691 The Esplanade,
Lota
Brisbane, Q4179
Australia

ISAAC Council and Executive Board

Dear Members of ISAAC,

According to the website *RationalWiki*, some "anti-science" groups argue that science can be shown to be in error because bumblebees should not be able to fly and kangaroos should not be able to jump.

RationalWiki's response is to observe that "*… it is possible to "*prove*" that* [a bumble bee can't fly and] *a kangaroo can't jump if you leave out a few key variables … but a full aerodynamic calculation (to say nothing of getting all empirical and watching a bumblebee fly) will show that the bumblebee's flight works perfectly fine.*"

In relation to FC we have seen researchers seek to "prove" that FC is not valid, but their experiments leave out "a few key variables". We are not suggesting this has been deliberate, but we do suggest that ignoring observational evidence is like refusing to believe in bumblebees and kangaroos. Observation of phenomena may sometimes precede understanding. This should not give rise to shrill accusations of "anti-science". Nor should it be met with shrieks of "heresy" when the observations do not fit with the received orthodoxy. If this occurs, we have learned nothing from the scientific revolution.

The history of human knowledge is full of examples where conventional wisdom is challenged by new thinking, leading to a clash of ideas, and to the old ideas falling from favour. Examples include the work of Copernicus, Galileo, Newton, Darwin and even Einstein. These scientists worked on the basis that they would propose a theory and **invite** the type of criticism that would help challenge and shape their ideas - even if that resulted in ultimate denial of their theory. The earlier members of this elite group worked on the basis of what they observed being in conflict with orthodox thinking of the time.

It is now accepted by science that where observations do not agree with existing research, both need to be seriously looked at.

FC has been developed independently by skilled teachers, therapists and parents all over the world, who knows how many times. The difference this time is that it has a history of over twenty years during which professional "best practice" has been shaped and developed. Researchers now have so much more to study.

Observation that people who have severe communication impairment may be able to communicate with FC support should be the stimulus for us to identify the "key variables" that make their communication possible. It should not be the stimulus to shut it down.

When we first received the ISAAC's draft position paper and the report on which it had been based, we pondered how we could best respond in a measured and considered manner. Fortunately, we discovered a process called EVIDAAC, a twenty point guide on how to review research on AACs. The Principal advisor on the development of EVIDAAC was the Chair of ISAAC's *ad hoc* FC Committee, Dr. Ralf Schlosser. One of ISAAC's "outside reviewers", Prof. Jeff Sigafoos, is also a Co-Director of EVIDAAC (Schlosser and Sigafoos, 2009).

We were also fortunate to come across a powerpoint presentation by Dr Schlosser and colleagues which helped us to understand the standards required to meet – or, more importantly, not meet - the criteria (Schlosser, Wendt and Boesch, 2009). Additional information has been taken from Schlosser, Wendt and Sigafoos (2007) and Schlosser and Sigafoos (2009).

We have not visited the EVIDAAC website itself as, according to Norton anti-virus software, it presents computer and identity threats.

The following document details an assessment of the *ad hoc* Committee's draft report against each of these criteria. The document comprises of a table which summarises why and how our responses have been reached followed by a set of Appendices that provide more detail and references to a range of authoritative sources from published material to support our conclusions.

Unfortunately our assessment results in a negative response to each question or standard. As such, we would submit that the entire document must be subject to serious question.

FC provides a perfect example where there is now ample observational and clinical evidence that it can work for some people. However there remains a school of thought that chooses to deny any evidence that does not totally accord with their world view. This is a belief system, not science.

It has been suggested that we should start afresh and look at individuals who have transitioned from requiring physical support to independent typing and even to speech and work backwards to understand what worked. We would commend this approach to the ISAAC Council and Executive Board as a way forward.

Finally the story of Gray's Paradox may add further illumination to our thinking:

In 1936, British zoologist Sir James Gray made an estimate of the power a dolphin could exert based on its physiology, and concluded that it was insufficient to overcome the drag forces in water to reach the speeds that had been observed. This was "Gray's Paradox". It took scientists until 2008 to work out how dolphins did it, but to my knowledge no-one ever suggested that the paradox was based on Gray's delusional need to believe in the impossible, nor did they expect dolphins to slow down. Observation that dolphins could swim fast provided the impetus to discover the "key variables" that had been missing from earlier calculations.

Of all organisations, surely ISAAC should be motivated to see this process through and start the research process afresh. Let dolphins swim as fast as they like, and let our kids communicate.

Yours sincerely,

Bill Kingston
(Member No. not yet issued.)

C.A. Davies
Cathie Davies
(Member No. 12860)

Executive Summary:

Items One to Six refer to the search process. Our response and the Appendices expand on the following conclusions:

- The research question or questions have not been clearly identified.
- There is no operational definition of the population or the intervention.
- The outcome measure chosen by the ad hoc Committee is inappropriate.
- The search method described gives the impression that a wide range of studies and other documents have been included. However, the process was open to source selection bias, publication bias and database bias.
- Search items used by the *ad hoc* Committee are excessively broad, likely to produce an unwieldy volume of material requiring subsequent filtering.
- A mismatch is apparent between described search methods and the surprisingly limited list of items identified on pages 4 and 5. This suggests that some studies were excluded before the explicit inclusion/ exclusion criteria were applied.
- No details of the filtering process have been made available. There is no evidence of a pre-existing search protocol document, and no transparent relationship between the search procedures described and the list of studies identified.

Items Seven to Ten refer to the criteria for inclusion and exclusion of studies. Our response and the Appendices expand on the following conclusions:

Inclusion and exclusion criteria should be pre-defined to avoid reviewers preferentially including studies that support their own prior conclusions.

- It is clear that some exclusions have occurred without explanation during the search phase. Other exclusions were "pre-defined":
    - Only written material is included. This is inappropriate because it precludes direct and video observation of "best practice" which, in the absence of operational definitions, can provide detailed information regarding the characteristics of FC users, the supports encompassed by FC; and the most meaningful outcomes.
    - Only peer reviewed material is included. This is inappropriate because the long history of dissonance been research and clinical observation will not be resolved without examination of a broader range of material.
    - Material "*that did not focus on FC or mentioned FC only in a tangential manner*" is excluded. This is inappropriate because, in the absence of quality research specifically targeting FC users, research on overlapping populations informs future research and current practice with clients who may benefit from FC.
- As only one review and three individual studies are ultimately considered, it is clear that further exclusions occurred during the "analysis" phase of the review. (Note – no inter-rater agreement data is available regarding these exclusions.)
    - Outcome and study design constraints result in exclusion of all material not deemed "Level 1" by the ad hoc Committee. This is inappropriate due to fundamental flaws in experimental design, and to the now widely recognised principle that assessments based on *a priori* criteria violate recommended practice principles by precluding consideration of individual needs.

- Reviews that had been subject to synopsis, and individual studies included in those reviews, are excluded from further consideration because they are deemed lower forms of evidence than the synopses. This is inappropriate because individual studies must be examined for quality and validity.

Items eleven and twelve refer to data extraction from included studies. Our response and the Appendices expand on the following conclusions:

It has been recognised that: "[i]*f … an outcome of an intervention contradicts the coder's expectation, he or she may not code other variables of this study accurately"* (Schlosser, Wendt and Sigafoos, 2007, p145).

- The synopsis of (Probst (2005) notes that "[t]*he author seems to have made decisions regarding … the extraction of data from included studies"* (Schlosser and Wendt, 2008, p83). The only studies coded for potential threats to validity were those that demonstrated some support for FC. Similar bias has been identified in previous reviews of FC research, with detailed analysis applied only to studies potentially supportive of FC (Tehan & Senior, 2006, p12).
- The lack of an independent second rater in Probst (2005) is identified by Schlosser and Wendt (2008, p83) as a weakness.
- In the ISAAC Review, *a priori* coding categories were used only when coding material against "criteria for inclusion", not for data extraction. The Report does not indicate who extracted the data, nor the process that was followed. Inadequate information is provided regarding inter-rater agreement and dispute resolution.
- Serious methodological shortcomings were noted in the three included individual studies. These shortcomings were taken to discredit only the two studies that showed some support for FC. Conclusions of the third study were considered sound because they agreed with other controlled studies, systematic reviews, and several position statements published by academic and professional groups (ISAAC, 2014, p10).

Differential treatment of studies on the basis of their conclusions appears to be straightforward evidence of bias based on the reviewers' expectations and personal viewpoint. Acceptance of a flawed study's conclusions because they agree with expert opinion turns conventional hierarchies of evidence on their head.

Item thirteen relates to the quality of included studies. Our response and the Appendix expand on the following conclusions:

As noted by the Chairman of the *ad hoc* Committee and colleagues in 2007: "*Because a systematic review can only be as sound as the included studies, an investigation of the quality of those studies is an important component of the data extraction and appraisal process"* (Schlosser, Wendt and Sigafoos, 2007, p145-6).

- The four studies ultimately included by the *ad hoc* Committee were not subjected to systematic quality appraisal.
- The only studies Probst (2005) examined for validity were those with results that implied a level of support for FC. There was no consideration of factors such as subject selection, treatment integrity, relevance of outcome measures, confounding factors related to intrusive control procedures, and other significant threats to validity that characterised all included studies.

- The *ad hoc* Committee's Report commented adversely on the quality of the three individual studies they include alongside Probst (2005). The only study whose conclusions they considered sound was the one that supported their own position.

My original submission to the ISAAC review stressed that it was of the utmost importance that ISAAC's reviewers examine original reports to assess the quality of experimental studies, rather than relying on "reviews" that uncritically collect and collate findings. There is no evidence that my submission was ever read.

