# Cognitive Functioning in Children and Adults With Nonsyndromal Cleft Lip and/or Palate: A Meta-analysis

Rachel M. Roberts, PhD, Jane L. Mathias, PhD, and Patricia Wheaton, PhD
*School of Psychology, University of Adelaide*

All correspondence concerning this article should be addressed to Rachel M. Roberts, School of Psychology, University of Adelaide, Adelaide, South Australia, 5005, Australia. E-mail: rachel.roberts@adelaide.edu.au

Received October 25, 2011; revisions received February 1, 2012; accepted February 27, 2012

**Objective** To provide a meta-analysis of research that has examined the cognitive functioning of children and adults with cleft lip (CL), cleft palate (CP), and cleft lip and palate (CLP). **Methods** Data from 29 studies, which compared persons with a cleft to a control group on tests of cognitive functioning, were analyzed. The data were obtained from 1,546 persons with cleft and 279,805 controls. **Results** Participants with a cleft performed significantly worse on 7 cognitive domains. However, the only moderate and significant deficit, which was based on nonheterogeneous study findings and not subject to publication bias, was in the language domain. CL, CP, and CLP were all associated with cognitive impairments, although the profiles for the groups differed. **Conclusions** Cross-sectional studies suggest that persons with clefts experience poorer cognitive functioning across a range of domains, although large-scale longitudinal studies are needed to more definitively differentiate outcomes by cleft type.

**Key words** cleft lip; cleft lip and palate; cleft palate; cognitive; meta-analysis.

A cleft, affecting the lip and/or palate, is one of the most common birth defects and the most common cause of craniofacial anomalies in infants (Canfield et al., 2006). Cleft lips (CLs) and cleft palates (CPs) both result from abnormal facial development during gestation. A CL involves one or more clefts in the upper lip due to a failure in the two sides of the upper lip and nose to fuse correctly and a CP involves a fissure in the midline of the roof of the mouth where the two sides of the palate have failed to join normally, which may additionally be associated with a cleft lip and palate (CLP) (Anderson, 1998). Clefts can either be of nonsyndromal origin, where the etiology is unknown, or syndromal, where the cleft is part of a known syndrome. Syndromal clefts are associated with a range of other congenital malformations and more diverse outcomes (e.g. Apert Syndrome is associated with respiratory, hearing, speech, attention, and language difficulties), consequently they are not the focus of this study. In the largest U.S. study of infants with nonsyndromal orofacial clefts, the prevalence of CL, CP, and CLP was estimated to be 0.3, 0.4, and 0.5 per 1,000 live births, respectively (total prevalence: 1.2 per 1,000 live births) (Genisca et al., 2009).

Over the past 30 years, there have been significant improvements in our ability to correct some of the more severe orofacial deformities, such as clefts, with an associated decrease in the morbidities that are associated both with the condition itself and the corrective surgery (Monro, 1995). Despite this, most children with CL, CP, or CLP continue to experience one or more of a range of associated difficulties, including chronic ear infections and problems with feeding and growth, dental development, and speech and language development.

Cognitive and academic problems have also been reported in children and adults with CL/CP/CLP by researchers who have attempted to document the full extent of problems that can arise from these three conditions (e.g., Broder, Richman, & Matheson, 1998). In addition, recent research has found that children and adults

*Journal of Pediatric Psychology* 37(7) pp. 786–797, 2012
doi:10.1093/jpepsy/jss052
Advance Access publication March 26, 2012

with CLP have abnormal brain structure. For example, children with clefts reportedly have smaller brain volumes, with the frontal lobes and certain subcortical nuclei (caudate, putamen, and globus pallidus) being most affected (Nopoulos, Langbehn, Canady, Magnotta, & Richman, 2007). In contrast, adult males with clefts have been found to have normal brain volumes but enlarged frontal and parietal lobes, and smaller temporal and occipital lobes (Nopoulos et al., 2001). These differences in brain volume and structure may be related to the cognitive problems that have been reported in people with orofacial clefts (Nopolous et al., 2002).

Other factors that may contribute to the development of cognitive problems in people with clefts include hearing or speech deficits and the psychological effects of living with a facial disfigurement. Moreover, there are a number of variables that are thought to affect the likelihood of developing cognitive problems, including the age at which a cleft is surgically repaired (e.g., Murray et al., 2008), whether there is a history of articulation or hearing problems (e.g., Jocelyn, Penko, & Rode, 1996), cleft type (CL, CP, or CLP) (e.g., Nakajima, Mutsudome, & Yosikawa, 2001; Persson, Becker, & Svensson, 2008), and severity of the cleft(s) (e.g., Fox, Lynch, & Brookshire, 1978). Gender has also been related to cognitive outcome in some studies. For example, Broder et al. (1998) found a significant interaction between cleft type and gender, with males with CP having the highest rate of learning problems. However, other studies have not found any gender differences (e.g., Nakajima et al., 2001).

Although the cognitive abilities of children and adults with CL/CP/CLP have been examined, the findings are inconsistent. Indeed, some studies report superior performance by individuals with clefts (compared to healthy controls) on specific cognitive tasks (e.g., Pecyna, Feeney-Giacoma, & Neiman, 1987), some report no cognitive impairments (e.g., Persson et al., 2008; Starr, Chinsky, Canter & Meier, 1977) or only verbal impairments (e.g., Ruess, 1965), and others report more global cognitive deficits (e.g., Nopolous et al., 2002; Speltz et al., 2000). In addition, there are a number of literature reviews on cognitive functioning in CL/CP/CLP (e.g., De Soussa, Devare, & Ghanshani, 2009; Hunt, Burden, Hepper, & Johnston, 2005) but, not surprisingly, the conclusions drawn from these reviews are inconsistent. Furthermore, the selective inclusion of research and/or the failure to specify the criteria by which research was selected, make it difficult to evaluate and compare these reviews.

Variation in the research findings on the cognitive functioning of children and adults with CL/CP/CLP may, in part, result from the large number of tests that have been used to assess a range of different cognitive functions (e.g., development, intelligence, achievement, or specific cognitive functions) together with the fact that not all studies have separately examined outcome for different types of clefts (i.e., CL, CP, CLP). In addition, different control groups have been used in order to determine whether cognitive deficits exist or not. Specifically, a number of studies have compared participants with clefts to published normative data rather than to matched controls (e.g., Kapp-Simon & Kreukeberg, 2000; Starr et al., 1977). Alternatively, where studies have used a control group, some have recruited healthy controls (e.g., Murray et al., 2008), some have used sibling controls (e.g., Ruess, 1965), and others have used controls who have been diagnosed with another chronic condition (e.g., Lovius, 1971). Moreover, most studies are based on a small number of participants, especially when grouped according to type of cleft. In combination, these methodological differences are likely to have contributed to variation in the research findings. Finally, published research from previous decades may not be equally relevant today, due to recent improvements in both surgical techniques and the management of hearing and speech impairments, the introduction of early academic interventions, and the possibility that there is a greater acceptance of (and less stigma associated with) orofacial clefts in schools and communities.

