

SPECIAL EDUCATION PERSPECTIVES

Analysis and Critique of the Advocacy Paper *Towards Inclusive Education: A Necessary Process of Transformation*[†]

Jennifer Stephenson^{1*}  and Rahul Ganguly² 

¹Macquarie University, Australia, and ²Charles Sturt University, Australia

*Corresponding author. Email: jennifer.stephenson@mq.edu.au

(Received 28 April 2021; revised 26 July 2021; accepted 8 August 2021)

Abstract

The increasing inclusion of students with disabilities in regular classes still leads to debate and many advocate for full inclusion of all students. Arguments for full inclusion are generally rights-based, but proponents also claim research supports the effectiveness of full inclusion over specialist provision for all students with disabilities. In this article, we analyse and critique the use of the research literature in an Australian advocacy paper as an example of the broad claims made concerning full inclusion. We examine the extent to which the sources used provide conclusive evidence about the merits of full inclusion. We find the advocacy paper relies heavily on opinion and non-peer-reviewed literature, with little use of quantitative research that compares outcomes for students in different settings. We suggest that policymakers should treat the conclusions drawn in this paper cautiously and give due consideration to the literature that is not supportive of full inclusion.

Keywords: inclusion; schools; advocacy; disability

Australia has been part of the global movement towards the inclusion of students with disabilities into regular schools and classes, having signed the Salamanca Agreement in 1994 (United Nations Educational, Scientific and Cultural Organization) and the United Nations Convention on the Rights of Persons With Disabilities in 2008 (United Nations Committee on the Rights of Persons with Disabilities, 2006). The Disability Standards for Education were introduced in 2005 (Australian Government, 2005), requiring schools to make appropriate adjustments to curriculum and pedagogy for students with disabilities to enable them to access education on the same basis as other students. Despite the national commitment, the states differ in approaches to inclusion, in definitions of disability, and in resourcing, and all states retain specialised settings for some students (Anderson & Boyle, 2019).

Debate continues, sometimes passionately, about the merits of full inclusion in mainstream classes for all students, including those with the most severe and complex disabilities, compared to specialised or a mix of specialised and mainstream provisions designed to meet the needs of specific individual students (Carter & Abawi, 2018; Kauffman et al., 2020). In Australia, there are several groups advocating for full inclusion and the closure of all specialised settings, including groupings allowing for targeted instruction within classrooms (Australian Coalition for Inclusive Education, n.d).

Arguments for full inclusion from advocacy groups and others generally derive from a social justice and rights-based approach, often appealing to Article 24 of the United Nations Convention on the

[†]This manuscript was accepted under the Editorship of David Paterson.

Rights of Persons With Disabilities. Inclusion is seen as necessary for people to participate fully and actively in their communities and to be respected and valued as members of that community (Kliewer, 1998). Those opposed to full inclusion note the right of all students to appropriate and effective education (Kauffman et al., 2018; McLesky & Waldron, 2011), and the rights of parents to choose an appropriate education setting, and the rights of other students (Farrell, 2000). They also note that education is a means to an end, and full inclusion in a community after school may be possible after a specialised education (Farrell, 2000).

Proponents of inclusion use the research literature to make claims about the effectiveness of inclusion in the education of students with disabilities. Unfortunately, the interpretation of the inclusion research literature that reports student outcomes is recognised as being problematic in that theoretical positions, definitions of inclusion, outcome measures, participants, and circumstances differ across papers. In comparative research, where outcomes from settings with different degrees of inclusion are compared, the comparability of samples and settings may be questionable and the research may not be high quality (Dyson, 2014; Farrell, 2000; Florian, 2014; Foreman, 2015; Lindsay, 2007; Nilholm & Göransson, 2017). In addition, although proponents of inclusion claim research reviews generally support inclusion (e.g., Freeman & Alkin, 2000; Ruijs & Peetsma, 2009), other reviewers find the evidence is inconclusive (e.g., Göransson & Nilholm, 2014; Limbach-Reich, 2015).

In Australia, there are ongoing advocacy efforts to increase inclusion rates and to abolish specialist settings altogether. Given the ongoing debate about rights and evidence in relation to inclusion, and that policy development and implementation is influenced by both values and evidence (Lindsay, 2007), it is of interest to consider how research evidence is used in advocacy for full inclusion in Australia. One influential advocacy paper (Cologon, 2019) was produced by the organisation Children and Young People with Disability Australia (CYDA). CYDA is a not-for-profit organisation advocating for the rights of children and young people with disability, including the right to a fully inclusive education for all (CYDA, 2020a). It is a member of the Australian Federation of Disability Organisations and receives funding from the Australian Government. During 2019 and 2020, CYDA lobbied government departments and relevant ministers, participated in advisory committees, made submissions to inquiries and forums, and worked collaboratively with a large number of other organisations (CYDA, 2019, 2020b). CYDA also distributes fact sheets (<https://www.cyda.org.au/resource/fact-sheets?start=15>) based on Cologon (2019).

Cologon (2019) is an updated version of a similar paper by Cologon (2013). Both papers were produced under the auspices of CYDA, and funding was received from the Australian Government to support their production. The position taken in these papers is strongly rights-based, citing the Convention on the Rights of Persons With Disabilities (United Nations Committee on the Rights of Persons with Disabilities, 2006). Both papers provide a consideration of the outcomes of inclusive education as well as discussion of issues around defining inclusion, attitudes to inclusion, means to increase inclusion, and the implications of research for policy development related to inclusion. Cologon advocates strongly for full inclusion and draws on the document produced by the United Nations Committee on the Rights of Persons with Disabilities, General Comment No. 4 (GC4), which provides further detail and explanation of Article 24 of the Convention on the Rights of Persons With Disabilities (United Nations Committee on the Rights of Persons with Disabilities, 2016).