Item fourteen relates to pre-definition and operationalization of "judgements of effectiveness". Our response and the Appendix expand on the following conclusions:

- Experiments are designed to test hypotheses regarding the relationship between independent and dependent variables. Thus, operational definition of "*judgements of effectiveness*" requires clear, operational definition of all variables.
- "Dependent variables" in the studies reviewed by the ad hoc Committee are measured as a percentage of "correct responses". I have argued in Appendix 1, Section 1.3 that this is not a legitimate measure of "communicative competence" for this population.
- "Independent variables" are attributes or characteristics that affect the outcome (or dependent variable). Experiments focus on the treatment variable, which in FC research is "*often manipulated under 'facilitator non-blind', 'facilitator-blind', 'facilitator double-blind', or 'communication without FC' (non-FC) conditions* (Schlosser and Wendt, 2008, p82).
- "Secondary independent variables" are any influences in the selection of participants, procedures, statistics, or design likely to affect the outcome and provide an alternative explanation for the results (Creswell, 2002, p621). Researchers seek to neutralise these through statistical or design procedures. However, as discussed in Appendix 1, the characteristics of the population from which subjects may legitimately be selected, and the nature of the supports and accommodations encompassed by the term "Facilitated Communication", are not yet adequately defined. Statistical and design procedures cannot address that which is not understood.
- In the absence of good quality research, "best practice" provides the most detailed information available regarding the supports encompassed by FC; the characteristics of FC users; and the outcomes that may be expected from use of the strategy.
- Professor Pat Mirenda has suggested a way forward:

  > *One way to approach this task is to identify people with ASD who have become competent (and independent) communicators through the use of AAC (including FC), and then to work backwards to answer the question "Are there common factors that appear to have contributed to these good outcomes?" If we find any (and I believe we will), we can then design longitudinal hypothesis-driven studies to examine these factors in natural contexts* (Mirenda 2008, p229).

Items fifteen to twenty relate to meta-analysis, and do not apply to this study.

Appraisal of the ISAAC ad hoc Committee Draft Report on Facilitated Communication

Bill Kingston (Membership Number not yet issued)
Cathie Davies (Membership Number 12860)

30 June 2014

Appraisal of the ISAAC *ad hoc* Committee Draft Report on Facilitated Communication

Appraisal Item	Rating
Item One: *The review addresses a clearly focussed question? (e.g. was there enough information on the population studied, the intervention given, the outcomes considered?* (Schlosser, Wendt and Boesch, 2009, p11). **Rationale:** • In Facilitated Communication (FC) research, adequate definitions of the population, the intervention, and meaningful outcomes are absent. "*A focused question will define not only the population but also the intervention and targeted outcomes. If even one of these is missing, mark 'No' "*. (Schlosser, Wendt and Boesch, 2009, p11). • The absence of an adequate definition of the population "*could affect the internal validity of the review, if the subjects were misdiagnosed in the original studies, or could affect the external validity because of difficultly determining the applicability of the evidence to specific clients*" (Schlosser, Wendt and Sigafoos, 2007, p143). • The complex, person-centred strategies and supports that make up Facilitated Communication, and the way they are flexibly applied by skilled facilitators, are yet to be operationally defined. Thus, it is not possible to "*determine the homogeneity of the studies and the efficacy of specific treatments" (Schlosser, Wendt and Sigafoos, 2007, p144).* The procedures implemented and subjected to investigation in the vast majority of experimental studies bear only superficial resemblance to FCT best practice. • Only one outcome measure is considered in the studies and reviews ultimately included by the ISAAC *ad hoc* Committee: "communicative competence", measured as a percentage of "correct responses" in message-passing or confrontational testing. There is no evidence, nor any attempt to demonstrate, that this is a valid test for the study population, nor of any individual's capacity for communication more broadly defined. • A directive from the ISAAC Council to the ISAAC Executive Board, ratified on July 29, 2012, significantly broadened the range of evidence to be considered: *The first modification is that the scope of the review will include both quantitative and qualitative studies. The second modification is that both published and unpublished sources will be included in the search and appraisal process in developing the draft* (ISAAC, 2013a). The *ad hoc* Committee has disregarded this directive and pursued an agenda formed around their unexamined assertion that "*the validity of FC stands and falls with the issue of authorship*", measured through confrontational testing under contrived conditions (ISAAC, 2014, p6). • While a broader research question or questions may not lead to a definitive position on FCT, a statement recognising the need for more research is undoubtedly preferable to a Position Statement that has been based on	No

seriously inadequate research. *Please see Appendix 1 for more detailed discussion.*	
Item Two: *The search methods were pre-defined (Schlosser, Wendt and Boesch, 2009, p12).* **Rationale:** • There is no evidence of a pre-existing search protocol document, nor that the search methods had been approved by any organisation, including ISAAC International, prior to implementation. • There is no transparent relationship between the methods described in the “Search for Studies and Reviews” section of the Report and the list of material ultimately considered by the *ad hoc* Committee. *Please see Appendix 2 for more detailed discussion.*	No
Item Three: *Multiple complementary sources that collectively minimize source selection bias are consulted. (Consider databases, hand searches, ancestry searches, contacting authors, forward citation sources.) (Schlosser, Wendt and Boesch, 2009, p12).* **Rationale:** Although the methods described under “Search for Studies and Reviews” (page 2-3) appear to be “multi-faceted”, it is clear that a very narrow approach has been taken. • A mismatch is apparent between the search methods described on pages 2 and 3 of the report, and the surprisingly limited list of items identified on pages 4 and 5. It would seem, either that the search methods described have not been followed, or that some items identified through the search process have been excluded without mention prior to application of the identified inclusion/exclusion criteria. • The vast bulk of submissions to the review by ISAAC members have either been explicitly or effectively excluded, or not referred to at all. • There is also reason to suspect publication bias and database bias, as discussed in relation to EVIDAAC Items 4 and 5. • In essence, the ad hoc Committee has, explicitly or effectively, excluded everything except a synopsis of a German-language review of FC and three individual studies that were not included in the German review. • The Chair of the *ad hoc* Committee co-authored the synopsis and the review was written by a member of the committee. • The synopsis was published in a journal edited by the Chair of the *ad hoc* Committee and one of the “outside readers”. • Authority to rely on a synopsis comes from an unconventional “hierarchy of evidence” described for AAC application in a paper authored by the Chair of the *ad hoc* Committee and the same “outside reader”. The hierarchy is expressly intended for use by busy practitioners seeking to make evidence-based decisions in their clinical practice. There is no indication that the hierarchy may be an appropriate approach for research.	No

It would, therefore, appear that the ISAAC Committee - while charged with undertaking "*a detailed examination of existing research and literature on FC …*" (ISAAC, 2012) - has done nothing to "*minimizes the oversight of studies outside the researcher's regular purview*" (Schlosser, Wendt and Sigafoos, 2007, p141). *Please see Appendix 3 for more detailed discussion.*	
Item Four *An attempt was made to locate unpublished studies (e.g. trial registers, dissertations etc.)* (Schlosser, Wendt and Boesch, 2009). **Rationale:** Despite: • a long-running disconnect between research and clinical observation that demands examination; • the well-recognised gate-keeper role of researchers and academics in peer-reviewed literature; and • the acknowledged phenomenon of publication bias, that is: ''*what occurs whenever the research that appears in the published literature is systematically unrepresentative of the population of completed studies''* (Rothstein, Sutton, & Borenstein, 2005, p 1, quoted in Schlosser, Wendt and Sigafoos, 2007, p141), the *ad hoc* Committee has summarily rejected any material not "peer-reviewed". This is despite the directive from the ISAAC Council to the ISAAC Executive Board, ratified on July 29, 2012, that instructed reviewers to consider a much wider range of material, including unpublished material. *Please see Appendix 4 for more detailed discussion.*	No
Item Five: *Databases are carefully selected so that they, together, minimize the potential of systematically excluding studies* (Schlosser, Wendt and Boesch, 2009). **Rationale:** • The long disconnect between research and clinical observation demands explanation. When the results of experimentation and experience cannot be reconciled, it is necessary to revisit both. • The "gatekeeping" role of researchers and academics in peer reviewed journals is particularly important in this scenario - even more so as Facilitated Communication draws little attention from researchers outside AAC. Contrary to accepted "hierarchies of evidence" in health and medical research, "expert opinion" may be more influential than critical evaluation of research when decisions are being made regarding publication. • The broader questions implied by ISAAC International's "*Adjustments to the Framework Within Which the FC Committee is Conducting its Efforts*" (ISAAC, 2013a) require a broader range of databases to be searched and a wider range of evidence to be considered if an unbiased review of evidence is genuinely being attempted. *Please see Appendix 5 for more detailed discussion.*	No

