 (0.387)</td>
<td>0.259 (0.384)</td>
<td>1.247 (0.405)*</td>
<td>0.541 (0.387)</td>
<td>NM</td>
<td>0.418 (0.386)</td>
<td></td>
</tr>
<tr>
<td><strong>GGE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gomaz-Ibanez et al.,</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2014)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.890 (0.315)*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Realmuto et al., (2015)</td>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.007 (0.335)*</td>
<td>0.409 (0.318)</td>
<td>0.730 (0.325)*</td>
<td>0.059 (0.315)</td>
<td>0.013 (0.315)</td>
<td>0.044 (0.315)</td>
<td>NM</td>
<td></td>
</tr>
<tr>
<td>Reynders et al., (2005)</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.226 (0.452)*</td>
<td>0.272 (0.414)</td>
<td>1.223 (0.451)*</td>
<td>0.000 (0.412)</td>
<td>0.437 (0.417)</td>
<td>0.321 (0.415)</td>
<td>NM</td>
<td></td>
</tr>
<tr>
<td><strong>GGE Overall</strong></td>
<td>48</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.085 (0.269)*</td>
<td>0.566 (0.197)*</td>
<td>0.898 (0.264)*</td>
<td>0.037 (0.250)</td>
<td>0.167 (0.251)</td>
<td>0.145 (0.251)</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td><strong>Overall Sample</strong>*</td>
<td>654</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.630 (0.130)*</td>
<td>0.698 (0.117)*</td>
<td>0.598 (0.124)*</td>
<td>0.427 (0.121)*</td>
<td>0.646 (0.129)*</td>
<td>0.543 (0.142)*</td>
<td>0.793 (0.168)*</td>
<td></td>
</tr>
</tbody>
</table>

### Notes:

- Overall effect size for each emotion adjusted to account for weighted effect by epilepsy group when analysed together, in studies that included both right and left hemispheric epileptogenic zones in analyses of patients with TLE, FCE, and generalised epilepsies.
- #: Not enough data provided to calculate an effect size; unable to obtain data from author.
- NM: Not measured.
- * p < 0.05
- TLE: temporal lobe epilepsy; FCE: frontocentral epilepsy; GGE: genetic generalised epilepsy.
3.4 Meta-Analyses by Clinical Epilepsy Variables

3.4.1 Temporal lobe epilepsy.

Patients with TLE were less accurate on facial emotion perception tasks compared to controls (see Figure 2). When performance was examined separately for each emotion, patients with TLE obtained lower accuracy scores for angry, disgusted, fearful, happy, and sad facial expressions, but not for surprised facial expressions (see Table 2).

Next, analyses were performed to determine whether the side of epilepsy focus (left, right, bilateral) impacted perception in patients with TLE. Patients with left-TLE, right-TLE, and bilateral-TLE performed significantly below controls (see Figure 4). Effect sizes for left-TLE (n = 11, g = 0.867, 95% CI 0.655 – 1.080, z = 8.011, p < 0.001), right-TLE (n = 11, g = 0.882, 95% CI 0.579 = 1.185, z = 5.708, p < 0.001), and bilateral-TLE (n = 2, g = 2.410, 95% CI 0.172 – 4.647, z = 2.111, p = 0.035) were large, and did not differ significantly from each other (Q = 1.809, df = 2, p = 0.405).

Finally, patients with TLE were investigated as to whether surgical status (pre-surgical studies [n = 5]; post-surgical studies [n = 7]) impacted accuracy. Temporal lobe surgery involved excision of the medial and antero-medial regions (n = 175, 7 studies). Both pre- and post-temporal lobe surgery patients performed significantly below controls (see Figure 5). Effect sizes of pre-surgical (n = 5, g = 0.876, 95% CI 0.649 – 1.103, z = 7.569, p < 0.001) and post-surgical (n = 7, g = 1.135, 95% CI 0.860 – 1.409, z = 8.108, p < 0.001) groups were large, and comparable in size (Q = 2.879, df = 1, p = 0.090).

Meta-regression of studies involving patients with TLE showed that associations between effect sizes and moderator variables (age of seizure onset, disease duration, mean seizure frequency per month, mean number of antiepileptic drugs, and/or surgical status) were non-significant (all p > 0.05) except for one variable. A significant positive association
was found between effect size and age at testing ($R = 0.04$, 95% CI $0.004 – 0.084$, $z = 2.140$, $p = 0.032$), with younger age at testing associated with lower accuracy.