The position presented is that research demonstrates the superiority of inclusive settings for all students, and services such as intensive specialised instruction based on need for individual or small group lessons whether in specialised settings or mainstream schools are opposed because they represent forms of exclusion or micro-exclusion. All necessary supports should be ‘embedded inclusively within everyday practices’ (Cologon, 2019, p. 23).

The report has been widely disseminated and links to it are provided on the websites of other disability organisations lobbying for inclusion (e.g., Australian Alliance for Inclusive Education; <https://allmeansall.org.au/research/>). It is referred to as an ‘evidence review’ by the Australian Coalition for Inclusive Education, a coalition co-convened and chaired by CYDA (<https://acie.org>).

[au/2019/11/09/disability-royal-commission-education-and-learning-frequently-asked-questions-faqs/](https://www.cyda.gov.au/2019/11/09/disability-royal-commission-education-and-learning-frequently-asked-questions-faqs/)). CYDA claims its work is 'extensively cited' (2020b, p. 14) in government and parliamentary reports.

Because this CYDA paper makes a clear claim that full inclusion for all is supported by research, we provide a detailed and specific analysis and critique of this aspect of the advocacy paper. The particular focus of our analysis is the use of inclusion research in advocacy in relation to the outcomes of inclusion, which is addressed in the second chapter of Cologon (2019). In setting out to critique the use of research evidence in Cologon (2019), and more specifically in the chapter addressing the outcomes of inclusion, we asked the following research questions:

1. What are the characteristics of the sources relating to inclusion that are used in this paper?
2. To what extent do sources cited as supporting inclusion provide conclusive evidence for the claims made about benefits and drawbacks of full inclusion and of specialised settings?

Method

The first stage of the analysis was to examine the reference list for the whole document to establish the nature of all the sources cited. Each item in the reference list was classified as being an original research paper in a refereed journal or not. Research sources included research articles, research reviews, or meta-analyses. Furthermore, the type of research reported in refereed journals was classified as original quantitative research (group or single case studies reporting results as descriptive or analytic statistics, survey or questionnaire, or analysis of data not collected by the authors), original qualitative research (reporting on interviews, observations, documents, etc., and analysed by qualitative methods), or mixed methods research (using both quantitative and qualitative methods). Review articles were those that contained a method that clearly described the search process and selection criteria for articles included and were classified as to whether they reviewed quantitative research, qualitative research, or both. Meta-analyses were identified separately. Articles that were published in refereed journals but that did not report original research were placed in the discussion category and included discussion articles, opinion pieces, or reviews that did not include a formal search or criteria for article selection, and position papers. Other references grouped together as agency reports included reports produced by government and non-government agencies, policy documents, articles in non-refereed journals, and curriculum documents. Books and book chapters were grouped together and no further analysis of the use of books as sources was carried out. Conference papers and theses were also identified.

Reliability of article classification by the first author was established by the second author independently classifying a random sample (chosen using random.org) of 20% (81) of the references. The classifications of the first and second authors for this sample were compared and percentage reliability calculated according to the following formula: agreements divided by agreements plus disagreements multiplied by 100. Reliability for classification as original research or not was 91%. Reliability for the more detailed classification was 81.5%.

The next stage was to analyse the references used in Chapter 2 addressing the outcomes of inclusive education to determine the number of references used, the frequency of their use, and the characteristics of sources used to support the claims contained in this chapter. Each statement that was footnoted with a citation or citations from the reference list was extracted. Each source cited for each statement was listed and the number of different sources used in the chapter and the number of times each source was cited were calculated.

As there were 93 footnotes, many with more than one citation, we then used random.com to select a sample of 20% (19) of the footnotes for further analysis. The first determination was to ascertain the standard of evidence provided by the source(s) for each claim made. The standard of evidence was established according to the criteria in Table 1.

Table 1. Criteria Used to Determine the Standard of Evidence

Standard of evidence	Criteria
Definitive	Claim is consistent with very clear high-quality empirical evidence reported in results from experimental or survey research or with clear and consistent evidence in the case of reviews and meta-analyses.
Secondary source	Claim is supported by statements in a quantitative research study, quantitative review, or meta-analysis published in a refereed journal, but the statement does not come from the results reported; instead, it comes from the literature review or discussion.
Discussion paper or non-refereed source	Claim is supported by citing from a discussion or opinion article published in a refereed journal or from an unrefereed book or agency report.
Unsupported claim	The claim is not supported by the cited source. This category includes results from qualitative studies, as they are not designed to establish cause-and-effect relationships.

We also extracted relevant data from research studies, including the definition of inclusion used by the authors. For research studies other than reviews and meta-analyses, we also extracted information on research design (group studies with comparison groups, single case designs, survey or questionnaire, or other), participants (students with disabilities, typically developing students, regular classroom teachers, special educators, paraprofessionals, student teachers, parents/carers/family, other), country of study, and setting of study (early childhood, elementary/primary/middle school, secondary school). For research studies (quantitative and qualitative) that included participants in settings described as inclusive, we extracted the nature of any supports provided, and the nature and number of participants with disability. For discussion papers and agency reports, we extracted the definition of inclusion used in the source. It was impractical to analyse books.

We also analysed the sources cited five times or more in Chapter 2 by extracting data as described in the above paragraph. We made a determination of the level of evidence that could be provided generally by each of these sources, according to the criteria in Table 1, but taking into consideration the paper as a whole.

Reliability of coding and data extraction for the random sample, for the most frequently cited papers, and for additional coding for research papers including students with disabilities was established by independent coding by the authors. Reliability of coding and data extraction for the random sample was 84.7% and was 85.5% for the most cited papers. Reliability for additional coding of quantitative and qualitative papers that included students with disabilities as participants was 100% for the number and nature of students with disabilities. Reliability for the nature of supports was below 80%, primarily because information about supports was scattered through the articles and little detail was provided. Extraction of data on supports was completed by consensus coding between the authors.