Item Six *The search items are stated and appropriate for each database (Schlosser, Wendt and Boesch, 2009, p14).* **Rationale** • The search terms advised by the *ad hoc* Committee are excessively broad, likely to produce an unwieldy volume of material requiring subsequent filtering. • More detailed definition of the research question or questions may have allowed identification of key search items suitable for a more targeted and transparent search process, however the research questions have not been articulated by the *ad hoc* Committee. (See Appendix 1, Section 1.4). • Any filtering of search results should be subject to scrutiny, in the same way that the overall search protocol should be subject to scrutiny (see my comments in respect of EVIDAAC Item 2). • The synopsis of the review on which the *ad hoc* Committee's report is primarily based (Probst, 2005) states that "*it would have been beneficial to report* [search terms] *for each specific database*" (Schlosser and Wendt, 2008, p83). In a later presentation the same authors (with a colleague) state that: "*A listing of search terms in general without cross-referencing it with certain databases would trigger a "no" response*" (Schlosser, Wendt and Boesch, 2009). This must reflect on the *ad hoc* Committee's report, which considered only three individual studies in addition to the synopsis of Probst (2005). *Please see Appendix 6 for more detailed discussion.*	No
Item Seven: *The criteria for inclusion and exclusion of studies are pre-defined (Schlosser, Wendt and Boesch, 2009).* **Rationale:** Inclusion and exclusion criteria should be pre-defined because: "*… if the authors state their inclusion criteria, it is less likely they will (as they are wont to do) preferentially cite studies that support their own prior conclusion"* (Oxman, Cook, & Guyatt, 1994, cited in Schlosser, Wendt and Sigafoos, 2007, p142). • It seems apparent that some exclusions occurred before initial coding took place (see Appendix 2). Criteria that may have informed this process are not identified. • The only constraints identified prior to coding were that material must be in written form, and material "*that did not focus on FC or mentioned FC only in a tangential manner*" would be excluded. • Exclusion of material that is not peer reviewed appears to have occurred after coding but before analysis. As only one review (or synopsis of that review) and three individual studies are ultimately considered by the ad hoc Committee, it is clear that further exclusions occurred during the "analysis" phase of the review: • Only one outcome measure is ultimately considered: percentage of "correct"	No

responses under controlled experimental conditions, from which conclusions regarding authorship are drawn. This outcome constraint dictates the study design constraint, which resulted in exclusion of all material not deemed "Level 1" by the ad hoc Committee. • Population and intervention constraints are not reported, because neither has been operationally defined. As discussed in respect of EVIDAAC Item 1, these shortcomings, together with the very dubious validity of the outcome constraint, raise serious questions over the quality of experimental studies. • Reviews that had been subject to synopsis, and individual studies included in those reviews, are excluded from further consideration because, under the 5-S hierarchy of evidence, they are deemed lower forms of evidence than the synopses. This is immensely significant, as it precludes the *ad hoc* Committee from examining original studies for quality and validity. These exclusions only helps to reinforce previously strongly expressed views of significant members of the *ad hoc* Committee and to exclude any "inconvenient truth" issues. • There were no apparent geographic constraints. Some linguistic and temporal constraints were pre-defined, others were not. These appear unlikely to have significantly impacted the outcome of the review. *Please see Appendix 7 for further discussion of this information.*	
Item 8: *The criteria for inclusion and exclusion are appropriate given the purpose of the review (Schlosser, Wendt and Boesch, 2009).* **Rationale:** Scope and selection bias is introduced through inappropriate inclusion and exclusion criteria: • *Exclusion of all but written material* precludes video and direct observation of "best practice" which, in the absence of quality research, may provide the most detailed information available regarding the characteristics of FC users, the supports encompassed by FC; and the most meaningful outcomes. • *Exclusion of material not about FC.* In the absence of quality research specifically targeting FC users, research on overlapping populations should be sought to inform future research and current practice in relation to clients who may benefit from FC. • *Exclusion of materials not peer reviewed.* The long history of dissonance been research and clinical observation demands explanation. The potential for publication and database bias is widely recognised, as is the gate-keeper role of researchers and academics in peer-reviewed literature and research institutions. • *Inclusion of material deemed "Level 1".* Exclusive consideration of studies that involve *a priori* confrontational testing appears to be in utter contempt of the ISAAC Council's July 2012 directive to broaden the range of submittable material. It is also highly inappropriate due to fundament flaws in experimental design (discussed in some detail in Appendix B of my original submission to ISAAC: Davies, 2012), and the now widely	No

recognised principle that assessments based on *a priori* criteria violate recommended practice principles by precluding consideration of individual needs. More appropriate outcome measures have been identified by peak bodies. See Appendix 1, Section 1.3.2. • *Exclusion of reviews that have been subject to synopsis, and individual studies included in those reviews.* Individual studies must be examined for quality and validity, instead of relying on reviews that merely collate the findings of fundamentally flawed experimental designs. • *Exclusion of material deemed "Level 2".* Level 2 studies offer legitimate, alternative quantitative approaches to examining authorship without relying on highly contrived experimental conditions or *a priori* confrontational testing or message passing tests. • *Exclusion of material deemed "Level 3" - qualitative research.* The serious flaws that render invalid most experimental designs to date can only be addressed through qualitative research. • *Exclusion of material deemed "Level 4" – anecdotal reports.* Practice based evidence can test "*the impact of an intervention in more typical settings and conditions*" (Lof, 2011, p.193). Also, anecdotal reports may point the way for future research. *Please see Appendix 8 for further discussion.*	
Item 9: *A log of rejected studies is reported/available upon request* (Schlosser, Wendt and Boesch, 2009). **Rationale:** • It is apparent that some exclusions occurred before initial coding took place (see Appendix 2). These include an estimated 224 items submitted by members in response to ISAAC International's invitation (ISAAC, 2013a), but never mentioned in the Report. Thus, there is no log of these rejected studies, nor any explanation given for their omission. • Pages 4 to 6 of the Report imply that 78 items were excluded for not being about FC or not peer reviewed, however this is misleading. As discussed in relation to EVIDAAC Items 7 and 8, further exclusions occurred during the "analysis" phase of the review, leaving only four papers that were ultimately analysed by the *ad hoc* Committee. *Please see Appendix 9 for further discussion.*	No
Item 10: *A reasonable percentage of studies (≥ 20%) is evaluated reliably for inclusion by more than one rater.* (Schlosser, Wendt and Boesch, 2009). **Rationale:** The *ad hoc* Committee report states that: "*The Chair of the ad hoc Committee coded all potential written documents found through the search or submitted to the ISAAC office for inclusion. A large percentage (approximately 60%) of written documents was independently coded by a second member of the ad hoc Committee*", (ISAAC, 2014, pp3-4).	No

• Confirmation that a large percentage of documents was independently coded by a second member of the *ad hoc* Committee provides no information regarding the decision-making process at other stages of the Review. As discussed in respect of EVIDAAC Items 7 to 9, it is clear that some exclusions occurred before coding and others after coding, during the analysis phase of the review. • No inter-rater agreement percentages are made available. • Inclusion of Probst (2005) is intended to capture data from 37 individual studies covered by that review. As it is afforded such prominence in the ad hoc Committee's study, is important to examine whether it meets the requirements of EVIDAAC Item 10. It does not. As noted in the synopsis of Probst: *The author seems to have made decisions regarding the inclusion or exclusion of studies … . The involvement of an independent second rater and the gathering of inter-rater agreement data would have allowed for a better assessment of the reliability of the data generated by this review* (Schlosser and Wendt, 2008, p83). *Please see Appendix 10 for further discussion.*	
Item 11 *The coding categories refer to the types of information or data that are being extracted from each included study. These categories need to be stated upfront/*a priori (Schlosser, Wendt and Boesch, 2009). **Rationale:** • There should be "*a clear outline of the process by which data was extracted from the original studies*" (Schlosser, Wendt and Sigafoos, 2007, p145). This should include details of who coded the data and how, as it has been recognised that: "[i]*f … an outcome of an intervention contradicts the coder's expectation, he or she may not code other variables of this study accurately*" (Schlosser, Wendt and Sigafoos, 2007, p145). • Bias has previously been identified in reviews of FC research, with detailed analysis reserved for studies potentially supportive of FC. "[V]*ery little attention is paid to the methodology or the validity of the results of the studies where negative effects have been found*" (Tehan & Senior, 2006, p12). All studies must be examined for quality (see Appendix 13). • Probst (2005) is intended by the *ad hoc* Committee to capture data from 37 individual studies, however the rationale behind data extraction is not fully explained. There was no second rater. (Schlosser and Wendt, 2008, p83). • Some individual studies reviewed by Probst (2005) produced results that may be interpreted as supportive of FC and, as has occurred in other reviews, only these have been examined for threats to validity. • In the ISAAC Review, *a priori* coding categories were used only when coding of material against "criteria for inclusion", not for data extraction. • The *ad hoc* Committee reviewed three studies in addition to Probst (2005) and noted serious methodological shortcomings in all three. While the shortcomings were taken to discredit the two studies that showed some	No

support for FC, conclusions of the third study were considered sound because they agreed with "*numerous other controlled studies*" and with "*systematic reviews and in several position statements published by academic and professional groups* (ISAAC, 2014, p10). • Differential treatment of studies on the basis of their conclusions appears to be straightforward evidence of bias based on the reviewers' expectations and personal viewpoint. • Acceptance of a flawed study's conclusions because they agree with expert opinion turns conventional hierarchies of evidence on their head. • Findings demonstrating a level of facilitator influence cannot be interpreted as a "*lack of evidence of validity*" (ISAAC, 2014, p8). Facilitator influence is not the same as "facilitator control". The possibility of some influence is widely accepted among FC supporters, however research (e.g. Cardinal, Hanson, Wakeham, 1996) shows that it may be significantly reduced through best practice and appropriate facilitator training. The more important question is: "*does the intervention (on balance) infringe or enhance the rights of the client?*" (Grayson, 1997, p232). *Please see Appendix 11 for further discussion.*	
Item 12 *At least a 20% sample of the data are extracted reliably by more than one rater (blinded to the treatments, if applicable)* (Schlosser, Wendt and Boesch, 2009). **Rationale:** • The synopsis of (Probst (2005) notes that the review: "*makes explicit the coding categories used for extracting data from each study*", however it notes that "[t]*he author seems to have made decisions regarding … the extraction of data from included studies*" (Schlosser and Wendt, 2008, p83). • The lack of an independent second rater is identified as a weakness, as "*the gathering of inter-rater agreement data would have allowed for a better assessment of the reliability of the data generated by this review* (Schlosser and Wendt, 2008, p83). • In the ISAAC Review, *a priori* coding categories were used only when coding of material against "criteria for inclusion", not for data extraction. • No coding categories were provided to guide data extraction from the three individual studies identified for inclusion in the ISAAC review. The Report does not indicate which members of the *ad hoc* Committee were involved in reviewing those papers, nor what process was followed for the review. • The ISAAC Report provides inadequate information regarding inter-rater agreement and dispute resolution. Although the Report states that: "*A democratic process was adopted throughout the review process* (ISAAC, 2014, p2), domination of the Committee by members who have made their anti-FC views public over many years makes it impossible to avoid the impression of a "stacked" committee. Under those circumstances, abandonment of attempts to reach consensus and reversion to voting is simply one more means of silencing dissenting voices.	No