Meta-regression was undertaken only for patients with TLE, as the minimum number of studies needed to examine the impact of moderators of facial emotion perception ($n \geq 3$) was insufficient for FCE and GGE (Borenstein et al., 2009).

Figure 4. Individual and mean weighted effect sizes (Hedges' $g$ and 95% CIs) of facial emotion perception accuracy (collapsed across emotions) for temporal lobe epilepsy (TLE) patients with left, right, or bilateral seizure focus.
3.4.2 Frontocentral epilepsy.

The total facial emotion perception accuracy score of patients with FCE and controls did not differ significantly (see Figure 2, and Appendix A). As emotion-specific data was obtained for only one study, meta-analyses by emotion type could not be performed.

3.4.3 Genetic generalised epilepsy.

Total facial emotion perception accuracy score obtained by patients with GGE was significantly lower than controls (see Figure 2, and Appendix A).

When facial emotion perception was examined for specific emotions, patients with GGE were impaired for anger, disgust, and fear. In contrast, no impairments were found for happiness, sadness, or surprise (see Table 2).

3.5 Systematic Review of Moderator Variables Not Included in Meta-Analyses

3.5.1 Hippocampal pathology in TLE.
Three studies (Hennion et al., 2015; Meletti et al., 2009; Tanaka et al., 2013) examined the impact of hippocampal pathology on facial emotion perception in patients with TLE with mixed results. Two studies (Meletti et al., 2009; Tanaka et al., 2013) found no difference in facial emotion perception scores for patients with hippocampal sclerosis, compared to patients without evidence of hippocampal sclerosis when moderator variables (e.g., age of epilepsy onset, epilepsy duration, age at testing, gender, education) were controlled. In contrast, Hennion and colleagues (2015) found patients with left-TLE with hippocampal sclerosis had significantly greater impairments in recognition of fearful facial expressions compared to patients with left-TLE without hippocampal sclerosis ($p = 0.0076$). In the same study, no difference was found in facial emotion perception accuracy of any emotion between patients with right-TLE, with or without hippocampal sclerosis.

### 3.5.2 Medication.

Only one study by Meletti and colleagues (2009) examined the relationship between antiepileptic drug type and facial emotion perception. They found that patients with TLE who were prescribed phenobarbital had significantly reduced facial emotion perception accuracy compared to patients with TLE who were not prescribed phenobarbital ($p = 0.003$).

### 3.5.3 Functional brain organisation of facial emotion perception.

Two studies (Meletti, Benuzzi, Nichelli, & Tassinari, 2003a; Szaflarski et al., 2014) examined neural correlates of facial emotion perception in patients with epilepsy. One study (Meletti et al., 2003a) observed similar activation patterns, but with overall reduced blood oxygenation level dependent (BOLD) activation, for left-TLE patients compared to controls. However, activation was also increased uniquely to left-TLE patients in the right cerebellum, hippocampus, and parahippocampus for happy expressions, while the right occipital gyrus was activated for fearful faces. In direct contrast with the findings of Meletti and colleagues (2003a), another study (Szaflarski et al., 2014) found a decrease in BOLD responses to the
hippocampus, parahippocampus, and cerebellum for happy expressions in patients with left-TLE, compared to controls.

### 3.5.4 Gender.

Four studies (Braams et al., 2015; Hlobil et al., 2008; Meletti et al., 2009; Tanaka et al., 2013) reported no difference in facial emotion perception accuracy with respect to the gender of patients.

### 3.5.5 Children with epilepsy.

Only three of thirty studies examined facial emotion perception in children with epilepsy (Braams et al., 2015; Golouboff et al., 2008; Pinabiaux et al., 2013). Braams and colleagues (2015) tested facial emotion perception in children who underwent epilepsy surgery ($n = 41$) at 24 months post-surgery. Children who underwent epilepsy surgery were less accurate in facial emotion perception than controls overall, and on recognition of several, but not all, specific emotions (sadness, disgust, and surprise). A small sub-group of children with epilepsy (11 of 41), dichotimised by age at surgery, were followed-up longitudinally from pre-surgery to, and reviewed at six, 12, and 24 months post-surgery. A subset of control children was also followed-up longitudinally. Control children showed an increase in facial emotion perception accuracy over the follow-up period. In contrast, children who underwent surgery at an earlier age ($<12.1$ years) showed a decline in facial emotion perception accuracy (especially for surprise, anger, disgust and happiness, but not for fear and sadness) from pre- to 6-months post-surgery. Nevertheless, the facial emotion perception accuracy of these children returned to, but did not surpass, pre-surgical accuracy at 24-months post-surgery, and remained below that of the control group. In contrast, children who underwent surgery at an older age showed a significant increase in facial emotion perception accuracy at all follow-up time points. This increase in facial emotion perception accuracy was comparable to (but not greater than), the increase in facial emotion perception accuracy in the
control group. Hence, facial emotion perception deficits did not resolve, but persisted post-
surgery, even in the group operated on at an older age.