An improved understanding of the cognitive and academic consequences of CL, CP, and CLP is essential in order to be able to educate and inform families, use early interventions to prevent or minimize the development of cognitive problems, and provide educational support to those children, adolescents and adults whose cognitive problems are affecting their learning. The current study, therefore, meta-analyzed existing published research that has examined the cognitive functioning of infants, children or adults with CL, CP, or CLP in order to establish both the nature and severity of cognitive impairments that are associated with nonsyndromal orofacial clefts, when compared to a control group. As recent research suggests that CL, CP, and CLP may be etiologically distinct (e.g., Genisca et al., 2009). A secondary aim was to compare the cognitive profiles of the three cleft types. Although limited, existing research has produced mixed findings with regard to differences in cognitive functioning between cleft types. For example, whereas Persson et al., (2008) and Conrad et al., (2009) both found that CP had poorer cognitive outcomes than those with CL and CLP, Collett, Leroux et al., (2010) have reported that children with CP had stronger, and children with CLP had equivalent, reading skills than control children.

## Methods

### *Literature Search and Inclusion Criteria*

The PubMed and PsychINFO electronic databases were searched from January 1960 to January 2011 in order to identify studies that examined cognitive functioning in people with CL and/or CP. The initial search used a large number of search terms in order to capture all relevant articles ($N = 190$ search terms; refer to Table A of the Supplementary Material). A study was eligible for inclusion if it met the following criteria: (a) was published in a journal in English; (b) examined a CL, CP, and/or CLP group, together with an age-matched noncleft control group (excluding norms); (c) was not a case study; (d) no participant was identified as having a known syndrome (i.e. Smith-Lemili-Opitz syndrome) or comorbid condition that might affect cognition (i.e. microcephaly); (e) tests of cognitive functioning were administered; and (f) the results were reported in a format that enabled the calculation of an effect size (i.e. means, *SD*s, *t*-tests, one-way *F*-statistics, odds ratios, exact *p*-values).

The literature searches initially identified 3,683 articles (refer to Figure A of the Supplementary Material). A preliminary application of the inclusion criteria to the titles and abstracts of these articles identified 160 studies that were potentially relevant. Reapplication of the inclusion criteria to the full-text versions of these papers revealed that 129 did not meet the inclusion criteria, leaving 31 studies. An examination of the reference lists of all retrieved articles identified a further 26 studies; 14 of which were potentially relevant. Two of these additional references met all of the eligibility criteria, increasing the number of studies to 33 (see Figure A of the Supplementary Material for details of study numbers by data base).

Meta-analytic procedures require that all samples are independent of one another (Rosenthal, 1995). For this reason, two studies by Lewis (1971, 1961), two studies by Conrad and colleagues (Conrad et al., 2008; Conrad et al., 2009) and three studies by Nopoulos and colleagues (Goldsberry, O'Leary, Hichwa, & Nopoulos, 2006; Nopoulos et al., 2001, 2002), which used the same participants, were combined and treated as one; reducing the final number of independent studies to 29.

### *Data Collection and Preparation*

Demographic and medical data (e.g., age, gender, diagnosis, number of craniofacial surgical procedures, age at first surgery), methodological information (e.g., cognitive test, participant numbers), and the data required for the calculation of the effect sizes were extracted from each study. Each cognitive test was subsequently classified into 1 of 10 categories, reflecting the main cognitive domain that it assessed: processing speed, memory—immediate recall and delayed recall, language, attention/executive functions, sensorimotor function, motor function, general cognition, academic ability, and visuospatial ability (refer to Supplementary materials, Table B for the test classifications).

In some studies, children underwent repeated testing at different ages in order to evaluate developmental changes in performance. Assessments undertaken at older ages are thought to provide a better predictor of outcome in children (Richman & Eliason, 1986), consequently only data from the final assessment were analyzed when multiple assessments were reported.

### *Effect Size Calculation and Interpretation*

Cohen's *d* effect sizes were used to compare the difference between the cognitive performance of the cleft and control groups (Zakzanis, 2001) for the 10 broad cognitive domains. A negative *d* indicates that the cleft group performed more poorly than the controls. According to Cohen (1992), a small effect is defined as $d = .2$, a medium effect as $d = .5$ and a large effect as $d = .8$, with an effect size of .5 indicating that the average scores of the two groups differ by one half of a pooled standard deviation. If means and standard deviations were not provided, *t*-values, one-way *F*-statistics, exact *p*-values and odds ratios were converted to *d* (Chinn, 2000; Lipsey & Wilson, 2001; Zakzanis, 2001).

Mean effect sizes were calculated for each of the cognitive domains (i.e., processing speed, memory, language, etc.) using a multi-stage process. First, effect sizes were calculated for every cognitive test that was used by a study and these were then sorted by cognitive domain. If an individual study provided multiple effect sizes (i.e., multiple tests or scores) for a cognitive domain, these were averaged so that each study only contributed a single effect size to the calculation of an average for that cognitive domain (i.e., one study, one vote principal; Turner & Bernard, 2006). As the reliability of an effect size is affected by sample size, it is important to weight effect sizes from different studies before they are combined to calculate a mean effect size (Lipsey & Wilson, 2001). This was done by weighting effect sizes by their inverse variance (i.e., the inverse of the squared standard error), as this provides a more precise measure of reliability than sample size (Hedges & Olkin, 1985). The weighted effect sizes for all studies that assessed a particular cognitive domain were then averaged in order to evaluate the effects of CL, CP, and CLP on cognitive functioning.

A number of additional statistics were also calculated to aid in the interpretation of the results. More specifically, 95% confidence intervals (95% CIs) were calculated to provide a measure of the statistical significance of the effect sizes. If a CI does not span zero, this indicates that there is a significant difference in the performance of the cleft and control groups (Hedges & Olkin, 1985; Lipsey & Wilson, 2001). In addition, percentage overlap scores (%OL) were calculated to provide a measure of the extent to which the scores of the two groups overlapped (Zakzanis, Leach, & Kaplan, 1999); $d = 0$ signifies complete overlap between the groups (%OL = 100%) and $d = 4.00$ almost complete discrimination (%OL = 2.3%). Fail safe *N*s ($N_{fs}$) were also calculated to address the bias toward publishing studies with significant findings, which may inflate the final effect size calculations. The $N_{fs}$ statistic provides a hypothetical measure of the number of unpublished studies with nonsignificant results (i.e., small effects: $d \leq .2$) that would need to exist (in file drawers) in order to call the current findings into question (Rosenthal, 1995; Zakzanis, 2001). The larger the $N_{fs}$, the more confidence we have in that finding. Finally, a measure of heterogeneity ($I^2$) was also calculated to determine whether the variability in the effect sizes that were obtained from different studies (and then averaged to obtain an overall effect size for each cognitive domain) was due to normal sampling error or caused by systematic differences between studies, such as differences in the inclusion and exclusion criteria, recruitment methods or participation rates (Sutton, Abrams, Jones, Sheldon, & Fong, 2000).