Please see Appendix 12 for further discussion.	
Item 13 *Criteria used to arrive at judgments of quality are pre-defined and appropriate for the types of included designs* (Schlosser, Wendt and Boesch, 2009). **Rationale:** • As noted by the Chairman of the *ad hoc* Committee and colleagues in 2007: "*Because a systematic review can only be as sound as the included studies, an investigation of the quality of those studies is an important component of the data extraction and appraisal process"* (Schlosser, Wendt and Sigafoos, 2007, p145-6). • According to the synopsis of Probst (2005): "… *the criteria used to assess study quality are listed only for some of the subgroups of studies, and the reader is left to wonder whether the same criteria applied to other subgroups"* (Schlosser and Wendt, 2008, p83). • The only studies examined for quality were those concerning unexpected communicative performance. *"Although it was stated that data concerning quality of the studies would be extracted, it is unclear what criteria were employed* (Schlosser and Wendt, 2008, p81-2). • The *ad hoc* Committee's Report comments adversely on the quality of the three individual studies they include alongside Probst (2005). • Conclusions drawn from fundamentally flawed studies must be disregarded. However, it is important to recognise that disregarded conclusions cannot be taken as support for the opposite conclusion. It is therefore highly inappropriate for the *ad hoc* Committee's "*overall appraisal*" of these additional studies to state that they "*support the conclusions reached by Probst (2005) and subsequent reviews*" (ISAAC, 2014, 11). *Please see Appendix 13 for further discussion.*	No
Item 14 *Methods used to arrive at judgments of effectiveness for each study are pre-defined and operationalized* (Schlosser, Wendt and Boesch, 2009). **Rationale:** • Judgements of effectiveness are derived from outcome measures – the "dependent variables" in experimental studies. Experiments are designed to test hypotheses regarding the relationship between independent and dependent variables. Thus, operational definition of "*judgements of effectiveness*" requires clear, operational definition of all variables. • "Dependent variables" in the studies reviewed by the ad hoc Committee are measured as a percentage of "correct responses". (I have argued in Appendix 1, Section 1.3 that this is not a legitimate measure of "communicative competence" for this population.) • "Independent variables" are attributes or characteristics that affect the outcome (or dependent variable). Experiments focus on the treatment variable, which in FC research is "*often manipulated under 'facilitator non-blind', 'facilitator-blind', 'facilitator double-blind', or 'communication*	No

without FC' (non-FC) conditions (Schlosser and Wendt, 2008, p82). • "Secondary independent variables" are any influences in the selection of participants, procedures, statistics, or design likely to affect the outcome and provide an alternative explanation for the results (Creswell, 2002, p621). Researchers seek to neutralise these through statistical or design procedures. However, as discussed in Appendix 1, the characteristics of the population from which subjects may legitimately be selected, and the nature of the supports and accommodations encompassed by the term "Facilitated Communication", are not yet adequately defined. Statistical and design procedures cannot address that which is not understood. • Experimental studies have overwhelmingly been based on poor understanding of FC, resulting in poor practice. Facilitators were poorly trained, "best practice" was ignored, and inappropriate subjects were selected. Experiments featured intrusive "controls", highly likely to alter both the support needs of the subjects and the nature of support that could be provided. Such experiments are inadequate to eliminate the obvious rival explanations for their results: that they are artefacts resulting from the flaws in experimental design. • It is currently difficult to imagine how a phenomenon as complex as person-centred communication support in a dynamic social environment may be reduced to a controlled experiment. In the absence of good quality research, "best practice" provides the most detailed information available regarding the supports encompassed by FC; the characteristics of FC users; and the outcomes that may be expected from use of the strategy. • Professor Pat Mirenda has suggested a way forward for FC research: *One way to approach this task is to identify people with ASD who have become competent (and independent) communicators through the use of AAC (including FC), and then to work backwards to answer the question "Are there common factors that appear to have contributed to these good outcomes?" If we find any (and I believe we will), we can then design longitudinal hypothesis-driven studies to examine these factors in natural contexts* (Mirenda 2008, p229). *Please see Appendix 14 for further discussion.*	
Items 15 to 20: Meta-analysis *Systematic reviews may involve meta-analysis; however, meta-analysis requires specific additional appraisal considerations. … Meta-analyses use statistical methods to aggregate individual studies* (Schlosser, Wendt and Sigafoos, 2007, p139.) **Rationale:** It is clear that the *ad hoc* Committee's report is not, and does not have the capacity to be, a meta-analysis. Therefore, EVIDAAC Items 15 to 20 do not apply to this study.	No

Appendices

Appendix 1

Item One:

> *The review addresses a clearly focussed question? (e.g. was there enough information on the population studied, the intervention given, the outcomes considered?* (Schlosser, Wendt and Boesch, 2009, p11).

Rating: No.

In a 2009 presentation, the Chair of the ad hoc Committee (Dr. Schlosser), one of the ISAAC "outside reviewers" (Prof. Sigafoos), and a colleague noted that:

> *A focused question will define not only the population but also the intervention and targeted outcomes. If even one of these is missing, mark "No".* (Schlosser, Wendt and Boesch, 2009, p11).

This section will consider each of these factors in turn. It will also examine whether the *ad hoc* Committee has addressed the research questions it had been tasked with, and whether a systematic review may legitimately focus on broad research questions.

1.1 Population Studied

For the purposes of medical research, the "*population*" is described as "*the disease group or a spectrum of the well population"* (National Health and Medical Research Council, 2000a, p13). In Facilitated Communication (FC) research, the population is made up of individuals who may benefit from FC. However, the characteristics of this group have not been defined in a manner that allows clear delineation of the sample frame.

In a publication referring to reviews of Augmentative and Alternative Communication (AAC) research, the ISAAC *ad hoc* Committee Chair and colleagues noted:

> *… there could be a great degree of variability across included studies in terms of the methods used to* [identify the target population]. *This difficulty could affect the internal validity of the review, if the subjects were misdiagnosed in the original studies, or could affect the external validity because of difficultly determining the applicability of the evidence to specific clients* (Schlosser, Wendt and Sigafoos, 2007, p143)

The heterogeneity of AAC users raises doubt over the value of any group-based experimental design: an issue that has been recognized in peer reviewed literature. For example, Higginbotham and Bedrosian (1995) note that:

> *In the case of persons who are communicatively impaired, it is generally agreed that AAC users comprise a rather heterogeneous group of individuals, and the possession of some sort of communication disability and use of a communication technology may be their only commonalities. When AAC users are incorporated in research without stringent subject-selection criteria, investigators frequently are presented with a highly diverse set of performance-related factors (e.g., sensory, perceptual, and physical status; language and cognitive skills; educational level), making the results difficult to interpret and generalize … . Even when such studies produce statistically or clinically significant findings, it is difficult to relate these findings to other individuals, particularly if they do not*

> *share the same characteristics of the subjects studied* (Higginbotham and Bedrosian, 1995, p11).

Similarly:

> *When considering possible research designs, the use of group-level designs involving AAC subjects has often been criticized … due to the potential limitations in the homogeneity of the subjects participating. … Because of the relatively low incidence of AAC users in the general population, it is frequently difficult to find enough subjects required for a particular group-level research design who also can meet the subject-selection criteria needed to resolve the research question posed* (Higginbotham and Bedrosian, 1995, p12).

Please see comments on possible alternative research design in Appendix 14, Section 14.7.

The principal references used by the *ad hoc* Committee in their Report to ISAAC are a German language review of FC research (Probst, 2005) and the English language synopsis of that review (Schlosser and Wendt, 2008). According to the synopsis, Probst (2005) has extracted information from individual studies on the characteristics of "the sample", in terms of age, gender, and diagnosis.

Facilitated Communication "best practice" dictates that assessment of potential candidates for FC should be carried out by practitioners who have achieved a high level of training and experience, drawing on their clinical expertise. Identification of movement disturbance, as well as evaluation of how well an individual's existing communication strategies are meeting his or her needs, are central to such assessment.

In the absence of good quality research, such practitioners are the best available source of information on the characteristics of the population from which FC users may be drawn, and should be a starting point in any attempt to develop an operational definition of this population.

While some of the individual reports summarized by Probst (2005) make brief introductory comments regarding movement difference as a potentially significant factor in FC, no understanding of this is demonstrated in recruitment of subjects for the studies. The synopsis of Probst (2005) makes no mention of movement difference. I am unable to comment on whether it is mentioned in the review, as I am unable to read German.

Some of the studies reviewed by Probst (2005) sought to introduce FC to groups of subjects who had not previously used the strategy, in an effort to monitor any emergence of new skills (e.g. Bebko, Perry and Bryson, 1996; Bomba, O'Donnell, Markowitz, & Holmes, 1996; Eberlin, McConnachie, Ibel, & Volpe, 1993; Myles, Simpson & Smith, 1996; Smith & Belcher, 1993). These studies typically defined subjects on the basis of primary diagnoses, mental age, adaptive or "aberrant" behaviours, unfacilitated expressive language abilities, etc.