Given the high number of sources cited that did not support the claims made, five sources judged not to support the claim they referred to were randomly selected using random.org. These are used as examples to illustrate the use of sources and are related to footnotes 30, 34, 42, 58, and 98.

Finally, we selected five broad, general claims that were made in Chapter 2 that we considered potentially misleading. The selected claims drawn from pages 9–11 are listed in Table 8.

Results

Complete Reference List

There were 404 items listed in the reference list for the whole paper. Of these, 208 (51.5%) were research papers published in refereed journals (93 original quantitative research, 76 qualitative research, 17 mixed methods, eight quantitative reviews, one qualitative review, four meta-analyses, and nine mixed reviews). The remaining items consisted of 109 discussion papers in refereed journals, 34 agency reports, 49 books, two conference papers, and two unpublished master's theses.

Table 2. Sources Cited Five Times or More in Chapter 2

Source	Number of times cited	Classification	Level of evidence
Hehir et al. (2016)*	19	Agency report	Discussion paper or non-refereed source
Kliwer (1998)	18	Discussion	Discussion paper or non-refereed source
Finke et al. (2009)	10	Qualitative research	Unsupported claim
Stahmer et al. (2003)	9	Qualitative research	Unsupported claim
Szumski et al. (2017)*	8	Meta-analysis	Definitive
Cologon (2012)*	8	Mixed research	Unsupported claim
Baker-Ericzén et al. (2009)*	8	Quantitative research	Definitive
Jordan et al. (2009)*	7	Discussion	Discussion paper or non-refereed source
Jordan et al. (2010)	7	Discussion	Discussion paper or non-refereed source
Giangreco et al. (1993)	5	Qualitative research	Unsupported claim
Cologon (2014)*	5	Book	Discussion paper or non-refereed source

Note. Items marked with an * were included in the random sample.

Table 3. Level of Evidence Provided by Random Sample of Sources in Chapter 2

Level of evidence	Number of sources
Definitive	7
Secondary source	2
Discussion paper or non-refereed source	19
Claim not supported	27

Chapter 2: All Sources

For Chapter 2, there were 93 footnotes with a total of 320 citations using 149 different sources. One footnote was not a citation and two citations could not be identified, as there were two items in the reference list with the same author and date, leaving 147 that could be coded. There were 96 (65.3%) research papers (47 original quantitative research papers, 29 original qualitative research papers, eight mixed methods, four quantitative reviews, five mixed reviews, and three meta-analyses), 36 discussion papers, six agency reports, eight books, and one conference paper. The most frequently cited sources (five uses or more) are shown in Table 2 and comprise 32.5% (104 out of 320) of the citations. None of the most cited papers included students with disabilities as participants, and only two were judged able to provide definitive support for a claim.

Random Sample From Chapter 2 Citations

In the sample randomly selected for more detailed analysis, 56 sources were cited with some sources used more than once. Sources used more than once supported different claims, so each use is counted separately. Only seven sources provided definitive evidence for the claim made, and 27 sources did not support the claim made. Table 3 provides summary statistics of the level of evidence provided by the 56 sources.

Table 4. Participants, Country, and Settings of Research Studies in Chapter 2 Sample

Details of study	Number
Participants	
Students with disabilities	12
Students who are typically developing	11
Regular classroom teachers	11
Special educators	3
Teacher aides (paraprofessionals)	1
Student teachers	1
Parents/carers/families	9
Other	6
Country of study	
Australia	5
Belgium	1
Canada	1
Greece	1
Netherlands	2
New Zealand	1
Switzerland	1
United Kingdom	5
United States of America	10
Setting of study	
Early childhood	5
Primary/elementary/middle school	16
Secondary school	7
Other	2

The majority of the 18 different quantitative studies in the random sample were surveys ($n = 11$), and two mixed research papers both included a survey and a qualitative component. Four studies compared groups of students in inclusive and specialised settings. There was one intervention study, and two papers analysed large datasets.

Table 4 provides information about the participants, country, and setting of quantitative ($n = 18$), qualitative ($n = 7$), and mixed ($n = 2$) research papers. Only half of these papers included students with disabilities as participants and most commonly studies were carried out in the United States in elementary/primary school settings.

The definitions of inclusion, when provided, varied considerably across the 18 quantitative papers, and it appears few, if any, students were fully included. Table 5 summarises the definitions.

The information provided about supports and special education services provided in the studies in the Chapter 2 sample that included students with disabilities in mainstream settings was also very variable (Table 6).

Table 5. Description of Inclusion in Quantitative Studies in Chapter 2 ($n = 18$)

Selected characteristics	n	%
Being in general education settings for more than 80% of the day	3	16.7
Receiving educational services in a general education classroom for 50% or more of the school day for at least one academic subject area	1	5.6
Enrolled in at least one general education classroom with support from paraprofessional	2	11.1
General education classroom (details not provided)	2	11.1
No information	10	55.6

Table 6. Special Education Services Provided in Inclusive Settings for Studies Including Students With Disabilities in Chapter 2 Sample ($n = 13$)

Services provided	Number of studies
Special educator support	6
Paraprofessional support	6
Therapy services	6
No support	1
No detail provided	4

Note. Some studies included students receiving more than one service.

Random Sample of Sources That Did Not Support the Claim Made

The sample of five sources that were judged not to support the claim made, along with the relevant claim and footnote number and reasons the claim was not supported, is provided in Table 7.

Claims Made in Chapter 2 for Which There Is Contradictory Evidence

Table 8 lists each claim and provides examples of contradictory evidence that should be considered when evaluating these general claims.