The conclusions drawn from this meta-analysis are based on the combined interpretation of the Cohen's *d*, 95% CIs, $N_{fs}$, and $I^2$ statistics. We argue that we can be more confident that cognition is impaired in the cleft group if there are moderate to large ($d \geq .5$) and significant (95% CIs $\neq 0$) differences in the cognitive performance of the cleft and control groups. In addition, the $N_{fs}$ statistic needs to be large enough to indicate that it is unlikely that there would be sufficient unpublished studies with small effects to draw the current findings into question ($N_{fs} > N_{studies}$). Finally, there should be statistical homogeneity between studies ($I^2$ nonsignificant) to enable the conclusion that variability in effect sizes was due to normal sampling error, rather than systematic differences between studies. The Cochrane collaboration (Higgins & Green, 2011) suggests $I^2$ values over 40% may represent heterogeneity of a moderate (30–60%), substantial (50–90%), or considerable (75–100%) degree. While the main focus is on those results that meet all of these statistical criteria, the data from all studies that met our inclusion criteria are provided. In a number of instances there were only a small number of studies that had assessed a cognitive domain. This data is provided in the interests of completeness, as this is an emerging area of research. However, the greater the number of studies and participants, the more confident we can be of a finding.

The data from all 30 studies, regardless of cleft type were initially analyzed together, after which the data for CL, CP, and CLP were analyzed separately in order to evaluate the cognitive profiles of each cleft type.

## Results

### *Participant Characteristics*

In total, data from 29 studies were analyzed. Twenty-five studies used healthy controls, three used sibling controls and one used orthopedic patients. The methods for recruiting control participants was not always reported but included recruitment via friends ($N_{studies} = 1$), health services ($N_{studies} = 7$), advertising ($N_{studies} = 2$), and schools ($N_{studies} = 7$). Demographic data for the cleft and noncleft groups are reported in Table I which shows that participant's mean age ranged from 1 to 30 years, with a mean age for the cleft and noncleft groups of 11 years ($SD = 8$)

Table I. *Demographic and Surgery Data for the Cleft and Noncleft Control Groups*

| | Cleft group | | | | Noncleft group | | | |
|---|---|---|---|---|---|---|---|---|
| | $N_{studies}$ | $N_{participants}$ | *M (SD)* | Range | $N_{studies}$ | $N_{participants}$ | *M (SD)* | Range |
| Sample size | 29 | 1,546 | 53 (39) | 8–151 | 29 | 279,805[a] | 9,648 (50,857) | 4–278,598 |
| Age (years) | 16 | 810 | 10 (8) | 1–30 | 17 | 279,203 | 10 (8) | 1–29 |
| Males | 21 | 642 | 29 (32) | 4–151 | 20 | 312,068 | 16,425 (71,513) | 2–311,738 |
| Females | 21 | 305 | 15 (16) | 0–72 | 20 | 217 | 11 (13) | 0–45 |
| Participants with cleft repair surgery (*N*) | 12 | 453 | 38 (40) | 11–135 | | | | |
| Repair age (months)[b] | 6 | 2,589 | 12 (11) | 0–33 | | | | |

*Note*. $N_{studies}$ = Total number of studies; $N_{participants}$ = Total number of participants; *M* = mean; *SD* = standard deviation.

[a]Large *N* because included one population cohort study.

[b]Age at first operation.

and 10 years ($SD = 8$), respectively. Participants were young adults in four studies, school-aged in 13 studies, and under 4 years in 10 studies, and a further two studies included participants aged from infancy to childhood. There were more males than females in both groups, consistent with the higher prevalence of clefts among males. The mean age of the first cleft repair operation was 12 months ($SD = 11$), however, there was a wide range across studies (0–33 months). There was no significant difference in age between cleft and control groups [$t(30) = -.01$, $p = .99$]. However, control groups had a higher proportion of males than the cleft groups ($\chi^2 = 58{,}610.93$, $df = 1$, $p < .001$). Specific demographic data for all studies that were included in the meta-analysis, including type of cleft, sample sizes, age, and the cognitive tests that were used are provided in Table C of the Supplementary Material.

### *Cognitive Effects of Clefts*

The weighted mean effect sizes for each cognitive domain, comparing all participants with and without a cleft (CL, CP, and CLP), are provided in Table II. As can be seen, general cognitive ability and language were the most commonly assessed domains ($N_{studies} = 17$ and 15, respectively), followed by academic ability ($N_{studies} = 6$), attention ($N_{studies} = 5$) and motor skills ($N_{studies} = 5$), with only a few studies assessing the other domains ($N_{studies} = 1$–3). Importantly, with one exception (processing speed), the cleft samples were reasonably sized ($N_{participants} = 50$–946).

Participants with a cleft were found to have moderate and significant deficits in immediate memory, language, and attention/executive abilities (Table II). However, there was a high degree of heterogeneity for the immediate memory and attention effect sizes obtained by individual studies (i.e., $I^2 > 40\%$). There were also small but significant negative effects (deficits) for the sensorimotor, motor, general cognition, academic and delayed memory measures. The forest plot in Figure 1 shows the mean effect size and CIs for each cognitive domain. Thus, while there were small to moderate significant effects on 7 of the 10 cognitive domains, only the moderate deficit in the language domain was based on nonheterogeneous study findings and not subject to publication bias.

### *Cognitive Effects of Different Types of Clefts*

The cognitive data for three different cleft groups (CL, CP, and CLP) were also analyzed separately in order to determine whether there were different patterns of performance for the three cleft types. In total, there were 20 studies that reported data separately for specific cleft groups. Summary demographic data for these studies are provided in Table III (refer to Supplementary Materials, Table D, for effect sizes for each study and cognitive measure). Age and gender were not consistently reported for specific cleft groups but the available data is presented in Table III.

Three studies specifically assessed the cognitive performance of participants with a CL, compared to a control group. No cognitive domain was assessed by more than two studies and some domains were not investigated at all (i.e., processing speed, sensorimotor and motor functions, general cognition, delayed memory). Attention and general cognitive ability ($N_{studies} = 2$) and immediate memory, language, and visuospatial abilities ($N_{studies} = 1$) have all been examined. Sample sizes were small ($N_{participants} = 8$–22), reflecting the low number of studies with participants with a CL.