The second study (Pinabiaux et al., 2013) compared facial emotion perception of
children with TLE who had undergone surgery, to healthy controls. Controls identified
happy, fearful, and neutral expressions more accurately than children with TLE. Children
with left-TLE did not differ from children with right-TLE in facial emotion perception.

Finally, Golouboff and colleagues (2008) examined facial emotion perception in
children with early-onset (5 to 7 years) TLE or FCE, either pre- or post- surgery. Overall,
children with left-TLE were impaired in facial emotion perception compared to controls.
Post-hoc comparisons found children with left-TLE were impaired in recognition of fearful
and neutral faces, children with right-TLE were impaired in recognition of disgust, and
children with FCE were impaired in recognition of happiness, relative to controls.

In children, no moderator variables, including age at testing, duration of epilepsy, side
or site of surgery, or number of antiepileptic drugs were associated with facial emotion
perception accuracy, except for age of epilepsy onset in two of the three studies. Golouboff
and colleagues (2008) found earlier age of first temporal lobe seizure was related to lower
facial emotion perception accuracy. Pinabiaux and colleagues (2013) found earlier epilepsy
onset was related to reduced fear recognition. In contrast, no relationship between age of
onset and facial emotion perception accuracy was found by Braams and colleagues (2015).

3.6 Publication Bias

No significant publication bias was detected with Egger’s regression test ($p = 0.876,
2-tailed). The number of non-significant studies required to nullify the effect was estimated at
1275 studies using the classic failsafe N. This was greatly in excess of Rosenthal’s (1979)
“$5k + 10$” criteria, where $k$ is the number of studies included in the meta-analysis. Funnel
plots were also visually inspected, with no evidence of publication bias.
3.7 Quality Ratings

The Downs and Black (1998) Checklist assessed the overall risk of bias across all studies as moderate to low (see Table 3). Cross-sectional \((n = 27)\) and longitudinal \((n = 3)\) study ratings ranged from 7 to 15 out of 16 \((m = 12.5)\), and 12 to 13 out of 18 \((m = 12.7)\), respectively. Reporting quality was adequate; five studies received perfect scores. Shortcomings related to incomplete reporting included omitting measures of variance (Brierley et al., 2004) or exact \(p\)-values (Carvajal, Rubio, Martín, Serrano, & García-Sola, 2009; Meletti et al., 2003a; Meletti, Benuzzi, Rubboli, Cantalupo, Stanzani Maserati, et al., 2003b; Pinabiaux et al., 2013). No longitudinal studies reported on the characteristics of patients lost to follow-up. One-third of studies adequately reported external validity (e.g., consecutive participant selection). Internal validity, and statistical and methodological bias were reported well, with all studies receiving perfect scores, except for two studies (Banks, Bellerose, Douglas, & Jones-Gotman, 2013; Meletti et al., 2003a). In contrast, only one study (Brierley et al., 2004) reported fully on selection bias criteria for population and time period of recruitment, and appropriate adjustments for confounding factors. All studies appeared to have sufficient power.
Table 3. Quality assessment: Adapted from Downs and Black (1998) checklist.