Discussion

The aim of the analysis of Cologon (2019) was to establish the characteristics of the sources used in advocating for full inclusion for students with disabilities and provide a basis for a critique of the evidence used to support the claims for full inclusion for all students. The implications of this analysis may go beyond this single advocacy paper. If other advocacy papers are similarly flawed, misrepresenting the research literature and making claims that appear to be unsupported, policymakers cannot use them for decision-making without further scrutiny of the relevant research. If policymakers and others use Cologon, and papers like it, as sources of evidence about the outcomes of inclusion, they need to consider how accurately they reflect the research evidence available.

In considering the certainty of evidence provided by this paper, we have analysed in some detail a sample of the sources used in the chapter relating to the outcomes of inclusion, as well as reviewing the sources used in the paper as a whole. Although it is claimed that research supports full inclusion, research articles (quantitative, qualitative, reviews, or meta-analyses) comprised little more than half of all sources cited for the whole paper. The percentage of research papers was higher for Chapter 2

Table 7. Claims, Sources, and Reasons Claim Is Not Supported

Claim (footnote number)	Source	Reason claim is not supported
'For students who do not experience disability, research finds that inclusive education results in: enhanced learning opportunities and experiences; education that is more sensitive to differing student needs; growth in interpersonal skills and greater acceptance and understanding of human diversity; and increased flexibility and adaptability'. (30)	Dessemonnet et al. (2012)	Study compared the progress of students with intellectual disability in regular classes with that of students with intellectual disability in special schools. The study was not about students who do not experience disability.
'However, despite the misunderstandings and the associated discrimination, research evidence tells a different story. In fact, aside from the ethical and philosophical concerns regarding excluding students who have been categorised as "too disabled" for inclusion, decades of research demonstrates that inclusive education has benefits for the academic, communication, and positive behavioural and social development of students labelled with "severe" and "multiple" "disabilities"'. (34)	Kurth & Mastergeorge (2012)	Study explored the effect of setting on participation patterns of adolescents with autism and mild intellectual disability (not severe and multiple disabilities). It did not measure any student outcomes related to setting.
'While there is some variation in individual studies, particularly based on teacher or parent ratings, growing evidence suggests that children and young people who attend "special" settings are more likely to experience bullying than their peers in "mainstream" settings, and that inclusive education is a key factor in reducing or eliminating bullying'. (42).	Woods & Wolke (2004)	Study assessed the relationship between bullying behaviour and National Curriculum Standard Assessment Tasks (SATs) test results and teacher assessments, and considered variables that predicted SATs and teacher assessment. All students were in regular classes; there was no comparison between settings.
'Given the importance of the role of teachers in relation to supporting positive behaviour, and the challenges that can be presented through behaviour, this is unsurprising. However, research investigating actual student behaviour compared with teacher concerns highlights a disconnect between the level of concern and the reality of classroom behaviour'. (58)	Sullivan et al. (2014)	A questionnaire study where teachers reported the frequency of problem behaviour of students, which behaviours were most difficult to manage, and what strategies they used to manage behaviour. Teachers did not report their level of concern about behaviour.
'Overall, research provides evidence that inclusive education results in higher quality education for students who do and do not experience disability'. (98)	Jordan & Stanovich (2001)	Study explored the relationship between teacher beliefs about working with exceptional and at-risk students and their interactions with those students and student self-concept. There was no comparison between settings.

(just over 65%), but just over 30% of these articles reported qualitative research, which cannot give evidence of cause-and-effect relationships. Similarly, there was only one quantitative research paper among the most cited sources, and the most frequently cited source (Hehir et al., 2016) was an unrefereed agency report that specifically set out to summarise research that 'demonstrates the benefits of inclusion' (Hehir et al., 2016, p. 2). The heavy reliance on opinion/discussion articles and unrefereed sources and the small number of research articles including students with disabilities in inclusive settings as participants is a concern when claims about outcomes are purported to be based on evidence.

Cologon (2019) is described as 'an extensive systematic literature review' that 'addresses the outcomes of inclusive education' (p. 4). No information about the search procedure, nor about the criteria for including sources in the review, is provided, and no further information could be obtained from the author or CYDA. A systematic review is usually considered as one that is conducted in a methodical way to reduce bias in the selection of articles to appraise. As such, it will include a search strategy that is described in sufficient detail to be reproducible, a predefined set of eligibility criteria,