The mean weighted effect sizes for all cognitive domains for the CL group are reported in Table IV. Participants with a CL showed moderate and high-moderate statistically significant deficits in language and general cognitive ability, respectively, with low but acceptable $N_{fs}$ statistics and statistical homogeneity for general cognitive ability. However, this finding was based on only two studies and only one study examined language ability. Thus, of the domains that have been examined, the data suggests that a CL may be associated with specific language problems and more pervasive cognitive deficits, however, there are insufficient studies to draw any reliable conclusions.

There were 12 studies that compared the cognitive functioning of participants with a CP to that of a control group (refer to Table IV). However, only two domains were assessed by more than two studies; namley language ($N_{studies} = 7$) and general cognitive ability ($N_{studies} = 6$). Attention, motor ability, and academic ability were all assessed by two studies, while immediate memory and visuospatial ability were only assessed by one study. No study included measures of processing speed, sensorimotor function or delayed memory. Sample sizes were variable, ranging from small for the cognitive domains that were assessed by only one or two studies ($N_{participants} = 22$–67) but larger for language ($N_{participants} = 289$) and general cognitive ability ($N_{participants} = 238$).

Participants with a CP were found to have large impairments in immediate memory, as seen by the large Cohen's $d$, 95% CIs that did not span zero, the $N_{studies} > N_{fs}$, and the nonsignificant $I^2$; all of which met the study criteria. In addition, language and motor functioning met all but one of the study criteria; with high $I^2$ values for these domains indicating a high degree of variability in the effect sizes found by individual studies.

Table II. *Cohen's d Effect Sizes Comparing Cleft Groups and Controls on Cognitive Domain*

| Cognitive Domain | $N_{Studies}$ | $N_{Participants}$ | *Mean* $d_w$ *(SD)* | 95% CIs | $N_{fs}$ | %OL | $I^2$ (%) | Study references |
|---|---|---|---|---|---|---|---|---|
| Processing Speed | 1 | 14 | −0.75 | | 5 | 53 | | Brennan and Cullinan (1974) |
| Memory—immediate | 3 | 185 | −0.66* (.18) | −0.90 to −0.42 | 7 | 57 | 75.2* | Collett, Stott-Miller, Kapp-Simon, Cunningham, and Speltz (2010); Conrad, Richman, Nopoulos, and Dailey (2009); Nopoulos et al. (2002) |
| Language | 15 | 643 | −0.52* (.32) | −0.64 to −0.40 | 24 | 67 | 0 NS | Broen, Devers, Doyle, Prouty, and Moller (1998); Collett, Leroux, and Spelz (2010); Collett, Stott-Miller et al. (2010); Conrad et al. (2009); Ebert, McWilliams, and Woolf (1974); Fox et al. (1978); Jocelyn et al. (1996); Lamb, Wilson, and Leeper (1972); Lovius (1971); Nopoulos et al. (2002); Pannbacker (1975); Philips and Harrison (1969); Scherer and D'Antonio (1995); Smith and McWilliams (1966); Snyder and Scherer (2004) |
| Attention/Executive | 5 | 260 | −0.51* (.23) | −0.73 to −0.29 | 8 | 67 | 42.9* | Conrad et al. (2009); Lemos and Feniman (2010); Laasonen et al. (2004); Nopoulos et al. (2002); Smith and McWilliams (1966) |
| Sensorimotor function | 1 | 77 | −0.47* | | 2 | 67 | | Conrad, Canady, Richman, and Nopoulos (2008) |
| Motor | 5 | 211 | −0.42* (.59) | −1.31 to −0.02 | 6 | 73 | 74.1* | Jocelyn et al. (1996); Nopoulos et al. (2002); Scherer, Williams, and Proctor-Williams (2008); Snyder and Scherer (2004); Speltz et al. (2000) |
| General cognition | 17 | 946 | −0.38* (.62) | −0.48 to −0.28 | 16 | 73 | 74.1* | Broen et al. (1998); Conrad et al. (2009); Eide, Skjaerven, Irgens, Bjerkedal, and Oyen (2006); Fox et al. (1978); Goodstein (1961); Jocelyn et al. (1996); Lamb et al. (1972); Lewis (1971/1961); Lovius (1971); Murray et al. (2008); Nakajima et al. (2001); Nopoulos et al. (2002); Pecyna et al. (1987); Ruess (1965); Scherer et al. (2008); Snyder and Scherer (2004); Speltz et al. (2000) |
| Academic | 6 | 278 | −0.31* (.53) | −0.51 to −0.11 | 4 | 79 | 0 NS | Collett, Stott-Miller et al. (2010); Collett, Leroux et al. (2010); Richman (1976); Richman and Harper (1978); Ruess (1965); Stackhouse (1982) |
| Memory—delayed | 1 | 50 | −0.31 | | 1 | 79 | | Nopoulos et al. (2002) |
| Visuospatial | 3 | 140 | −0.10 (.45) | −0.83 to 0.45 | 2 | 92 | 33.8* | Lamb et al. (1972); Laasonen et al. (2004); Nopoulos et al. (2002) |

*Note.* $N_{Studies}$ = number of studies included in analysis; $N_{Participants}$ = number of participants in cleft group; *M* $d_w$ = mean weighted effect size; *SD* $d$ = standard error of effect size, NS = Nonsignificant, $^*p < .05$.