<table>
<thead>
<tr>
<th>Studies</th>
<th>Quality Assessment Criteria from the Adapted Downs and Black (1998) Checklist</th>
<th>Quantity of Reporting</th>
<th>External Validity</th>
<th>Internal Validity</th>
<th>Statistical and Methodological Bias</th>
<th>Selection Bias</th>
<th>Power</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Quality of Reporting</td>
<td>1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 Sum</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amlerova et al., (2014)</td>
<td>1 1 1 1 1 1 1 0 1 0 1 1 1 1 1 0 0 1 13/18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ammerlaan et al., (2008)</td>
<td>1 1 1 2 1 1 NA 1 1 1 0 1 1 1 1 0 0 1 NA 1 13/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Banks et al., (2014)</td>
<td>1 1 1 1 1 1 1 NA 1 0 0 1 1 0 0 0 0 1 NA 1 10/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bonora et al., (2011)</td>
<td>1 1 1 1 1 1 1 NA 1 1 1 1 1 0 0 1 NA 1 12/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boucher et al., (2015)</td>
<td>1 1 1 2 1 1 NA 1 1 1 1 1 1 0 1 NA 1 15/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Braams et al., (2015)</td>
<td>1 1 1 1 1 1 1 0 1 1 1 1 1 0 0 1 0 1 NA 1 13/18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Brierley et al., (2004)</td>
<td>1 1 1 1 1 1 1 NA 1 0 1 1 1 1 1 1 1 NA 1 14/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Broicher et al., (2012)</td>
<td>1 1 1 2 1 1 NA 1 1 1 1 1 1 0 1 NA 1 15/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Carvajal et al., (2009)</td>
<td>1 1 1 2 1 1 NA 0 1 1 1 1 1 0 0 0 NA 1 13/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farrant et al., (2005)</td>
<td>1 1 1 1 1 1 NA 1 0 1 1 1 0 0 1 NA 1 12/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fowler et al., (2006)</td>
<td>1 1 1 1 1 1 NA 1 0 1 1 1 1 0 0 1 NA 1 13/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Golouboff et al., (2008)</td>
<td>1 1 1 1 1 1 NA 1 0 1 1 1 0 0 1 NA 1 12/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gomez-Ibanez et al., (2014)</td>
<td>1 1 1 1 1 1 0 NA 1 0 1 1 1 1 0 1 NA 1 12/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gosselin et al., (2011)</td>
<td>1 1 1 1 1 1 NA 1 0 1 1 1 0 0 1 NA 1 12/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hennion et al., (2015)</td>
<td>1 1 1 2 1 1 NA 1 1 1 1 1 1 0 0 1 NA 1 14/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hlobil et al., (2008)</td>
<td>1 1 1 2 1 1 NA 1 1 1 1 1 1 0 1 NA 1 15/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McClelland et al., (2006)</td>
<td>1 1 1 1 1 1 NA 1 1 1 1 1 0 0 1 NA 1 13/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meletti et al., (2003a)</td>
<td>1 1 1 1 1 1 0 NA 0 0 0 1 0 0 0 0 NA 1 7/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meletti et al., (2003b)</td>
<td>1 1 1 1 1 1 1 NA 0 0 1 1 1 0 0 1 NA 1 11/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Meletti et al., (2009)</td>
<td>1 1 1 1 1 1 1 NA 1 1 1 1 1 0 0 1 NA 1 13/16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study Authors</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>NA</td>
<td>1</td>
</tr>
<tr>
<td>-------------------------</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>----</td>
<td>---</td>
</tr>
<tr>
<td>Meletti et al., (2014)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>NA</td>
<td>1</td>
</tr>
<tr>
<td>Pinabiaux et al., (2013)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>Realmuto et al., (2015)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>Reynders et al., (2005)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>Sedda et al., (2013)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>NA</td>
</tr>
<tr>
<td>Shaw et al., (2007)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Szaflarski et al., (2014)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>Tanaka et al., (2013)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>Walpole et al., (2008)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>NA</td>
</tr>
<tr>
<td>Wendling et al., (2015)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>NA</td>
</tr>
</tbody>
</table>

Items: (1) Is the hypothesis/aim/objective of the study clearly described? (2) Are the main outcomes to be measured clearly described in the Introduction or Methods section? (3) Are the characteristics of the patients included in the study described clearly? (4) Are the distributions of principal confounders in each group of subjects to be compared described clearly? (5) Are the main findings of the study described clearly? (6) Does the study provide estimates of the random variability in the data for the main outcomes? (7) Have the characteristics of patients lost to follow-up been described? (8) Have actual probability values been reported (for example, 0.035 rather than <0.05) for the main outcomes except where the probability value is less than 0.001? (9) Were the subjects asked to participate in the study representative of the entire population from which they were recruited? (10) If any of the results of the study were based on ‘data dredging’, was this made clear? (11) Were the statistical tests used to assess the main outcomes appropriate? (12) Were the main outcome measures used accurate (valid and reliable)? (13) Were the patients in different groups recruited from the same population? (14) Were study subjects recruited over the same period of time? (15) Was there adequate adjustment for confounding in the analyses from which the main findings were drawn? (16) Were losses of patients to follow-up taken into account? (17) Did the study have sufficient power to detect a clinically important effect where the probability value for a difference being due to chance is less than 5%?