Other studies reviewed by Probst (e.g. Cabay, 1994; Crews, Sanders, Hensley, Johnson, Bonaventura, Rhodes & Garren, 1995; Beck and Pirovano, 1996) and those reviewed by the *ad hoc* Committee (Olney, 2001; Schiavo, Tressoldi, & Martinez, 2005; and Perini, Rollo and Gazzotti, 2010) chose subjects who had already been using FC and who *may*, therefore, have undergone appropriate assessment. However there are few if any records in the literature of such assessment occurring, and no details of movement differences that may thus have been identified.

There is no guarantee that the subjects in these studies were appropriately assessed, nor that their support followed "best practice" guidelines. Studies reviewed by Probst date from 1993 to 1998. One of the problems faced by experimenters in the 1990s was the astoundingly haphazard spread of FC through North America, resulting in confusion over the nature of the strategy. In short, while research from the 1990s – the bulk of "controlled studies" to date - may have reflected the state of knowledge and regional differences at the time, it is by no means comparable with – and therefore cannot be taken as commentary on – current "best practice" in facilitated communication.

Please note that Appendix B, Section B3 of my 2012 submission to the ISAAC review (Davies, 2012) examined issues around selection of subjects. The *ad hoc* Committee report provides no acknowledgement of these issues.

1.2 Intervention Given

> *Systematic reviews of studies regarding treatment efficacy and effectiveness should define the interventions they target. … This allows the consumer to determine the homogeneity of the studies and the efficacy of specific treatments (Schlosser, Wendt and Sigafoos, 2007, p144).*

For the purposes of medical research, the "*study factor*" is "*the intervention, diagnostic test, or exposure*" (National Health and Medical Research Council, 2000a, p.13). In FC research, this refers to the complex, person-centred strategies and supports that make up Facilitated Communication, and the way they are flexibly applied by skilled facilitators. Once again, "best practice" must be the starting point in any attempt to develop an operational definition of "the intervention", and qualitative research will be needed to achieve this.

The synopsis of Probst (2005) states that:

> *… the author provided an excellent context for this review in terms of describing the method and its theoretical background, along with socio-cultural aspects surrounding FC* (Schlosser and Wendt, 2008, p83).

I am unable to comment on the context provided by Probst, as I am unable to read German. The synopsis gives no details, nor does it seek to compare Probst's account with "best practice". No categories of data are listed that might assist in confirming that the interventions applied in each study complied with either.

In fact, this omission is unsurprising. The studies reviewed by Probst (2005) date from 1993 to 1998. As noted in Section 1.1, one of the problems faced by experimenters in the 1990s was the rapid and unrestrained spread of FC through North America:

> *…the procedure has the likely potential of being used in an inconsistent and potentially wanton fashion until it is better understood. In some instances facilitators have almost no training in facilitated communication; in other cases individuals are trained to use facilitated communication with only minimal regard for specified methodology; and users of facilitated communication routinely fail to follow consistent routines and guidelines. That most colleges and universities have been reluctant to train individuals to use facilitated communication until the procedure is better understood and shown to be efficacious only intensifies this problem* (Simpson and Myles, 1995).

The nature of the intervention being provided in each study reviewed by Probst – and indeed, each of the three individual studies reviewed by the ISAAC *ad hoc* Committee – is likely to have been impacted by the experience, training, and attitudes of facilitators and researchers. It is also likely to have been impacted by the settings in which studies were conducted (naturalistic or contrived), and – very significantly – the restrictions "control procedures" place on facilitator's flexible provision of support strategies.

In short, as noted in Appendix B of my initial submission to ISAAC (Davies, 2012), the procedures implemented and subjected to investigation in the vast majority of quantitative studies of FC to date bear only superficial resemblance to FCT best practice, which is yet to be operationally defined to support quantitative research.

1.3 Outcomes Considered

1.3.1 Outcome measures

The Australian National Health and Medical Research Council (NHMRC), in respect of medical research, notes that:

> *Often, the types of outcomes measured and reported in clinical trials are those that are easy to measure and that can be expected to show changes or differences in a relatively short period of time. These may, or may not be, important to the recipients of the intervention. … The outcomes being measured may not reflect how a patient feels, or whether they can go about their usual activities in a satisfactory way. Factors that relate to an improved quality of life … may be more patient relevant … . These types of measures are indicators of the balance between the treatment's benefits and harms* (NHMRC, 2000b, p24).

In medical research, carefully defined treatments can be randomly allocated to large groups of subjects drawn randomly from a well-defined study population. Yet, as the previous quote demonstrates, even in such ideal circumstances it is recognized that the outcome of clinical trials is not the whole story. Models for "evidence based" decision-making make it very clear that factors other than research outcomes – practitioner experience and expertise together with client needs and preferences – are essential to the decision-making process.

AAC research cannot be undertaken in the same rigorous manner as medical research. It is therefore even more important to recognize that "*the intended emphasis of EBP* [Evidence Based Practice] *rests on the shared integration of the three cornerstones*" (Schlosser, 2004) – that is, on research evidence, clinical expertise, and client needs and preferences. In the words of one commentator:

> *Elevating research findings to a position of predominance or supremacy over these other factors, or to the exclusion of these factors – though commonly the case – constitutes a misuse of EBP* (Prizant, 2011, p46).

Despite this, only one measure of outcome is ultimately considered by the ISAAC *ad hoc* Committee: "communicative competence", measured as a percentage of "correct responses" in:

> *an* a priori *controlled manipulation of what knowledge/stimuli was presented to the facilitator and the individual using FC in an attempt to empirically establish who was authoring the messages produced in response to the stimuli* (ISAAC, 2014, p3).

1.3.2 Validity of outcome measure

I am not aware of any evidence, nor any attempt to demonstrate, that confrontational testing or message passing are valid tests for the study population, nor of any individual's capacity for communication more broadly defined. The designs of most of the studies deemed "Level 1" by the *ad hoc* Committee are incapable of demonstrating anything other than the widely accepted facts that FC users have difficulty with confrontational testing or message passing in contrived circumstances, and that facilitators can influence communications under those conditions.

Speech Pathology Australia (SPA), in their Clinical Guideline for AAC, has clearly recognised the need to broaden the outcomes that are measured to include practice based evidence from real-world contexts and impacts on quality of life:

> *Outcome measurement in the field of AAC has, to date, focused primarily upon 'operational competence' (i.e., skills in access and use of an AAC system or strategy) and 'linguistic competence' (i.e., skills in receiving language through the spoken modality and producing language using the AAC system). … Successful AAC interventions resulting in better communication can lead to improved personal well-being and quality of life and reduction in behaviours of concern (Sigafoos et al., 2003). Therefore, it is recommended that speech pathology practice includes measurement of outcomes. There is also a need to broaden the scope of measures, in particular, to address the extent to which AAC interventions lead to improved participation in all aspects of daily life.* (Speech Pathology Australia, 2012, p28).

The Clinical Guideline for AAC also makes the following observations:

> *The optimal processes for assessment, intervention, and monitoring/evaluating outcomes when providing AAC to people with complex communication needs can be based on principles drawn from the Participation Model (Beukelman &*

> *Mirenda, 2005) and Dynamic Assessment (Hasson & Joffe, 2007; Iacono & Caithness, 2009). The Participation Model includes principles for identifying how various AAC systems and options can address current and future needs of people with complex communication needs, with a focus on facilitating their participation within their chosen communities. Dynamic assessment targets the person's potential for learning rather than the conduct of a static assessment of current skills, and aligns with the principles and philosophy of the Participation Model (Beukelman & Mirenda, 2005)* (Speech Pathology Australia, 2012, p.17).

It is now widely recognised that:

> *… Eligibility determinations based on* a priori *criteria violate recommended practice principles by precluding consideration of individual needs* (National Joint Committee for the Communication Needs of Persons With Severe Disabilities, 2003).

The limitations related to standardised testing for autistic individuals in particular have been recognised by SPA in their position statement: "Evidence Based Speech Pathology Practice for Individuals with Autism Spectrum Disorder":

> *Formal assessments of language, including the use of standardised tests … may be helpful in gauging a child's development with reference results to age-based norms. However formal assessments of this nature tend to rely on child cooperation and may be of limited use with students with significant intellectual disabilities, as well as those with no functional speech (National Research Council, 2001; Wetherby, Prizant & Shuler, 2003). Formal assessment may also be of limited use for some individuals with autism because of the largely social-pragmatic nature of the disorder, which is often beyond the scope of these assessments.*
>
> *As a result, a range of approaches may be appropriate in assessing individuals with autism (Wetherby, Prizant and Schuler, 2003), including formal and informal assessments, observational, and structured and unstructured assessment tasks. In addition, dynamic assessment, in which the speech pathologist assesses an individual's potential for learning rather than his or her static level of performance at a particular point in time (Hasson & Joffe, 2007), may also be appropriate for individuals with autism. Regardless of the tools used, the assessment must account for the individual's skills and difficulties, and fluctuations and differences in these skills and difficulties across a range of contexts and with a variety of communication partners. In addition, the assessment must consider the activity limitations, participation restrictions, and the personal and environmental factors which impact on the individuals functioning (Australian Institute of Health and Welfare, 2003, Filipek, 1999; National Research Council, 2001)* (Speech Pathology Australia, 2009, p9).

Quite clearly, the approaches described by SPA (2009 and 2012) are very different from those described in the studies that have been included by the *ad hoc* Committee. However, even within included (Level 1) studies, it is important to note the variety of methods used to assess “communicative competence”. The validity of individual instruments may also be questioned, as noted in Section 1.3.4.