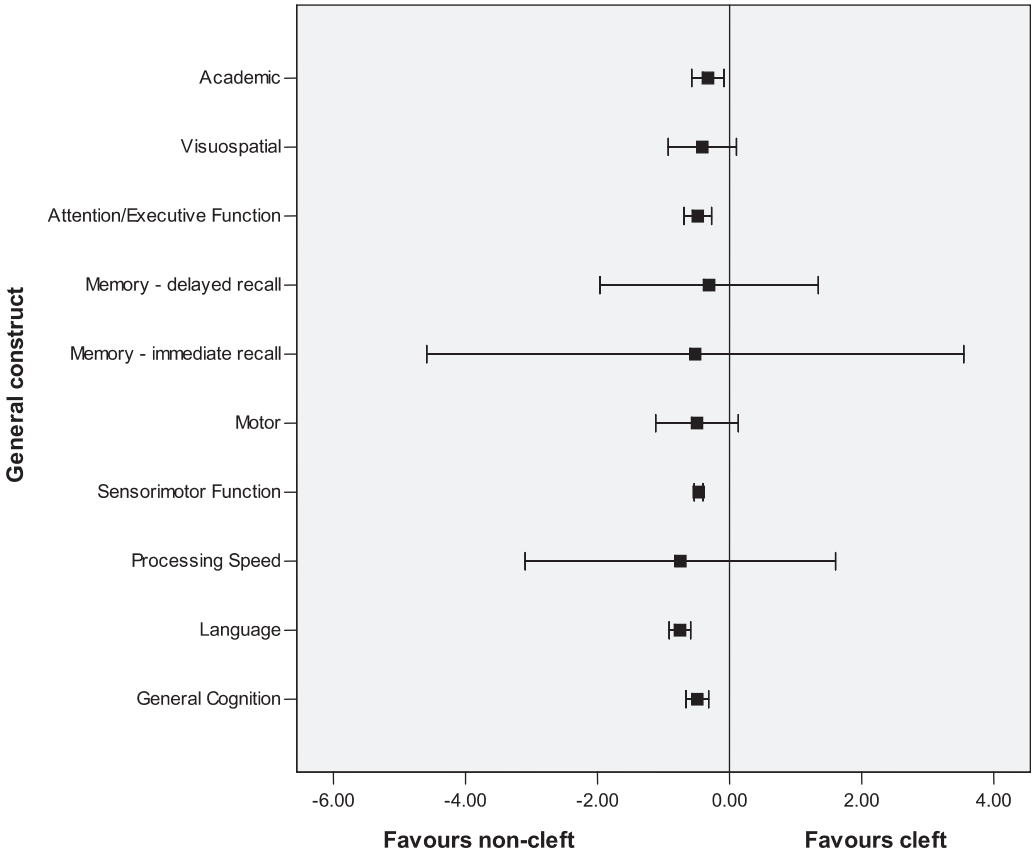


**Figure 1.** Error bar showing mean effect size and CIs for each cognitive domain.

Table III. *Summary Demographic Data for the Specific Cleft Groups and their Noncleft Control Groups*

| | Cleft group | | | | Control group | | | |
|---|---|---|---|---|---|---|---|---|
| | $N_{studies}$ | $N_{participants}$ | *M (SD)* | Range | $N_{studies}$ | $N_{participants}$ | *M (SD)* | Range |
| CL | 3 | 30 | 10 (3) | 8–14 | 3 | 189 | 63 (49) | 7–95 |
| Age (years) | 0 | 0 | | | 2 | 93 | 11 (2) | 10–13 |
| Males | 1 | 8 | 6 | | 1 | 66 | 31 | |
| Females | 1 | 8 | 2 | | 1 | 66 | 35 | |
| CP | 12 | 626 | 52 (46) | 11–154 | 12 | 279,211 | 23,268 (80,408) | 7–278,598 |
| Age (years) | 2 | 182 | 10 (12) | 2–19 | 6 | 278,837 | 9 (7) | 2–19 |
| Males | 7 | 325 | 34 (52) | 5–151 | 8 | 278,861 | 38,992 (102,060) | 17–311,738 |
| Females | 7 | 325 | 12 (9) | 0–24 | 8 | 278,861 | 19 (16) | 0–43 |
| CLP | 13 | 292 | 30 (26) | 8–103 | 13 | 526 | 40 (36) | 4–100 |
| Age (years) | 2 | 132 | 2 (1) | 1–2 | 10 | 340 | 5 (6) | 1–13 |
| Males | 6 | 113 | 13 (5) | 7–20 | 10 | 340 | 25 (20) | 4–65 |
| Females | 6 | 113 | 6 (4) | 3–13 | 10 | 340 | 18 (15) | 2–43 |

*Note:* $N_{studies}$ = Total number of studies; $N_{participants}$ = Total number of participants; *M* = mean; *SD* = standard deviation.

Overall, these results suggest people with CP experience large deficits in immediate memory and, possibly, language and motor functions. However, these findings can only be considered tentative as not all statistical criteria were met.

Finally, there were 13 studies that examined the cognitive functioning of people with a CLP (refer to Table IV). General cognitive ability, language, and attention were the most commonly assessed domains ($N_{studies}$ = 8, 5, 4, respectively), with only a few studies assessing motor skills,

Table IV. *Cohen's d Effect Sizes Comparing Cleft Groups and Controls on each Cognitive Domain*

| Cognitive Domain | $N_{Studies}$ | $N_{Participants}$ | *M (SD)* $d_w$*(d)* | 95% CIs | $N_{fs}$ | OL% | $I^2$ (%) | Studyreference |
|---|---|---|---|---|---|---|---|---|
| CL | | | | | | | | |
| Memory—immediate | 1 | 14 | −0.24 | | 1 | 85 | | Conrad et al. (2009) |
| Language | 1 | 14 | −0.70* | | 3 | 57 | | Conrad et al. (2009) |
| Attention/Executive | 2 | 22 | −0.04* (0.16) | −0.65 to −0.15 | 2 | 100 | 0 $^{NS}$ | Conrad et al. (2009); Laasonen et al. (2004) |
| General Cognition | 2 | 22 | **−0.52* (0**.04) | −0.97 to −0.07 | 3 | 67 | 0 $^{NS}$ | Conrad et al. (2009); Goodstein (1961) |
| Visuospatial | 1 | 8 | −0.67 | | 2 | 57 | | Laasonen et al. (2004) |
| CP | | | | | | | | |
| Memory—immediate recall | 1 | 22 | −0.89* | | 3 | 48 | | Conrad et al. (2009) |
| Language | 7 | 289 | −0.59* (0.09) | −0.68 to −0.50 | 14 | 62 | 70.75* | Broen et al. (1998); Collett, Leroux et al. (2010); Conrad et al. (2009); Ebert et al. (1974); Lovius (1971); Phillips and Harrison (1969); Snyder and Scherer (2004) |
| Attention/Executive | 2 | 64 | −0.38* (0.16) | −0.67 to −0.09 | 2 | 73 | 88.68* | Conrad, et al. (2009); Laasonen et al. (2004); |
| Motor | 2 | 39 | −0.91* (1.14) | −2.17 to −0.71 | 8 | 48 | 88.07* | Snyder and Scherer (2004); Speltz et al. (2000) |
| General cognition | 6 | 238 | −0.23* (0.44) | −0.37 to −0.09 | 1 | 85 | 84.58* | Broen et al. (1998); Conrad et al. (2009); Eide et al. (2006); Goodstein (1961); Lovius (1971); Snyder and Scherer (2004) |
| Academic | 2 | 67 | 0.00 (0.75) | −0.16 to 0.16 | 1 | 100 | 33.77* | Collett, Leroux et al. (2010); Richman and Harper (1978) |
| Visuospatial | 1 | 42 | −0.83 | | 3 | 53 | | Laasonen et al. (2004) |
| CLP | | | | | | | | |
| Memory—immediate | 1 | 30 | −0.63* | | 2 | 62 | | Conrad et al. (2009) |
| Language | 5 | 111 | −0.48* (0.12) | −0.60 to −0.36 | 7 | 67 | 0 $^{NS}$ | Jocelyn et al. (1996); Smith and McWilliams (1966); Snyder and Scherer (2004) |
| Attention/Executive | 4 | 96 | −0.27* (0.20) | −0.41 to −0.13 | 1 | 79 | 0 $^{NS}$ | Conrad et al. (2009); Laasonen et al. (2004); Lemos and Feniman (2010); Smith and McWilliams (1966) |
| Motor | 3 | 43 | −0.71 (0.11) | −1.45 to 0.03 | 8 | 57 | 0 $^{NS}$ | Jocelyn et al. (1996); Scherer et al. (2008); Snyder and Scherer (2004) |
| General Cognition | 8 | 287 | −0.40* (0.70) | −0.48 to −0.28 | 8 | 73 | 63.50* | Conrad et al. (2009); Goodstein (1961); Jocelyn et al. (1996); Murray et al. (2008); Pecyna et al. (1987); Scherer et al. (2008); Snyder and Scherer (2004); Speltz et al. (2000) |
| Academic | 2 | 39 | 0.21 (0.25) | 0 to 0.42 | 1 | 85 | 0 $^{NS}$ | Collett, Leroux et al. (2010); Stackhouse (1982) |
| Visuospatial | 1 | 14 | −0.22 | | 1 | 85 | | Laasonen et al. (2004) |