All items given 0 or 1 point, except for item 4, which was given 0, 1 or 2 points. Items 7 and 16 were not applicable to cross-sectional studies and were marked as NA (Not Applicable). Total possible score for longitudinal studies was 0 to 18 (1 study) and for cross sectional studies was 0 to 16 (11 studies).

Items 1 – 8 (quality of reporting), item 9 (external validity), items 10 – 12 (internal validity: statistical and methodological bias), items 13 – 16 (internal validity: selection bias), item 17 (power).
4 Discussion

The primary aim of the review was to establish whether patients with epilepsy display global or emotion-specific deficits in facial emotion perception. The secondary aim was to determine whether deficits in facial emotion perception are moderated by epilepsy factors (site and side of seizure focus, age of onset, duration of epilepsy disorder, and seizure frequency), treatment factors (number of antiepileptic drugs, surgery), functional organisation of facial emotion perception, and/or demographic factors (age at testing, sex).

4.1 Overall Facial Emotion Perception Accuracy

Our meta-analyses reveal that patients with epilepsy have significantly reduced facial emotion perception accuracy, with the overall mean score of patients with epilepsy falling almost one standard deviation below the mean score obtained by controls. Given that facial emotion perception is proposed to be supported by a distinct network of brain structures (Adolphs, 2002; Calder & Young, 2005), we could expect that facial emotion perception is spared when epilepsy focus/pathology does not involve this brain network (i.e., in patients with GGE). Conversely, facial emotion perception would be impaired when epilepsy focus/pathology does involve this brain network (i.e., in patients with TLE and FCE). Hence, it is of interest to review our findings on the presence and gravity of facial emotion perception in relation to the site of epilepsy focus/pathology. The majority of studies (18 out of 24) in this review include patients with TLE, with only a small number of studies including patients with extra-temporal epilepsies (FCE, 2 out of 24; GGE 3 out of 24). Facial emotion perception is significantly reduced in patients with TLE, but not in patients with FCE. While the lack of deficits in facial emotion perception was not expected in FCE, inspection of effect sizes reveals little difference in the magnitude of facial emotion perception deficits between TLE and FCE groups. Thus, the lack of statistically significant findings in patients with FCE is likely to be due to the small number of studies/patients (2
studies, 22 patients), resulting in a lack of power, and a false negative finding in our meta-analysis.

In contrast, in patients with GGE, whose epilepsy focus/pathology may not involve the facial emotion perception network, our meta-analysis shows significant deficits in facial emotion perception that are comparable in magnitude to deficits found in patients with TLE and patients with FCE. Currently, it is unclear why patients with GGE have impaired facial emotion perception. Three possibilities provide potential explanations. First, patients with GGE experience generalised seizures that may disrupt functional integrity of the facial emotion perception network. Nevertheless, as these seizures are typically mild and easy to control with medication (Bourgeois et al., 1987), we would expect that the magnitude of facial emotion perception deficits (if present) would be smaller than in epilepsies with focus/pathology within the facial emotion perception network. Second, it is possible that facial emotion perception is influenced by genetic factors that also underpin epilepsy. For example, children with epilepsy have a higher risk of neurodevelopmental disorders, such as autism spectrum disorder (Sundelin et al., 2016). Facial emotion perception accuracy in people with autism spectrum disorder is also reduced compared to controls (Law Smith, Montagne, Perrett, Gill, & Gallagher, 2010; Uljarevic & Hamilton, 2012). Interestingly, the siblings and children of patients with epilepsy also have an increased risk of autism spectrum disorder compared to population levels, suggesting a shared aetiology, and potential underlying genetic mechanism common across epilepsy, autism spectrum disorder, and facial emotion perception (Gilby & O’Brien, 2013; Lau et al., 2009). Thirdly, facial emotion perception may be impacted by social factors in addition to biological. For example, exposure to facial emotions may be reduced if seizure frequency and management has limited the quantity and quality of interactions occurring in school or work environments from absenteeism (e.g., Aguiar, Guerrero, McBrian, & Montenegro, 2007), or stigma (de Boer,
Mula, & Sander, 2008; Jacoby & Austin, 2007). Arguably, greater seizure frequency and severity may be associated with reduced interpersonal interactions, and opportunities for practice perceiving facial emotions in varying social contexts.