1.3.3 Validity of instruments

Instruments are tools for measuring, observing, or documenting quantitative data (Creswell, 2002, p622). These tools must be examined for validity, which means:

> *… the development of sound evidence to demonstrate that the test interpretation (of scores about the concept or construct that the test is assumed to measure) matches its proposed use.”* (Creswell, 2002, p159).

Approaches to providing such evidence include examination of test content, response process, and consequence of testing.

For **evidence based on test content**, the question asked is whether:

> *… the scores from the instrument show that the test’s content relates to what the test is intended to measure* (Creswell, 2002, p162).

As noted in Section 1.3.2, I am unaware of any evidence, nor any attempt to demonstrate, that the ability to produce a set answer in confrontational testing or to pass messages under controlled conditions is a valid indicator of any individual’s capacity for other forms of communication, for example self-motivated communication on topics of their own choice in supportive, naturalistic circumstances.

Nor am I aware of any evidence that such a test is meaningful or appropriate for use in the study population.

Against this, qualitative observational evidence together with anecdotal reports of successful communication using FC provide more than adequate reason to question the validity of confrontational testing, message passing, and highly contrived settings for assessment of communicative competence in the FC population.

Evidence based on response process requires that instruments are:

> *… evaluated for the fit between the construct being measured and nature of the responses of the individuals completing the instrument … . Validity evidence can be assembled through interviews of the participants to report what they experienced or were thinking when they completed the instrument* (Creswell, 2002, p162).

Clearly, when authorship is ‘the question’, interviews with participants may be problematic. However this does not mean that FC user’s reports of difficulties with confrontational testing, message passing, and contrived test conditions can simply be ignored. Clearly alternative experimental or observational approaches must be taken.

This makes the ISAAC *ad hoc* Committee’s decision to ignore alternative (Level 2) study designs, for example Grayson, Emerson, Howard-Jones and O’Neil (2011), Bernardi & Tuzzi (2011a), Bernardi and Tuzzi (2011b), and Tuzzi (2009), all the more inappropriate.

Evidence based on the consequences of testing asks: "*What benefits (or liabilities) have resulted from using the instrument?* (Creswell, 2002, p164).

Confrontational testing and message passing are not only used in research: they have also been used in clinical assessment to discredit Facilitated Communication Training (FCT). In some situations, individuals who have been using FCT over an extended period have lost support. Follow-up studies on the impact this has had on their lives may be informative.

Prizant, Wetherby, & Rydell (2000) make the following comment regarding the nature of intervention most likely to produce real benefit in the cognitive development of autistic individuals:

> *By creating contexts for joint action and joint attention … and by coaching peers and adults in how to sustain interactions, a greater sense of communicative efficacy is established. … Such an emphasis will reduce the transactional secondary effects of more primary disabilities, which may be more devastating in the long run than the initial limitations exhibited by the child. Ultimately, the individual's competence in social interaction, in developing relationships, and in the capacity to cope with stress using flexible communicative strategies will determine the level of independence that he or she can have beyond early childhood* (Prizant, Wetherby, & Rydell, 2000, p134-5).

Problems will inevitably arise when testing, rather than teaching, becomes the objective of an activity. The practice of scaffolded learning is very different from a testing process intended to identify the skills an individual can demonstrate without scaffolding. However, the fact that this is a challenge should not be used as an excuse for excluding individuals with severe communication impairment from enriched learning and training opportunities. **Such issues are a challenge that must be addressed by educators and researchers. They should not be made the learners' problem.**

1.3.4 Outcome assessment in the reviewed studies

As discussed in my original submission to ISAAC (Davies, 2012), many studies were based on activities that were unlikely to motivate subjects or to alleviate communication frustration (picture matching etc.); that relied on literacy (while FC requires an assumption of competence, it does not rely on literacy); that were time-limited (highly inappropriate for a population that may be characterised by motor planning difficulties); and that were based on questions that may have been inappropriate given the background of the subjects involved (for example, questions about preferences when subjects may never have experienced choice-making, or questions related to emotional reactions or requiring speculation about abstract concepts delivered alongside other questions that clearly had "right" or "wrong" answers).

To provide one specific example in which use of a specific instrument appears questionable, Beck and Pirovano (1996) selected twelve subjects who had been reported to have used FC with some success for at least a year. When the Peabody Picture

Vocabulary Test was administered to them using FC under controlled conditions, however:

> *No subject achieved a score high enough to be converted to a standard score or a percentile. Some subjects also did not score high enough to be able to convert their scores to age equivalencies* (p.504).

This did not lead the authors to question whether the instrument may not be appropriate for these subjects – instead they persisted and used raw scores to compare subject's response levels under various conditions.

The twelve subjects were tested with, at different times, visual and auditory stimuli. The authors described this as "*24 opportunities to show increased ability with FC*", and noted that "*improvement was noted 16 times*" (p.509). Most of these apparently positive findings, however, can legitimately be disregarded as only three met the conditions required to be statistically significant.

The statistical weakness of this study illustrates the difficulties that must be overcome if controlled, quantitative studies are to provide legitimate data in the behavioural sciences. An increase in the number of trials may have increased statistical reliability of the results; however such an increase is likely to be problematic given the nature of the task. Excessive repetition of simplistic tasks is likely to produce confounding behaviours from subject because the task is essentially meaningless or uninteresting, or because subjects take the repetition to indicate that there was something wrong with their earlier responses and react accordingly. Alternatively, subjects may learn from the repetition so that conditions change as testing proceeds.

Instead of acknowledging that no conclusion could legitimately be drawn from results that are not statistically significant, the authors state only that the apparently positive results cannot be taken as support for the validity of FC. This masks the obvious fact that the results cannot be taken as evidence against FC either.

1.4 The research question

Schlosser, Wendt and Sigafoos (2007) assert that:

> *A sound question is one of the outcomes of a well-conducted problem-formulation process, and the first step in planning a systematic review* (Schlosser, Wendt and Sigafoos, 2007, p140).

None-the-less, there has been no consistent advice regarding the question to be addressed by the review.

The press release which first announced the review, dated May 11th 2012, provided highly restrictive guidelines regarding material to be considered (ISAAC 2012). These guidelines were relaxed by a directive from the ISAAC Council to the ISAAC Executive Board, ratified on July 29, 2012, in response to member feedback:

> *The first modification is that the scope of the review will include both quantitative and qualitative studies. The second modification is that both published and*

> *unpublished sources will be included in the search and appraisal process in developing the draft* (ISAAC, 2013a).

Inexplicably, these modifications were not communicated to the membership until February 8, 2013, at which time I contacted ISAAC to request clarification. In part, my request read:

> *Looking back at the original guidelines for materials that may be submitted, the press release indicates that the first part of Section 3a has been modified. Will the other restrictions remain in place? That is, … Section 3b, which stated that material submitted:*
>
> > … must evaluate the authorship of the messages; descriptive analyses of output generated by someone using Facilitated Communication will not be reviewed unless there has been prior independent verification that the authorship has been attributed to the individual using Facilitated Communication [ISAAC, 2013a].
>
> *… Many published qualitative studies address issues other than authorship. Does ISAAC's decision to consider qualitative and unpublished materials reflect acceptance of wider research goals? If so, it is important to note that wider goals will not be served if the other restrictions on materials remain in place.*
>
> *If wider goals have been embraced, will they be articulated for the benefit of stakeholders?*
>
> *If wider goals have not been embraced, the impact this has on limiting material to be considered must be made quite explicit to stakeholders.* (Cathie Davies, email to ISAAC via Franklin Smith, 9 February 2013.)

Having received no clarification, I emailed ISAAC International (via Franklin Smith) as follows on 29 April 2013, regarding the nature of research questions to be addressed by the *ad hoc* Committee:

> *To date the FC debate has highlighted the difference between qualitative and quantitative research as if each has emerged from different and entirely irreconcilable philosophies. However it should be very clear that there is important interplay between the two. Qualitative research asks open questions to find out more about a phenomenon - information absolutely necessary before quantitative research can (or should) proceed. To start with one, fixed question (i.e. authorship) before the variables that may impact it have been properly defined (i.e. what, exactly, is FC?) is so wrong it should be laughable – yet this type of error typifies the experimental research conducted during the 1990s.*
>
> *The original review guidelines were framed to exclude any material that did not specifically address authorship. The current guidelines are not clear on whether this focus has changed, despite my request on 9 February for clear articulation of research goals. As a result, a somewhat scattergun approach to collecting material for the current phase of the review has been necessary. It has not been*

> *possible to assist the committee by arranging materials around particular research questions* (Cathie Davies, email to ISAAC via Franklin Smith, 29 April 2013.)

Even basic information regarding submission of new material had not been made available until 26 March 2013, with the deadline for submissions 30 April - although the material was to be dated no later than 15 April. As I noted in my email dated 29 April 2013, the twenty day period between publication of the new guidelines on 26 March and the deadline excluded anything written after 15 April was inadequate for collation of unpublished practice-based data into a form suitable for submission.

As it turns out, this is a moot point as such unpublished submissions would have been ignored by the *ad hoc* Committee, along with everything other than material they deemed "Level One" – please see my comments in respect of EVIDAAC Item 8 and Appendix 8 for further details. As a result, the Committee has produced a report entirely formed by their unexamined assertion that "*the validity of FC stands and falls with the issue of authorship*" (ISAAC, 2014, p6), with no effort made to address the broader, more open-ended research questions implied by the "*adjustments to the framework within which the FC Committee is conducting its efforts*" (ISAAC, 2013a).