Table 8. Generic Claims and Contradictory Evidence

Claim	Other evidence
<p>'Another common issue raised about inclusive education is the concern about the impact of including <i>all</i> students on the education of students who do <i>not</i> experience disability. However, contrary to common fears, research provides clear evidence of positive benefits of including everyone together, with no detrimental effects'.</p>	<p>Lindsay's (2007) review reported '[no] clear endorsement for the positive effects of inclusion' (p. 2); Dyson (2014) stated that existing research was 'unable to demonstrate that, when inclusive principles are turned into practice, learners benefit in some observable way' (p. 282). As noted in Cologon (2019), Fletcher (2010) found the presence of a student with emotional/behavioural difficulties in the early years of school decreases maths and reading scores of other students, and Gottfried (2014), for a similar population, found classmates of students with disabilities were more likely to have behaviour problems, less self-control, and poorer interpersonal skills. Reed et al. (2012) found students with autism in specialist classes made greater improvement in behaviour and socialisation than those in regular classes.</p>
<p>'... growing evidence suggests that children and young people who attend "special" settings are more likely to experience bullying than their peers in "mainstream" settings, and that inclusive education is a key factor in reducing or eliminating bullying'.</p>	<p>Rose et al. (2015) reported that students with autism spectrum disorder and learning difficulties were bullied more in inclusive settings, whereas students with intellectual disability and emotional and behaviour disorders were bullied more in specialised settings.</p>
<p>'The growing body of research into the outcomes of inclusive education for social development has also found that inclusion results in a more positive sense of self and self-worth for students who do and do not experience disability'.</p>	<p>Nepi et al. (2013), reporting on Italian schools where full inclusion has been practised for many years, found students with disabilities struggled to establish a social position, were less accepted, and did not feel connected to their school. Nugent (2008) found children with dyslexia who were in specialised settings were happier than those in inclusive settings and reported more positive experiences. Olsson et al. (2017) reported similar findings for deaf and hearing-impaired adolescents. Rathmann et al. (2018) reported that students with special education needs were more likely to report low life satisfaction when they attended mainstream schools. Crabtree and Meredith (2000) reported that students with intellectual disability in mainstream schools had lower self-perception in some areas than those in special schools, with no differences in other areas.</p>
<p>'In fact, research shows that students who experience disability who are educated in "mainstream" settings demonstrate better academic and vocational outcomes when compared with students educated in segregated settings'.</p>	<p>Foster and Pearson (2012) found students with autism in inclusive settings were no more likely to go on to college, not drop out, or have improved functional cognitive scores. Hienonen et al. (2021) found that although there was no difference in Finnish and maths learning between students with disabilities in mainstream and specialised settings, students in specialised settings had higher grade point averages and scored a little higher on some measures of learning motivation.</p>
<p>'... the lack of any research showing the superiority of segregation over 'mainstreaming' indicates that even poorly-done [<i>sic</i>] inclusion in the form of integration is still better'.</p>	<p>Carlberg and Kavale's (1980) review found positive effects of special class placement for students with emotional or behavioural disorders and those with learning difficulties. The papers already cited in this table show specific contexts where specialised settings had better outcomes than inclusive settings.</p>

and some assessment of the quality of research reviewed (Hanley & Cutts, 2013). Tools to improve the quality of systematic reviews are available, such as the PRISMA checklists (Moher et al., 2009) and the handbook for systematic reviews from the Cochrane Collaboration (Higgins & Green, 2008). One would expect that a methodical search would have turned up some of the studies illustrating positive outcomes in specialist settings, and some of the less positive findings on inclusion. It seems that little research reporting negative results, apart from Gottfried (2014) and Fletcher (2010) on the impact of students with emotional or behavioural disorders, was included. Discussion and opinion papers by vocal critics of full inclusion, such as James Kauffman, were absent, as were papers suggesting alternative views on rights-based arguments (e.g., Farrell, 2000).

When we looked more closely at a random sample of the sources used in Chapter 2, we found that only seven of the sources could provide definitive evidence for the claim made. There was a heavy reliance on discussion papers, which by their nature do not provide direct evidence, only the authors' opinions, and on non-refereed sources, which may not be of a high standard, or, like Hehir et al. (2016), offer only supportive evidence. Of most concern, we judged that 27 sources did not support the claims made for various reasons, such as not providing any empirical comparison data when claims were made that inclusive settings were superior or the article simply not addressing the claim made.

The relevance of the sources used for providing evidence about the outcomes of inclusion is also problematic. Claims about the relative benefits of inclusive compared to specialised settings must be based on studies that have compared outcomes for similar students in both settings, but there were very few sources proving this kind of data. None of the most cited sources and only four of the sources in the Chapter 2 sample included comparisons. Indeed, only 11 of the 29 research papers used in the sample from Chapter 2 included students with disabilities as participants. Göransson and Nilholm (2014) call this 'tumble-weed referencing' (p. 276), where research is used as though reliable evidence on the benefits of inclusion exists, but in actual fact the evidence is weak or non-existent.

When making blanket claims about the overall effectiveness of inclusion for all students, it is also important to look at the evidence in terms of the groups and contexts being examined. It may not be appropriate to generalise findings from country to country, from one disability group to another, or from setting to setting. The paper consistently cites sources describing specific students and contexts, with a preponderance of research from United States elementary schools, as providing generic evidence in favour of inclusion, but extreme caution must be used in using research findings in this way.

It is also important to consider the frequently noted problems with research on inclusion. Reviewers have found that definitions of inclusion vary considerably (Nilholm & Göransson, 2017), and the definitions of inclusion used in the 18 quantitative papers selected from Chapter 2 illustrate this variation. Ten papers (55%) gave no definition at all and others gave relatively vague definitions, such as more than 50% or more than 80% of the school day in mainstream classes. The Szumski et al. (2017) meta-analysis cited in the paper does not provide a definition of inclusion in the criteria given for selecting papers. Another aspect of the definition of inclusion is the nature of additional services and supports provided, and although these were not well described in the sample sources that included students with disabilities as participants, there were few studies where students did not receive support from a special educator or a paraprofessional within the mainstream class.

Cologon (2019) does address the limitations of the existing research and concludes that 'even poorly-done [*sic*] inclusion in the form of integration is still better [than specialised settings]' (p. 13). An alternative, and perhaps better explanation, given that few sources compared outcomes for students who were fully included at all times and who received no special education services with those in specialised settings, may be that students with disabilities do better when they are provided with access to both mainstream classes and to separate specialised settings and/or instruction. The support offered by several authors for inclusion within multi-tiered or response-to-intervention approaches to instruction is based on the evidence that students with disabilities may need intensive instruction delivered separately to regular class instruction (Fuchs & Fuchs, 2015). Others are of the view that the intensive and explicit instruction needed by some students with disabilities cannot be provided within regular classrooms (Kauffman et al., 2020), and certainly there are calls for more

research to establish effective tertiary practices, particularly for students with more severe disabilities (Snell, 2008). A huge research base on effective instruction for students with disabilities that has come from the special education research exists already (McLesky et al., 2017), and this literature has been almost entirely ignored in the paper, as have multi-tiered approaches and, indeed, separate small group or individual instruction is classified as micro-exclusion.