Meta-regression, which only included patients with TLE, found just one moderator variable is associated with facial emotion perception: younger age at assessment is related to lower facial emotion perception accuracy. Yet, only three studies involved children with epilepsy (Braams et al., 2015; Golouboff et al., 2008; Pinabiaux et al., 2013), representing a serious shortcoming in the literature, as facial emotion perception is critical for the development of social cognition (Rantanen, Eriksson, & Nieminen, 2012) and social functioning (Kok et al., 2014; Van Overwalle, 2009). Hence, impaired facial emotion perception may hinder the development of social cognitive skills, such as theory of mind, that are important for successful interpersonal relationships and social interaction (Rantanen et al., 2012; Realmuto et al., 2015). Turning attention to the developmental literature; facial emotion perception and categorisation improve with age in typically developing children (Gao, Maurer, & Nishimura, 2010; Tremblay et al., 1987). As our study shows that impaired recognition of facial emotions is present not only in children, but also in adults with epilepsy, it is possible that the developmental trajectories of facial emotion perception in children with epilepsy and typically developing children differ. It is difficult to ascertain, due to the limited number of studies of children with epilepsy, whether effect sizes have a trait-like stability present from childhood into adulthood, or if the deficit progressively increases, indicating a slow rate of developmental progress. Furthermore, it is unknown whether developmental gains in facial emotion perception made by children with epilepsy are secondary to functional reorganisation and/or employment of compensatory strategies (e.g., attention to situational factors).
Other moderator variables, such as age of onset, duration of epilepsy disorder, and seizure frequency, are not associated with facial emotion perception in patients with epilepsy. The lack of association between the age of epilepsy onset and facial emotion perception is surprising, as a younger age of onset has been repeatedly associated with poorer outcomes in other cognitive areas, such as IQ and memory (Jambaque, Dellatolas, Dulac, Ponsot & Signoret, 1993; Jocic-Jakubi & Jovic, 2006; Kernan et al., 2012; Nolan, 2004; Pavone et al., 2001; Reilly et al. 2014; Schoenfeld et al., 1999). It is possible that this lack of association between the age of epilepsy onset and facial emotion perception is due to the non-linear developmental trajectory of some emotions (e.g., anger; Thomas et al., 2007), which meta-regression is unable to detect. Facial emotion perception deficits present since childhood may result from seizure activity disrupting a critical acquisition period (e.g., prior to the age of five; Meletti et al., 2003b), that may be undetected if deficits fall within normal developmental ranges that are broad in childhood. It is also possible that patients with epilepsy onset in early childhood develop compensatory strategies, which are not measured by facial emotion perception accuracy scores. It is surprising that patients who underwent temporal lobectomy did not differ in facial emotion perception from patients who were awaiting surgery, as the temporal lobes are a key component of the brain network (Adolphs, 2002). These non-significant findings, however, were based on cross-sectional studies, but could also reflect that the areas identified for resection may be damaged and non-functional. The only study that examined changes in facial emotion perception pre- to post- temporal lobectomy included children who underwent surgery at an early age (<12.1 years), and found a post-surgery decline compared to pre-surgery accuracy (Braams et al., 2015). Facial emotion perception accuracy recovered to, but did not surpass, pre-surgical accuracy at 24 months post-surgery. Given facial emotion perception in typically developing children continued to improve throughout the same time period, children with epilepsy were likely to
have more severe deficits in facial emotion perception at 24-months post-surgery. However, there are currently no longitudinal neuroimaging studies available to determine functional facial emotion perception changes. Together, at a first glance, findings of our meta-analyses and in particular, a lack of significant associations between several moderator variables (site and side of epilepsy focus, and surgery), seem to provide support for a constructionist (Barrett & Bliss-Moreau, 2009; Calder & Young, 2005; Haxby et al., 2000) rather than locationist (Gur et al., 2002; Lindquist et al., 2012) model of facial emotion perception, which purports that facial emotion perception relies on the maturation and ongoing function of an interconnected network of brain regions, allowing for plasticity in function.

4.2 Emotion-Specific Facial Emotion Perception Accuracy

Across the total sample, deficits in facial emotion perception are evident for each emotion. However, further analysis reveals that patterns of facial emotion perception impairment differ in relation to the site of seizure focus (i.e., the epilepsy group; TLE, FCE, GGE or unspecified epilepsy). Patients with TLE demonstrate deficits in facial emotion perception for all emotions except surprise. This selective sparing of surprise may be considered evidence against surprise as a basic facial emotion, given its interpretation as a positive or negative expression depends upon the social context in which it occurs (Fowler et al., 2006).