*Note.* $N_{Studies}$ = number of studies included in analysis; $N_{Participants}$ = number of participants in cleft group; *M* $d_w$ = mean weighted effect size; *SD d* = standard deviation of effect size, NS = Nonsignificant, $^*p < .05$.

academic ability, visuospatial ability, and immediate memory. ($N_{studies} = 1–3$). Processing speed, sensorimotor functions, and delayed memory were not investigated by any of these studies. Once again, the CLP samples varied considerably in size ($N_{participants} = 14–287$). As seen in Table IV, people with a CLP showed moderate and significant impairments in immediate memory (based on only one study) and language. The findings for the remaining cognitive domains did not meet the study criteria. Therefore, CLP only appears to be associated with impaired language skills.

## Discussion

The current study analyzed data from a total of 29 studies that have examined the cognitive functioning of participants with nonsyndromal clefts, compared to that of age-matched controls, in order to determine the extent to which cognitive functioning is compromised by a cleft. The final data set included 1,546 persons with clefts (642 males, 305 females) with a mean age of 10 years ($SD = 8$) and a control group of 279,805 persons. The higher number of males with clefts is consistent with previous reports indicating a higher prevalence of cleft conditions in males (Genisca et al., 2009). Notably, people with a cleft of any kind (CL, CP or CLP) performed significantly worse on 7 out of 10 cognitive domains. However, only the moderate and significant deficit in the language domain met the criteria for study homogeneity and publication bias; making this the most compelling finding.

Only a small number of studies ($N < 3$) examined five of the 10 cognitive domains: processing speed, immediate memory, delayed memory, sensorimotor function, and visuospatial ability, limiting our ability to draw conclusions regarding these domains. As heterogeneity analyses have limited power when they are based on a small number of effects, we may have failed to detect heterogeneity in some cognitive domains. Moreover, the significant heterogeneity statistics for many cognitive domains may reflect the fact that a variety of different tests were used to assess each domain, particularly attention/executive functions. In addition, this heterogeneity may reflect the inclusion of studies that used both toddlers and older children and young adults. However, at present, there are insufficient data to analyze these age groups separately or to examine the impact of age on cognitive function in persons with a cleft. Longitudinal studies are needed in order to clarify the impact of a cleft on an individual's cognitive and motor development from infancy through to adulthood. At this stage, almost all studies have been cross-sectional or only followed participants over a few years. A further issue limiting our ability to draw firm conclusions from these findings was the low $N_{fs}$ for many of the cognitive domains. This is likely to have resulted from the small number of studies that have been published to date combined with small effect sizes, which reduced the robustness of the findings.

It has been recently suggested that etiological differences between the different cleft types may lead to differences in the neuro-development (Genisca et al., 2009) and cognitive functioning (e.g. Persson et al., 2008) of these groups. Consequently, we additionally examined cognitive impairments by cleft type. Importantly, the sample sizes in this part of the analysis were low: the mean number of participants was 10 for CL, 52 for CP and 30 for CLP. All three cleft types were associated with cognitive impairments, although the profile of deficits appears to differ. Language skills appear to be affected by CL and CLP, general cognition by CL, and immediate memory by CP and CLP. However, given the limitations of this analysis, these should only be considered preliminary findings as the only impairment by cleft-type finding that met all of the statistical criteria was the language impairment found for CLP.

The current study highlights the need for large-scale studies that include more participants for each cleft group (CL, CP, CLP) and use common measures to assess cognitive functioning across the full range of areas of cognition to clarify our understanding of the impact that a cleft has on a person's development. While participants in the current analysis were generally well-matched for age, it is also important to match for gender and other potential confounding variables (e.g., socioeconomic status, cleft severity). For example, clefts are more prevalent in lower socioeconomic groups (e.g., Durning, Chestnutt, Mogan, & Nester, 2006), making it important to additionally match samples on this variable. While more recent studies have controlled for the lower socioeconomic status of the cleft groups in their analyses (e.g., Conrad et al., 2009), older studies have not. Unfortunately, there were insufficient studies to consider this issue further. Indeed, many studies that were included in the current analysis did not report socioeconomic status. In addition, clefts can vary in severity, which may also impact on cognitive functioning, and should also be considered. Despite these limitations, this meta-analysis provides the first quantitative review of research in this area. Some of the strengths of this work include the fact that a careful, methodical, and extensive review of the literature were undertaken, there were clear inclusion criteria, and effect sizes were calculated, and it served to integrate the findings

of 29 studies in a way that is not possible in a narrative review.

It is worth noting that only the language impairments for the combined cleft groups and the CLP group alone met all of the statistical criteria. While it may be the case that hearing and speech deficits, together with the psychological effects of living with a facial disfigurement, predispose persons with clefts to developing specific language problems, moderate to large deficits in other cognitive functions were also frequently observed. However, the latter results were inconclusive because there were too few studies, low $N_{fs}$ statistics (which partly reflects the small number of studies), and significant heterogeneity in the data; suggesting the need for further research, which additionally examines some of the moderating variables that contribute to this heterogeneity.