1.5 Can a systematic review focus on broad research questions?

The quality of a systematic review depends on the quality of the studies appraised (Wright, Brand, Dunn, and Spindler, 2007, p24). We have seen that, in FC research, adequate definitions of the population, the intervention, and meaningful outcomes are absent. As a result, the evidence on which the *ad hoc* Committee has reported is not of adequate quality to support their conclusions. It is most certainly not of adequate quality to inform a Position Statement. The only position ISAAC can reasonably take is that existing research evidence is fatally flawed, and further research is required.

A broader approach is not incompatible with a systematic review of evidence:

> *Reviews may focus on treatments and their effectiveness, diagnosis, prognosis, epidemiology, perspectives based on qualitative studies, theories, and other areas such as methodological rigor of studies and reviews or measurement issues* (Schlosser, Wendt and Sigafoos, 2007, p139).

However, it is recognised that a broader approach will not lead to a definitive position on FCT:

> … [i]*f a research question necessitates the inclusion of low level evidence … then the systematic review is likewise low-level evidence. Such reviews can be important preliminary studies, and may identify incidence of results and areas for future research …* (Wright et al, 2007, p24).

A statement recognising the need for more research is surely preferable to a Position Statement that has been based on seriously inadequate research.

Appendix 2

Item Two:

The search methods were pre-defined (Schlosser, Wendt and Boesch, 2009, p12).

Rating: No

2.1 What is meant by "pre-defined search methods"?

The literature seems equivocal on the exact meaning of "*pre-defined search methods*". One source suggests that:

> *The methods for literature searching, screening, data extraction, and analysis should be contained in a written document to minimize bias before starting the literature search* (Wright, Brand, Dunn, and Spindler, 2007, p24).

On the other hand, the *ad hoc* Committee Chair and colleagues appear, in a 2009 presentation, to place the focus on reporting, not process:

> *"Pre-defined" means "*a priori*" or that the search methods are explained upfront before the authors report on included studies and results (rather than as an afterthought when the authors mention in the discussion section that they searched a particular database which had not been mentioned in the Methods section)* (Schlosser, Wendt and Boesch, 2009, p12).

Despite this, in 2007 two of the same authors (writing with a colleague who is one of the "outside reviewers" for ISAAC's review of FC) noted that "*… the presence of a protocol is essential for the rigorous implementation of a review*" (Schlosser, Wendt and Sigafoos, 2007, p139). They elaborate:

> *A protocol is developed* a priori *and serves as a 'road map' by outlining the essential procedures for conducting the review* (Schlosser, Wendt and Sigafoos, 2007, p139).

This section will consider the question of pre-determined search methods using three important issued identified by Schlosser, Wendt and Sigafoos (2007):

> *… the presence of a protocol; the peer-review and approval of the protocol by an organization; and the degree of adherence to the protocol* (Schlosser, Wendt and Sigafoos, 2007, p139).

It may be noted that these questions do imply that the focus must be on process rather than simply on reporting.

2.2 The presence of a protocol

The ISAAC *ad hoc* Committee's report does not refer to any written protocol document produced before starting the literature search in 2012. If such a document had existed, it seems likely that it would have required extensive review after the directive from the ISAAC Council to the ISAAC Executive Board, ratified on July 29, 2012, which broadened the range of evidence to include both quantitative and qualitative studies, and both published and unpublished sources (ISAAC, 2013a). As no such review is mentioned, it can perhaps be assumed that no protocol document existed before the report was written.

2.3 Approval of the protocol by an organization

> *If a protocol was used, the reader should examine whether it was approved by an organization such as the Cochrane Collaboration or the Campbell Collaboration* (Schlosser, Wendt and Sigafoos, 2007, p139).

The *ad hoc* Committee's Report makes no comment on this issue. Neither does it make any reference to negotiations between the *ad hoc* Committee and ISAAC International to develop and gain approval for a search protocol.

2.4 Adherence to the protocol

> *… readers should evaluate whether (and how closely) the authors adhered to the protocol. Protocols reduce the likelihood of a biased selection of studies, because the rules for eligibility are set up* a priori … (Schlosser, Wendt and Sigafoos, 2007, p139).

The *ad hoc* Committee outlines the method they used to "Search for Studies and Reviews" on pages 2 and 3 of their report. A multi-faceted search strategy is described, involving five databases, ancestry searches, contact with individual authors, and materials submitted by the ISAAC membership.

The databases chosen appear to be appropriate, but as discussed in my response to EVIDAAC Items 3,4 and 5, sources relied upon were inappropriately limited, resulting in source selection bias, publication bias and database bias.

Terms used to search the databases were "facilitated communication", "supported typing", and "assisted typing". As discussed in relation to EVIDAAC Item 6, these terms seem excessively broad, likely to produce an unwieldy volume of material requiring subsequent filtering. Clear statement of research questions, coupled with search terms tailored to target each question, may have produced more focused results.

Table One summarises the number of items derived from an attempt to replicate the search on 5 May 2014. It is understood that many papers may be listed on more than one database, and that some may not, on closer investigation, relate to the AAC strategy of Facilitated Communication: however Table One suggests that the search method described on pages 2 to 3 of the *ad hoc* Committee's Report could be expected to yield a considerably larger list of items from all sources (databases, ancestry searches, contact with individual authors, and materials submitted by the ISAAC membership) that that produced on pages 4 to 6 of the Report.

It therefore appears unlikely that the search method described in the "Search for Studies and Reviews" section has been adhered to.

Using Level 1 material as an example, the items listed on page 4 of the Report include only ten reviews and three individual studies that may have been identified in the manner described under "Search for Studies and Reviews". All remaining studies appear to have been identified "*per analysis of Probst (2005)*". This apparently late adjustment to the search method seems to fit with the description of "*search methods … explained … as an afterthought*" (Schlosser, Wendt and Boesch, 2009, p12).

Further, five of the nine items listed as peer reviewed "Level 4" material on page 5 of the Report are drawn from one edition of *Evidence-Based Communication Assessment and Intervention.* This seems an unlikely balance compared with results that may be expected from the search methods described. I note that Editors-in-Chief of *Evidence-Based Communication Assessment and Intervention* include the Chair of the *ad hoc* Committee (Dr. Schlosser) and an ISAAC "outside reviewer" (Prof. Sigafoos).

The response to EVIDAAC Item Two, "*the search methods were pre-defined*", must therefore be "no".

Table One: Attempt to replicate database search described by ad hoc Committee

Database	search term	Number of items identified
CINAHL	facilitated communication	72
	supported typing	0
	assisted typing	0
ERIC	facilitated communication	167
	supported typing	0
	assisted typing	0
PubMed	facilitated communication	140
	supported typing	1
	assisted typing	7
LLBA	facilitated communication	137
	supported typing	2
	assisted typing	0
psychInfo (via EBSCO)	facilitated communication	264
	supported typing	3
	assisted typing	1

Appendix 3

Item Three:

> *Multiple complementary sources that collectively minimize source selection bias are consulted. (Consider databases, hand searches, ancestory searches, contacting authors, forward citation sources.)* (Schlosser, Wendt and Boesch, 2009, p12).

Rating: No.

The *ad hoc* Committee Chairman and colleagues noted in 2007 that:

> *… a multi-faceted search strategy needs to be implemented in which search strategies complement each another. This increases the likelihood that the search is comprehensive and minimizes the oversight of studies outside the researcher's regular purview (White, 1994)* (Schlosser, Wendt and Sigafoos, 2007, p141).

As discussed in Appendix 2, Section 2.4, there is reason to doubt adherence to the "pre-defined" search methodology described on pages 2 and 3 of the Report. The method described appears "multi-faceted" and gives the impression that a wide range of studies and other documents have been included for consideration by the *ad hoc* Committee. However, a mismatch is apparent between the described search methods and the surprisingly limited list of items identified on pages 4 and 5.

It would seem, either that the search methods described have not been followed, or that some items identified through the search process have been excluded without mention prior to application of the identified inclusion/exclusion criteria. (For discussion of material that has been excluded according to identifiable criteria, please see Table Two and my comments in respect of EVIDAAC Items 7, 8, 9 and 10.)

There is also reason to suspect publication bias and database bias, as discussed in relation to EVIDAAC Items 4 and 5.

Certainly many items submitted by members have not been mentioned at all in the Report. I personally submitted 111 items in response to the 2013 "*Information on new submissions to be considered*" (ISAAC, 2013b), either electronically (by reference or PDF) or in hard copy. Of these:

- 5 were listed as excluded (I assume because they were books and therefore not peer reviewed);
- 2 were rated as Level 3 material and therefore not further considered by the committee;
- 8 were rated as Level 2 material and therefore not further considered by the committee;
- 3 were rated as Level 1 material, but two were considered only by mentioning that they had been included in one of the narrative reviews cited by the committee;
- 93 were not cited in any way, with no indication that they had ever been considered by the committee.

I am aware that material submitted by other members was also ignored. This does not support the suggestion that the reviewers have utilised "*multiple complementary sources*" (Schlosser, Wendt and Boesch, 2009, p12).

To summarise, although the method described under "Search for Studies and Reviews" (page 2-3 of the *ad hoc* Committee's Report) appears "multi-faceted", it is clear that a very narrow approach has in fact been taken. Following an unconventional hierarchy of evidence (the 5-S model) adapted for AAC clinicians by the Committee Chair (Dr. Schlosser) and one of the "outside reviewers" (Professor Sigafoos), the principal source for examination of Level 1 material (the only material ultimately taken into account in the Report) has been a synopsis co-authored by Dr. Schlosser and published in the peer reviewed journal *Evidence-Based Communication Assessment and Intervention*, which is edited by Dr. Schlosser and Prof. Sigafoos. The synopsis refers to a German language review written by a member of the *ad hoc* Committee, Probst (2005). Only three other individual studies have been considered by the *ad hoc* Committee.