In contrast to patients with TLE, patients with GGE have selective deficits in facial emotion perception of anger, disgust, and fear, but not happiness, sadness, or surprise. It is interesting to note that while the effect sizes for emotions in which deficits were found are large (g = 0.566 – 1.085), the effect sizes for spared emotions are quite small (g = 0.037 – 0.167). Effect sizes were inspected for patients with TLE and GGE for emotions that were impaired in TLE, but spared in GGE, namely happiness (TLE g = 0.554, GGE g = 0.037), and sadness (TLE g = 1.032, GGE g = 0.167). We notice that the lack of significant findings
(sparing) is not due to limited power, as effects sizes were very small in patients with GGE. These distinct patterns of impairment may be associated with different functional implications. For example, a recent study found that, compared to controls, patients with GGE who had greater difficulty recognising sad faces were also more impaired on tasks of cognitive empathy (Jiang et al., 2014). As deficits in sadness recognition in patients with TLE is almost one standard deviation below those with GGE, patients with TLE may have considerably greater difficulty initiating empathic responses to distress. This could contribute to increased social morbidity, as empathy is associated with prosocial behaviour (Deschamps, Schutter, Kenemans, & Matthys, 2014; Miller, Nuselovici, & Hastings, 2016).

Differences in the pattern of facial emotion perception deficits may also be in part related to developmental factors. For example, facial emotion perception of happiness may be spared as it is the first recognisable emotion, whereas anger, disgust, and fear have more elongated developmental trajectories that are more susceptible to disruption from seizure activity (Herba & Phillips, 2004). This marked difference in effect sizes for different emotions, even in patients with TLE who have deficits in recognition of all but one facial emotion, supports the possibility that facial emotion perception can be differentially impaired. Three of four studies (Hlobil et al., 2008; McClelland et al., 2006; Meletti et al., 2003a) that dichotomised patients with TLE into early (prior to 5 to 6 years) and late (after 5 to 6 years) epilepsy onset groups found patients with early onset right-TLE are more likely to have impaired fear recognition, while one study (Hennion et al., 2015) found no difference in facial emotion perception between early and late onset groups. This further suggests that either hemisphere appears capable of supporting facial emotion perception, but functional reorganisation may be less achievable with disruption in early facial emotion perception development. Alternatively, facial emotion perception deficits may be a biomarker (trait) of epilepsy, rather than a consequence of epilepsy.
4.3 Limitations and Future Directions

Findings of our study should be interpreted with caution for several reasons. First, while epilepsy is a heterogeneous disorder, only five papers involved patients with extra-temporal epilepsies. Future research should include patients with extra-temporal epilepsies and assess whether facial emotion perception is impaired globally or in emotion specific domains in each epilepsy group. This will allow further investigation and clarification of whether patients with different types of epilepsy present with distinct patterns of emotion specific deficits, as we found in our study.

Second, our work shows that research into emotion processing of children with epilepsy has been neglected. It is unclear how facial emotion perception changes in patients with epilepsy from childhood to adulthood. Longitudinal studies of patients with epilepsy need to be conducted to further investigate the impact of epilepsy on facial emotion perception development, and the impact of facial emotion perception deficits on development of higher social-cognitive skills. It is critical that these studies also include a control group of typically developing participants, to determine whether developmental trajectories of the two groups differ. Ideally, future studies would investigate mechanisms that underpin developmental changes using functional neuroimaging and behavioural tasks concurrently.

Third, while the overall quality of studies included in this paper, as determined by the Downs and Black (1998) Checklist, was adequate, specific areas of weakness were noticed. The main shortcomings relate to external validity and reporting of participant source and timing of recruitment. These shortcomings may contribute to selection bias, and therefore, limit the representativeness of the samples. Studies that reported on recruitment all identified patients through specialised epilepsy management services or hospitals, with none drawing patients from a community population, and only one third recruiting patients consecutively. Similarly, patients had often undergone, or were candidates for epilepsy surgery, indicating a
lack of response to polytherapy, and a classification of their seizures as “medically intractable”. Arguably, this constitutes a sample of patients with more severe epilepsy (e.g., greater frequency of seizures, resistance to polytherapy), so findings are likely to be most useful for patients who require specialised services for medically complex epilepsy. Future studies should consider direct comparison of facial emotion perception in patients with varying epilepsy severity.