In terms of the practical implications for clinicians working in this area, these results suggest that language functioning should be routinely assessed in all persons with a cleft, appropriate interventions provided, and outcomes closely monitored, including the impact of language impairments on the post school training and employment success of adults with clefts. At present, there is insufficient evidence regarding the other cognitive domains to provide clear advice but, where possible, a cautious approach would be to assess the full range of cognitive abilities, pending further research.

Future researchers should ensure that studies are conducted and reported in a way that will enable their data to be meta-analyzed by including descriptive data (means, SDs) for all tests (regardless of statistical significance) and broken down by cleft type. In addition, information relating to potential moderating variables (e.g., age, gender, age of cleft repair) should be reported to enable an analysis of their impact on cognitive outcome. It is a real concern that even very basic information, such as age and gender, were not reported in a substantial number of studies ($N = 12$ and $N = 7$ from a total of 29 studies, respectively). In addition, many studies did not report data by cleft type; this information should be routinely provided, along with demographic information. Finally, socioeconomic status is known to be related to clefting (e.g., Clark, Mossey, Sharp, & Little, 2003) and should, therefore, always be reported.

Given that clefts are one of the most common birth defects, it is surprising that the research conducted over the last 50 years has not yet reached a point where basic questions regarding the cognitive profiles of people with clefts can be answered confidently. This knowledge is needed in order to provide a sound foundation upon which clinical work with this population can be based.

## Supplementary Data

Supplementary data can be found at: http://www.jpepsy.oxfordjournals.org/.

## Funding

This study was supported by a grant from the Faculty of Health Sciences, University of Adelaide.

*Conflicts of interest:* None declared.

## References

Anderson, K. N. (Ed.), (1998). *Mosby's medical, nursing and allied health dictionary* (5th ed.). St Louis: Mosby.

*Brennan, D. G., & Cullinan, W. L. (1974). Object identification and naming in cleft palate children. *Cleft Palate Journal, 11*, 188–195.

Broder, H. L., Richman, L. C., & Matheson, P. B. (1998). Learning disability, school achievement, and grade retention among children with cleft: A two-center study. *Cleft Palate-Craniofacial Journal, 35*, 127–131.

*Broen, P. A., Devers, M. C., Doyle, S. S., Prouty, J. M., & Moller, K. T. (1998). Acquisition of linguistic and cognitive skills by children with cleft palate. *Journal of Speech, Language, and Hearing Research, 41*, 676–687.

Canfield, M. A., Honein, M. A., Yuskiv, N., Xing, J., Mai, C. T., Collins, J. S., ... Kirby, R. S.; the National Birth Defects Prevention Network. (2006). National estimates and race/ethnic-specific variation of selected birth defects in the United States, 1999-2001. *Birth Defects Research, Part A, 76*, 747–756.

Chinn, S. (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. *Statistics in Medicine, 19*, 3127–3131.

Clark, J. D., Mossey, P. A., Sharp, L., & Little, J. (2003). Socioeconomic status and orofacial clefts in Scotland, 1989 to 1998. *Cleft Palate-Craniofacial Journal, 40*, 481–485.

Cohen, J. (1992). Quantitative methods in psychology: A power primer. *Psychological Bulletin, 112*, 155–159.

*Collett, B. R., Leroux, B., & Spelz, M. L. (2010). Language and early reading among children with orofacial clefts. *Cleft Palate-Craniofacial Journal, 47*, 284–292.

*Collett, B. R., Stott-Miller, M., Kapp-Simon, K. A., Cunningham, M. L., & Speltz, M. L. (2010). Reading in children with orofacial clefts versus controls. *Journal of Pediatric Psychology, 35*, 199–208.

*Conrad, A. L., Richman, L., Nopoulos, P., & Dailey, S. (2009). Neuropsychological functioning in children with non-syndromic cleft of the lip and/or palate. *Child Neuropsychology, 15*, 471–484.

*Conrad, A. L., Canady, J., Richman, L., & Nopoulos, P. (2008). Incidence of neurological soft signs in children with isolated cleft of the lip or palate. *Perceptual and Motor Skills, 106*, 197–206.

De Soussa, A. D., Devare, S., & Ghanshani, J. (2009). Psychological issues in cleft lip and cleft palate. *Journal of the Indian Association of Paediatric Surgery, 14*, 55–58.

Durning, P., Chestnutt, I. G., Morgan, M. Z., & Nester, N. J. (2007). The relationship between orofacial clefts and material deprivation in Wales. *Cleft Palate-Craniofacial Journal, 22*, 203–207.

*Ebert, P. R., McWilliams, B. J., & Woolf, G. (1974). A comparison of the written language ability of cleft palate and normal children. *Cleft Palate Journal, 11*, 17–20.

*Eide, M. G., Skjaerven, R., Irgens, L. M., Bjerkedal, T., & Oyen, N. (2006). Associations of birth defects with adult intellectual performance, disability and mortality: Population-based cohort study. *Pediatric Research, 59*, 848–853.

*Fox, D., Lynch, J., & Brookshire, B. (1978). Selected developmental factors of cleft palate children between two and thirty-three months of age. *Cleft Palate Journal, 15*, 239–245.

Genisca, A. E., Frias, J. L., Broussard, C. S., Honein, M. A., Lammer, E. J., Moore, C. A., . . . Rasmussen, S. A.; The National Birth Defects Prevention Study. (2009). Orofacial clefts in the National Birth Defects Prevention Study, 1997–2004. *American Journal of Medical Genetics, Part A, 149*, 1149–1158.

*Goldsberry, G., O'Leary, D., Hichwa, R., & Nopoulos, P. (2006). Functional abnormalities in the neural circuitry of reading in men with nonsyndromic clefts of the lip or palate. *Cleft Palate Craniofacial Journal, 43*, 683–690.

*Goodstein, L. D. (1961). Intellectual impairment in children with cleft palates. *Journal of Speech and Hearing Research, 4*, 287–294.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press Inc.

Higgins, J. P. T., & Green, S. (Eds.), (2011). *Cochrane handbook for systematic reviews of interventions version 5.1.0*. The Cochrane Collaboration. Retrieved from www.cochrane-handbook.org

Hunt, O., Burden, D., Hepper, P., & Johnston, C. (2005). The psychosocial effects of cleft lip and palate: A systematic review. *European Journal of Orthodontics, 27*, 274–285.

Jocelyn, L. J., Penko, M. A., & Rode, H. L. (1996). Cognition, communication, and hearing in young children with cleft lip and palate and in control children: A longitudinal study. *Pediatrics, 97*, 529–534.

Kapp-Simon, K. A., & Krueckeberg, S. (2000). Mental development in infants with cleft lip and/or palate. *Cleft Palate-Craniofacial Journal, 37*, 65–70.

*Laasonen, M., Haapanen, M. L., Maenpaa, P., Pulkkinen, J., Ranta, R., & Virsu, V. (2004). Visual, auditory, and tactile temporal processing in children with oral clefts. *Journal of Craniofacial Surgery, 15*, 510–518.