Finally, perhaps the reported experiences of students with disabilities themselves do not always endorse full inclusion. Nugent (2008) interviewed children with dyslexia who were in specialised and mainstream settings and their parents. All children were generally positive about their experiences, but those in specialised settings were happier and found the work easier. The author notes the importance of being with 'like' others, and specialised settings may provide benefits for friendships. Olsson et al. (2018) surveyed over 7,000 adolescents with deafness or hearing impairment and found those in specialised settings were more satisfied and felt more included than those in mainstream schools. Rathman et al. (2018) surveyed over 5,000 German students and found that students with disabilities attending mainstream settings were more likely to report low life satisfaction than those in specialised settings.

At a broader level, our analysis of Cologon (2019) may provide some guidance for readers of advocacy documents and policymakers who need to evaluate information about the research evidence presented by advocates of full inclusion. Readers should look for evidence that the review was systematic and followed accepted guidelines for searching for and selecting sources (Higgins & Green, 2008; Moher et al., 2009). A systematic review should include studies that reveal benefits from specialised settings and less favourable outcomes from inclusive settings. The nature of the sources used needs scrutiny; many authors express opinions in favour of full inclusion, but these opinions are not research evidence and should not be treated as such. Given the lack of quality comparative research, readers should be wary of claims that one setting produces better results, especially when details about the contexts and students compared are not presented. Readers should also be wary of sweeping claims about benefits for all based on studies of a particular age group, in a particular country, or in a particular context, especially when it seems very few studies involve students who are fully included without some kind of individual or small group support outside regular classes. Finally, readers should consider what cited papers mean by inclusion. We have noted the lack of consensus on definitions of inclusion, and not all students who are described as being included are fully included all the time without any individual or small group supports.

We believe the overriding right of children with disability is to receive an education that meets their needs and leads to an adult life that is as independent as possible. For some students, their needs are more likely to be met if all or part of their education is in specialist settings rather than restricted to what is possible in a mainstream classroom (Calhoun et al., 2019). There is no doubt there is research that demonstrates that inclusion can produce benefits for many students with disabilities, but several reviewers are critical of the quality and generalisability of this research given the lack of agreed definitions of inclusion (Dyson, 2014; Lindsay, 2007; Shaw, 2017). Research that demonstrates positive effects for specialist provisions (Fuchs & Fuchs, 2015) also needs to be acknowledged. Any evaluation of the desirability of moving to a full inclusion model of school education must consider the whole spectrum of research. It is important to identify what works as well as what does not work if all students are to receive the best education. Those who make policy decisions relating to education and those, such as parents and teachers, who are the targets of advocacy papers such as Cologon (2019) must be confident that such papers fully represent the research available.

References

- Anderson, J., & Boyle, C. (2019). Looking in the mirror: Reflecting on 25 years of inclusive education in Australia. *International Journal of Inclusive Education*, 23(7–8), 796–810. <https://doi.org/10.1080/13603116.2019.1622802>
- Australian Coalition for Inclusive Education. (n.d). *Who we are*. <https://acie.org.au/who-we-are/>
- Australian Government. (2005). *Disability Standards for Education 2005*. <https://www.dese.gov.au/disability-standards-education-2005>