Fourth, while facial expressions of emotion are thought to be universally recognised, it may also be worthwhile assessing facial emotion perception deficits cross-culturally, as only two studies examined populations from non-Western nations (Hlobil et al., 2008; Tanaka et al., 2013). However, we recognise the potential exclusion of relevant studies from our analysis due to database searches with records of research conducted predominantly in the English language.

Our study also has some limitations. Non-significant effect sizes for moderator variables (apart from age at testing) may be an artefact of using means, rather than individual data, in our study. The use of means may have limited the spread of scores, and therefore power to detect significant effects. Overlap in control groups also prevented direct comparisons of individual facial emotions within epilepsy groups, to provide relative patterns of impairment and sparing, rather than absolute differences.

Finally, while research has indicated the modulatory role of expression intensity on facial emotion perception accuracy (Ammerlaan et al., 2008; Sedda et al., 2013; Shaw et al., 2007), the current study only analysed expressions presented at 100% intensity. Research has shown that modulating expression intensity through morphing techniques increases task sensitivity, and therefore, detection of subtle impairments of facial emotion perception, including potential gender differences (Hoffmann, Kessler, Eppel, Rukavina, & Traue, 2010;
Law Smith et al., 2010). As such, our study may have underestimated morbidity of emotion perception deficits in patients with epilepsy.

4.4 Conclusions

This study is, to our knowledge, the first comprehensive meta-analysis of facial emotion perception in patients with epilepsy. The meta-analysis reveals marked facial emotion perception deficits in patients with epilepsy, with differential patterns of emotion-specific facial emotion perception deficits in patients with TLE compared to GGE. In contrast, no significant differences in facial emotion perception are found in relation to the side of seizure focus or surgical status in patients with TLE. Thus, findings of our study are consistent with a more recent conceptualisation of facial emotion perception in the human brain, which suggests that facial emotion perception relies on a core visual processing system, complemented by an extended system of emotion processing areas (Calder & Young, 2005; Haxby et al., 2000). Critically, in patients with TLE, meta-regression shows that a younger age of testing is associated with lower facial emotion perception accuracy. These early facial emotion perception deficits may have a significant, flow-on effect to the development of higher-level social-cognitive skills, the acquisition of interpersonal abilities, and ultimately may contribute to the social morbidity experienced in patients with epilepsy.

In our study we have identified areas of research need and provided recommendations, including the need for studies examining facial emotion perception in children with epilepsy and patients with seizure focus outside the temporal lobes, and unique patterns of impairment that can inform remediation strategies for facial emotion perception to reduce social morbidity. By targeting these areas of need and developing studies with recommendations from this review, future studies can (i) provide more detailed characterisation of facial emotion perception deficits, (ii) advance our knowledge of risk factors, and, (iii) progress our understanding of neural architecture and plasticity of facial
emotion perception. In turn, these findings would advance clinical management and enable development of treatments aimed at minimising social morbidity in patients with epilepsy.
Acknowledgements

We thank the researchers who generously provided access to their data upon our request.
References


Neurology, Neurosurgery & Psychiatry, 75(4), 593–599.
http://doi.org/10.1136/jnnp.2002.006403


Figure Captions

Figure 1. Flow diagram of identification and selection of studies.

Figure 2. Individual and mean weighted effect sizes (Hedges' $g$ and 95% CIs) for facial emotion perception studies based on accuracy of responding (collapsed across emotions) for patients with temporal lobe epilepsy (TLE), frontocentral epilepsy (FCE), genetic generalised epilepsy (GGE), and unspecified epilepsy.

Figure 3. Individual and mean weighted effect sizes (Hedges' $g$ and 95% CIs) for facial emotion perception accuracy collapsed across epilepsy groups by emotion.

Figure 4. Individual and mean weighted effect sizes (Hedges' $g$ and 95% CIs) of facial emotion perception accuracy (collapsed across emotions) for temporal lobe epilepsy (TLE) patients with left, right, or bilateral seizure focus.

Figure 5. Individual and mean weighted effect sizes (Hedges' $g$ and 95% CIs) of facial emotion perception accuracy (collapsed across emotions) for temporal lobe epilepsy (TLE) patients pre- or post-temporal lobe surgery.
Tables

Table 2. Characteristics and patient demographics of studies included in the meta-analysis.

Table 2. Mean weighted effect sizes as a function of facial emotion by epilepsy group.