*Lamb, M. M., Wilson, F. B., & Leeper, H. A. Jr. (1972). A comparison of selected cleft palate children and their siblings on the variables of intelligence, hearing loss, and visual-perceptual-motor abilities. *Cleft Palate Journal, 9*, 218–228.

*Lemnos, I. C. C., & Feniman, M. R. (2010). Sustained Auditory Attention Ability test (SAAAT) in seven-year-old children with cleft lip and palate. *Brazilian Journal of Otorhinolaryngology, 76*, 199–205.

*Lewis, R. (1961). A survey of the intelligence of cleft palate children in Ontario. *Cleft Palate Bulletin, 11*, 83–85.

*Lewis, R. (1971). Survey of the intelligence of cleft-lip and cleft-palate children in Ontario. *British Journal of Disorders of Communication, 6*, 17–25.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications Inc.

*Lovius, B. B. J. (1971). Speech and intelligence in adult cleft-palate patients. *Dental Practitioner and Dental Record, 21*, 290–293.

Monro, I. R. (1995). A description of craniofacial anomalies: The mechanism and rationale of surgery. In R. Eder (Ed.), *Craniofacial anomalies: Psychological perspectives* (pp. 3–21). New York: Springer.

*Murray, L., Hentges, F., Hill, J., Karpf, J., Mistry, B., Kreutz, M., . . . Goodacre, T.; the Cleft Lip and Palate Study Team. (2008). The effect of cleft lip and palate, and the timing of lip repair on mother-infant interactions and infant development. *Journal of Child Psychology and Psychiatry, 49*, 115–123.

*Nakajima, T., Mutsudome, A., & Yosikawa, A. (2001). Postoperative speech development based on cleft types in children with cleft palate. *Pediatrics International, 43*, 666–672.

*Nopoulos, P., Berg, S., Van Demark, D., Richman, L., Canady, J., & Andreasen, N. C. (2001). Increased incidence of a midline brain anomaly in patients with nonsyndromic clefts of the lip and/or palate. *Journal of Neuroimaging, 11*, 418–424.

*Nopoulos, P., Berg, S., Van Demark, D., Richman, L., Canady, J., & Andreasen, N. C. (2002). Cognitive dysfunction in adult males with non-syndromic clefts of the lip and or palate. *Neuropsychologia, 40*, 2178–2184.

Nopoulos, P., Langbehn, D. R., Canady, J., Magnotta, V., & Richman, L. (2007). Abnormal brain structure in children with isolated clefts of the lip or palate. *Archives of Pediatric Adolescent Medicine, 161*, 753–758.

*Pannbacker, M. (1975). Oral language skills of adult cleft palate speakers. *Cleft Palate Journal, 12*, 95–106.

*Pecyna, P. M., Feeney-Giacoma, M. E., & Neiman, G. S. (1987). Development of the object permanence concept in cleft lip and palate and noncleft lip and palate infants. *Journal of Communication Disorders, 20*, 233–243.

Persson, M., Becker, M., & Svensson, H. (2008). General intellectual capacity of young men with cleft lip with or without cleft palate alone. *Scandinavian Journal of Plastic and Reconstructive Surgery and Hand Surgery, 42*, 14–16.

*Philips, B. J., & Harrison, R. J. (1969). Language skills of preschool cleft palate children. *Cleft Palate Journal, 6*, 108–119.

*Richman, L., & Harper, D. (1978). School adjustment of children with observable disabilities. *Journal of Abnormal Child Psychology, 6*, 11–18.

*Richman, L. C. (1976). Behavior and achievement of cleft palate children. *Cleft Palate Journal, 13*, 4–10.

Richman, L. C., & Eliason, M. J. (1986). Development in children with cleft lip and/or palate: Intellectual, cognitive, personality, and parental factors. *Seminars in Speech and Language, 7*, 225–239.

Rosenthal, R. (1995). Writing meta-analytic reviews. *Psychological Bulletin, 118*, 183–192.

*Ruess, A. L. (1965). A comparative study of cleft palate children and their siblings. *Journal of Clinical Psychology, 21*, 354–360.

*Scherer, N. J., & D'Antonio, L. L. (1995). Parent questionnaire for screening early language development in children with cleft palate. *Cleft Palate Craniofacial Journal, 32*, 7–13.

*Scherer, N. J., Williams, A. L., & Proctor-Williams, K. (2008). Early and later vocalization skills in children with and without cleft palate. *International Journal of Pediatric Otorhinolaryngology, 72*, 827–840.

Shipster, C., Hearst, D., Dockrell, J. E., Kilby, E., & Hayward, R. (2002). Speech and language skills and cognitive functioning in children with Apert syndrome: A pilot study. *International Journal of Language and Communication Disorders, 37*, 325–343.

*Smith, R. M., & McWilliams, B. J. (1966). Creative thinking abilities of cleft palate children. *Cleft Palate Journal, 3*, 275–283.

*Snyder, L. E., & Scherer, N. (2004). The development of symbolic play and language in toddlers with cleft palate. *American Journal of Speech-Language Pathology, 13*, 66–80.

*Speltz, M. L., Endringa, M. C., Hill, S., Marris, C. L., Jones, K., & Omnell, M. L. (2000). Brief reports: Cognitive and psychomotor development of infants with orofacial clefts. *Journal of Pediatric Psychology, 25*, 185–190.

*Stackhouse, J. (1982). An investigation of reading and spelling performance in speech disordered children. *British Journal of Disorders of Communication, 17*, 53–60.

Starr, P., Chinsky, R., Canter, H., & Meier, J. (1977). Mental, motor, and social behavior in infants with cleft lip and/or cleft palate. *Cleft Palate Journal, 14*, 140–146.

Sutton, A. J., Abrams, K. R., Jones, D. R., Sheldon, T. A., & Song, F. (2000). *Methods for meta-analysis in medical research*. Chichester: John Wiley & Sons, Ltd.

Turner, H. M., & Bernard, R. M. (2006). Calculating and synthesizing effect sizes. *Contemporary Issues in Communication Science and Disorders, 33*, 42–55.

Zakzanis, K. K. (2001). Statistics to tell the truth, the whole truth, and nothing but the truth: Formulae, illustrative numerical examples, and heuristic interpretation of effect size analyses for neuropsychological researchers. *Archives of Clinical Neuropsychology, 16*, 653–667.

Zakzanis, K. K., Leach, L., & Kaplan, E. (1999). *Neuropsychological differential diagnosis*. Lisse, The Netherlands: Swets & Zeitlinger.

** References marked with an asterisk indicate studies included in the meta-analysis*.