- Baker-Ericzen, M. J., Garnand Mueggenborg, M., & Shea, M. M. (2009). Impact of trainings on child care providers' attitudes and perceived competence toward inclusion: What factors are associated with change? *Topics in Early Childhood Special Education, 28*(4), 196–208. <http://tec.sagepub.com/content/28/4/196>
- Calhoun, M. B., Berkeley, S., & Scanlon, D. (2019). The erosion of FAPE for students with LD. *Learning Disabilities Research & Practice, 34*(1), 6–13. <https://doi.org/10.1111/ldrp.12188>
- Carlberg, C., & Kavale, K. (1980). The efficacy of special versus regular class placement for exceptional children: A meta-analysis. *The Journal of Special Education, 14*(3), 295–309. <https://doi.org/10.1177/002246698001400304>
- Carter, S., & Abawi, L.-A. (2018). Leadership, inclusion, and quality education for all. *Australasian Journal of Special and Inclusive Education, 42*(1), 49–64. <https://doi.org/10.1017/jsi.2018.5>
- Children and Young People with Disability Australia. (2019). *Annual report 2019*. <https://www.cyda.org.au/about/who-we-are#2>
- Children and Young People with Disability Australia. (2020a). *What we do*. <https://www.cyda.org.au/about/what-we-do>
- Children and Young People with Disability Australia. (2020b). *Annual report 2020*. <https://www.cyda.org.au/about/who-we-are#2>
- Cologon, K. (2012). Confidence in their own ability: Postgraduate early childhood students examining their attitudes towards inclusive education. *International Journal of Inclusive Education, 16*(11), 1155–1173. <https://doi.org/10.1080/13603116.2010.548106>
- Cologon, K. (2013). *Inclusion in education: Towards equality for students with disability*. Children with Disability Australia. <https://web.archive.org/web/20170312230854/https://www.cyda.org.au/inclusion-in-education>
- Cologon, K. (2014). “Not just being accepted, but embraced”: Family perspectives on inclusion. In K. Cologon (Ed.), *Inclusive education in the early years: Right from the start* (pp. 91–114). Oxford University Press.
- Cologon, K. (2019). *Towards inclusive education: A necessary process of transformation*. Children and Young People with Disability Australia. <https://www.cyda.org.au/resources/details/62/towards-inclusive-education-a-necessary-process-of-transformation>
- Crabtree, J. W., & Meredith, C. (2000, October 5–6). *Self-concept and social comparisons in learning disabled students attending mainstream and special schools: Does integration have an impact?* Self-concept theory, research and practice: Advances for the new millennium conference, University of Western Sydney, NSW, Australia. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.195.6552&rep=rep1&type=pdf#page=195>
- Dessemontet, R. S., Bless, G., & Morin, D. (2012). Effects of inclusion on the academic achievement and adaptive behaviour of children with intellectual disabilities. *Journal of Intellectual Disability Research, 56*(6), 579–587. <https://doi.org/10.1111/j.1365-2788.2011.01497.x>
- Dyson, A. (2014). A response to Göransson and Nilholm. *European Journal of Special Needs Education, 29*(3), 281–282. <https://doi.org/10.1080/08856257.2014.933542>
- Farrell, P. (2000). The impact of research on developments in inclusive education. *International Journal of Inclusive Education, 4*(2), 153–162. <https://doi.org/10.1080/136031100284867>
- Finke, E. H., McNaughton, D. B., & Drager, K. D. R. (2009). “All children can and should have the opportunity to learn”: General education teachers' perspectives on including children with autism spectrum disorder who require AAC. *Augmentative and Alternative Communication, 25*(2), 110–122. <https://doi.org/10.1080/07434610902886206>
- Fletcher, J. (2010). Spillover effects of inclusion of classmates with emotional problems on test scores in early elementary school. *Journal of Policy Analysis and Management, 29*(1), 69–83. <https://doi.org/10.1002/pam.20479>
- Florian, L. (2014). What counts as evidence of inclusive education? *European Journal of Special Needs Education, 29*(3), 286–294. <https://doi.org/10.1080/08856257.2014.933551>
- Foreman, P. (2015). *Social justice principles, the law, and research, as bases for inclusion: An update*. https://www.education.vic.gov.au/Documents/about/department/psdlitreview_Socialjusticeprinciples_thelaw_and_research.pdf
- Foster, E. M., & Pearson, E. (2012). Is inclusivity an indicator of quality of care for children with autism in special education? *Pediatrics, 130*(Suppl. 2), S179–S185. <https://doi.org/10.1542/peds.2012-0900P>
- Freeman, S. F. N., & Alkin, M. C. (2000). Academic and social attainments of children with mental retardation in general education and special education settings. *Remedial and Special Education, 21*(1), 3–26. <https://doi.org/10.1177/074193250002100102>
- Fuchs, D., & Fuchs, L. S. (2015). Rethinking service delivery for students with significant learning problems: Developing and implementing intensive instruction. *Remedial and Special Education, 36*(2), 105–111. <https://doi.org/10.1177/0741932514558337>
- Göransson, K., & Nilholm, C. (2014). Conceptual diversities and empirical shortcomings – A critical analysis of research on inclusive education. *European Journal of Special Needs Education, 29*(3), 265–280. <https://doi.org/10.1080/08856257.2014.933545>
- Gottfried, M. A. (2014). Classmates with disabilities and students' noncognitive outcomes. *Educational Evaluation and Policy Analysis, 36*(1), 20–43. <https://doi.org/10.3102/0162373713493130>
- Hanley, T., & Cutts, L. (2013). What is a systematic review? [Editorial]. *Counselling Psychology Review, 28*(4), 3–6.

- Hehir, T., Grindal, T., Freeman, B., Lamoreau, R., Borquaye, Y., & Burke, S. (2016). *A summary of the evidence on inclusive education*. Instituto Alana. https://alana.org.br/wp-content/uploads/2016/12/A_Summary_of_the_evidence_on_inclusive_education.pdf
- Hienonen, N., Hotulainen, R., & Jahnukainen, M. (2021). Outcomes of regular and special class placement for students with special educational needs – A quasi-experimental study. *Scandinavian Journal of Educational Research*, 65(4), 646–660. <https://doi.org/10.1080/00313831.2020.1739134>
- Higgins, J. P. T., & Green, S. (Eds). (2008). *Cochrane handbook for systematic reviews of interventions*. Wiley-Blackwell.
- Jordan, A., Glenn, C., & McGhie-Richmond, D. (2010). The Supporting Effective Teaching (SET) project: The relationship of inclusive teaching practices to teachers' beliefs about disability and ability, and about their roles as teachers. *Teaching and Teacher Education*, 26(2), 259–266. <https://doi.org/10.1016/j.tate.2009.03.005>
- Jordan, A., Schwartz, E., & McGhie-Richmond, D. (2009). Preparing teachers for inclusive classrooms. *Teaching and Teacher Education*, 25(4), 535–542. <https://doi.org/10.1016/j.tate.2009.02.010>
- Jordan, A., & Stanovich, P. (2001). Patterns of teacher–student interaction in inclusive elementary classrooms and correlates with student self-concept. *International Journal of Disability, Development and Education*, 48(1), 33–52. <https://doi.org/10.1080/10349120120036297>
- Kauffman, J. M., Felder, M., Ahrbeck, B., Badar, J., & Schneiders, K. (2018). Inclusion of *all* students in general education? International appeal for a more temperate approach to inclusion. *Journal of International Special Needs Education*, 21(2), 1–10. <https://doi.org/10.9782/17-00009>
- Kauffman, J. M., Travers, J. C., & Badar, J. (2020). Why *some* students with severe disabilities are not placed in general education. *Research and Practice for Persons With Severe Disabilities*, 45(1), 28–33. <https://doi.org/10.1177/1540796919893053>
- Kliwer, C. (1998). The meaning of inclusion. *Mental Retardation*, 36(4), 317–322. [https://doi.org/10.1352/0047-6765\(1998\)036<0317:TMOI>2.0.CO;2](https://doi.org/10.1352/0047-6765(1998)036<0317:TMOI>2.0.CO;2)
- Kurth, J., & Mastergeorge, A. M. (2012). Impact of setting and instructional context for adolescents with autism. *The Journal of Special Education*, 46(1), 36–48. <https://doi.org/10.1177/0022466910366480>
- Limbach-Reich, A., (2015). Reviewing the evidence on educational inclusion of students with disabilities: Differentiating ideology from evidence. *International Journal of Child, Youth and Family Studies*, 6(3), 358–378. <https://doi.org/10.18357/ijcyfs.63201513560>
- Lindsay, G. (2007). Educational psychology and the effectiveness of inclusive education/mainstreaming. *British Journal of Educational Psychology*, 77(1), 1–24. <https://doi.org/10.1348/000709906X156881>
- McLeskey, J., Barringer, M.-D., Billingsley, B., Brownell, M., Jackson, D., Kennedy, M., Lewis, T., Maheady, L., Rodriguez, J., Scheeler, M. C., Winn, J., & Ziegler, D. (2017). *High-leverage practices in special education*. Council for Exceptional Children & CEEDAR Center. <https://highleveragepractices.org/store/books/high-leverage-practices-special-education>
- McLeskey, J., & Waldron, N. L. (2011). Educational programs for elementary students with learning disabilities: Can they be both effective and inclusive? *Learning Disabilities Research & Practice*, 26(1), 48–57. <https://doi.org/10.1111/j.1540-5826.2010.00324.x>
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLOS Medicine*, 6(7), Article e1000097. <https://doi.org/10.1371/journal.pmed.1000097>
- Nepi, L. D., Facondini, R., Nucci, F., & Peru, A. (2013). Evidence from full-inclusion model: The social position and sense of belonging of students with special educational needs and their peers in Italian primary school. *European Journal of Special Needs Education*, 28(3), 319–332. <https://doi.org/10.1080/08856257.2013.777530>
- Nilholm, C., & Göransson, K. (2017). What is meant by inclusion? An analysis of European and North American journal articles with high impact. *European Journal of Special Needs Education*, 32(3), 437–451. <https://doi.org/10.1080/08856257.2017.1295638>
- Nugent, M. (2008). Services for children with dyslexia – The child's experience. *Educational Psychology in Practice*, 24(3), 189–206. <https://doi.org/10.1080/02667360802256741>
- Olsson, S., Dag, M., & Kullberg, C. (2018). Deaf and hard-of-hearing adolescents' experiences of inclusion and exclusion in mainstream and special schools in Sweden. *European Journal of Special Needs Education*, 33(4), 495–509. <https://doi.org/10.1080/08856257.2017.1361656>
- Rathmann, K., Vockert, T., Bilz, L., Gebhardt, M., & Hurrelmann, K. (2018). Self-rated health and wellbeing among school-aged children with and without special educational needs: Differences between mainstream and special schools. *Research in Developmental Disabilities*, 81, 134–142. <https://doi.org/10.1016/j.ridd.2018.04.021>
- Reed, P., Osborne, L. A., & Waddington, E. M. (2012). A comparative study of the impact of mainstream and special school placement on the behaviour of children with autism spectrum disorders. *British Educational Research Journal*, 38(5), 749–763. <https://doi.org/10.1080/01411926.2011.580048>
- Rose, C. A., Stormont, M., Wang, Z., Simpson, C. G., Preast, J. L., & Green, A. L. (2015). Bullying and students with disabilities: Examination of disability status and educational placement. *School Psychology Review*, 44(4), 425–444. <https://doi.org/10.11705/spr-15-0080.1>

- Ruijs, N. M., & Peetsma, T. T. D. (2009). Effects of inclusion on students with and without special educational needs reviewed. *Educational Research Review*, 4(2), 67–79. <https://doi.org/10.1016/j.edurev.2009.02.002>
- Shaw, A. (2017). Inclusion: The role of special and mainstream schools. *British Journal of Special Education*, 44(3), 292–312. <https://doi.org/10.1111/1467-8578.12181>
- Snell, M. E. (2008). Rethinking effective instructional practices: A response to Copeland and Cosbey. *Research and Practice for Persons With Severe Disabilities*, 34(1), 228–231. <https://doi.org/10.2511/rpsd.33.4.228>
- Stahmer, A., Carter, C., Baker, M., & Miwa, K. (2003). Parent perspectives on their toddlers' development: Comparison of regular and inclusion childcare. *Early Child Development and Care*, 173(5), 477–488. <https://doi.org/10.1080/0300443032000088267>
- Sullivan, A. M., Johnson, B., Owens, L., & Conway, R. (2014). Punish them or engage them? Teachers' views of unproductive student behaviours in the classroom. *Australian Journal of Teacher Education*, 39(6), 43–56. <https://doi.org/10.14221/ajte.2014v39n6.6>
- Szumski, G., Smogorzewska, J., & Karwowski, M. (2017). Academic achievement of students without special educational needs in inclusive classrooms: A meta-analysis. *Educational Research Review*, 21, 33–54. <https://doi.org/10.1016/j.edurev.2017.02.004>
- United Nations Committee on the Rights of Persons with Disabilities. (2006). *Convention on the Rights of Persons With Disabilities (CRPD)*. <https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html>
- United Nations Committee on the Rights of Persons with Disabilities. (2016). *Convention on the Rights of Persons With Disabilities: General Comment No. 4 (2016) on the Right to Inclusive Education*. http://tbinternet.ohchr.org/_layouts/treatybodyexternal/Download.aspx?symbolno=CRPD/C/GC/4&Lang=en
- United Nations Educational, Scientific and Cultural Organization. (1994, June 7–10). *The Salamanca statement and framework for action on special needs education*. World Conference on Special Needs Education: Access and Equity, Salamanca, Spain. <https://unesdoc.unesco.org/ark:/48223/pf0000098427>
- Woods, S., & Wolke, D. (2004). Direct and relational bullying among primary school children and academic achievement. *Journal of School Psychology*, 42(2), 135–155. <https://doi.org/10.1016/j.jsp.2003.12.002>