Neither the synopsis (Schlosser and Wendt, 2008) nor the review (Probst, 2005) would be admitted as evidence under the American Speech-Language-Hearing Association (ASHA) hierarchy of evidence, which places "*meta-analysis of randomized controlled trials*" as the highest level of evidence but makes no reference to systematic reviews of non-randomised trials (ASHA, undated). There are no Randomly Controlled Trials addressing FC, and Probst is not a meta-analysis. It is more correctly seen as a narrative review, as "*the conclusions are based more on descriptive data than on inferential analyses"* (Schlosser and Wendt, 2008, p83).

Further, although the synopsis states that Probst (2005) used "*multiple, complementary sources and databases … in order to minimize the danger of source-selection bias in general, and database bias in particular* (Schlosser and Wendt, 2008, p83), the central importance of Probst in the ISAAC review must be questioned. It does not address the open-ended questions implied by the directive from the ISAAC Council to the ISAAC Executive Board, which was ratified on July 29, 2012 and communicated to members through ISAAC International's "*Adjustments to the Framework Within Which the FC Committee is Conducting its Efforts*" (ISAAC, 2013a) and "*Information on new submissions to be considered*" (ISAAC, 2013b). (Please see Appendix 1, Section 1.4, for elaboration.)

In short, the approach taken by the *ad hoc* Committee appears to have done nothing to "*minimizes the oversight of studies outside the researcher's regular purview*" (Schlosser, Wendt and Sigafoos, 2007, p141).

Appendix 4

Item Four

> *An attempt was made to locate unpublished studies (e.g. trial registers, dissertations etc.)* (Schlosser, Wendt and Boesch, 2009).

Rating: No.

Unpublished studies have been explicitely ignored by the *ad hoc* Committee:

> *Given that our committee includes researchers/scholars from both qualitative and quantitative traditions in which peer-reviewed articles enjoy the highest regard, we restricted the inclusion of articles, across all four levels of analysis, to those that appeared in the peer-reviewed journal literature* (ISAAC, 2014, p3).

This approach is extremely inappropriate in such a polarised field. In the words of one commentator: "*… FC is one of the best exemplars of how practice can become absolutely dissociated from empirical research*" (Mostert, 2010, p39). Given this history and the gate-keeping role of researchers and academics in peer-reviewed literature, the restriction to peer reviewed material - almost by definition - admits only one side of the story for consideration.

The following quote from an earlier publication by the Chair of the *ad hoc* Committee and colleagues suggests that this outcome is well understood. The exclusion cannot, therefore, be seen as anything other than deliberate bias:

> *… publication bias … is ''what occurs whenever the research that appears in the published literature is systematically unrepresentative of the population of completed studies'' (Rothstein et al., 2005, p 1) … The prevalence of publication bias in the speech–language pathology literature is currently unknown, but probably does exist. Hence, when appraising systematic reviews, readers should determine whether the authors of a review attempted to locate unpublished studies such as conference proceedings, unpublished theses and dissertations, and other 'grey' literature. A bias through the exclusion of unpublished literature may threaten the validity of the systematic review* (Rothstein, Sutton, & Borenstein, 2005, cited in Schlosser, Wendt and Sigafoos, 2007, p141).

With regard to the principal reference consulted by the *ad hoc* Committee, it is once again important to stress that the research questions addressed by Probst (2005) do not coincide with those implied by the directive from the ISAAC Council to the ISAAC Executive Board, ratified on July 29, 2012 and communicated to members through ISAAC International's "*Adjustments to the Framework Within Which the FC Committee is Conducting its Efforts*" (ISAAC, 2013a) and "*Information on new submissions to be considered*" (ISAAC, 2013b). (Please see Appendix 1, Section 1.4, for elaboration.) The comment in the synopsis, that Probst: "*considered other sources as well, including bibliographies published on the internet, books, and professional journals"* (Schlosser and Wendt, 2008, p83), does nothing to improve the study's relevance to the broader questions at hand.

Turning now to one form of unpublished data - practice-based evidence: in my email to ISAAC International (via Franklin Smith) dated 9 February 2013, I advised:

> *As much of the material that may be of value to the committee is unpublished (please see the section of my submission* [Davies, 2012] *dealing with "practice based evidence"), please note that a time-frame longer than the stated sixty days may be needed to collate it into a form suitable for submission* (Cathie Davies, email to ISAAC via Franklin Smith, 29 April 2013.)

I received no response on this matter, and subsequently made the following comments in my email dated 29 April 2013, which accompanied a summary of the material I had submitted over the intervening weeks:

> ***Practice Based Evidence***
> *In my original* [2012] *submission, I wrote about "practice based evidence" - the kind of systematic data that has been collected on our kids at clinics and schools over many years. There is a vast wealth of this available, just waiting for analysis. Add to this the rich personal accounts and we have some wonderful raw data. However – as noted in my email of 9 February, sixty days would be inadequate to collate this kind of material into a form suitable for submission. The nineteen days between publication of the new guidelines on 26 March and the deadline excluding anything written after 15 April made any such effort utterly impossible.*
>
> *The reality is that most FC stakeholders are AAC users, parents, carers or clinicians. We don't have the time, resources - or even the current research skills - to do the analysis ourselves. For decades very little funding has been available for FC research - largely because influential individuals such as the Chair of the ISAAC review committee have told the world that the questions have all been answered. By calling for new submissions, ISAAC appears to have acknowledged that the experimental data may not tell the full story. The appropriate course would be to announce this to the world, so quality research can at last proceed in the appropriate manner* (Cathie Davies, email to ISAAC via Franklin Smith, 29 April 2013.)

For further information regarding the importance to Evidence Based Practice (EBP) of evidence derived from clinical practice and practitioner expertise. please see Section 3 of my original submission to the ISAAC review (Davies 2012).

Appendix 5

Item Five:

> *Databases are carefully selected so that they, together, minimize the potential of systematically excluding studies* (Schlosser, Wendt and Boesch, 2009, p13).

Rating: No.

5.1 Why are the chosen databases inadequate?

As noted by the Chair of the *ad hoc* Committee and colleagues in 2007:

> *… when appraising a review, a reader should look for a careful selection of multiple, appropriate databases, so that the yield of relevant studies is maximized and the risk of database bias minimized* (Schlosser, Wendt and Sigafoos, 2007, p141).

The "gatekeeping" role of researchers and academics in peer reviewed journals is particularly important in the context of FC research, due to the long disconnect between FC research and practice. The effect is likely to be amplified because Facilitated Communication draws little attention from researchers outside AAC. As a result there may be less critical academic scrutiny, so that poor quality research goes unchallenged. In direct contradiction of accepted "hierarchies of evidence", this results in greater reliance on "expert opinion". (Please see Appendix 8, Section 8.8 and Appendix 11, Section 11.4 for further comment regarding reliance on "expert opinion".)

The databases listed on page 2 of the Report appear most likely to yield material similar to that authored by members of the *ad hoc* Committee. The broader questions implied by the July 2012 directive from the ISAAC Council to the ISAAC Executive Board (communicated to members as "*Adjustments to the Framework Within Which the FC Committee is Conducting its Efforts*", ISAAC, 2013a) require a broader range of databases to be searched if an unbiased review of evidence is genuinely being attempted.

A broader search is necessary because, as noted in respect of EVIDAAC Item 1, adequate definitions of the population, the intervention, and the most meaningful outcomes are absent from most quantitative studies of FC. Fundamentally flawed experiments must be disregarded - even if their designers were aiming high in the "hierarchy of evidence". In the absence of valid quantitative research, it is necessary to consider studies that may rank lower in traditional hierarchies of evidence. Studies using methods considered less rigorous are, however, less likely to be found in the databases chosen by the *ad hoc* Committee.

5.2 How might additional databases be identified?

The *ad hoc* Committee Report notes that one of the three individual studies they have considered in their review was not included in previous reviews.

> *Likely this omission is due to it being published in* Disability Studies Quarterly. *This is a journal in the field of disability studies (more closely linked to literary analysis in the humanities) that is indexed by the Modern Languages Association*

> *database rather than databases commonly searched in the field of communication sciences and disorders* (ISAAC, 2014, pp7-8).

In 2013, ISAAC International provided an expanded list of materials to be considered:

- *published peer reviewed reports on FC*
- *unpublished reports on FC*
- *written anecdotal reports on FC*
- *qualitative reports on FC*
- *any other written evidence on FC* (ISAAC, 2013b)

Many of these items are unlikely to be found in the databases "*commonly searched in the field of communication sciences and disorders*" (ISAAC, 2014, p8). As noted in Section 5.1, the broader range of included materials demands a broader search methodology.

One means by which alternative data bases may have been identified is hinted at in the description of search methods on page 2 of the *ad hoc* Committee's Report. Here, provision is made for ancestry searches:

> … *the searching of bibliographies of obtained studies, reviews, previous position statements, and websites for additional studies that may qualify for inclusion*" (ISAAC, 2014, p2).

A logical next step may be identification and searching of databases that index the material found through ancestry searches and other means.

ISAAC members supplied many items that may have provided a starting point for identification of suitable databases in this manner, if an unbiased review of evidence was genuinely being attempted. There is no evidence that any such efforts have been made.

Appendix 6

Item Six

The search items are stated and appropriate for each database (Schlosser, Wendt and Boesch, 2009).

Rating: No.

6.1 Were the search items used by the ad hoc *Committee stated and appropriate?*

As described on page 2 of the *ad hoc* Committee report, databases were searched for the free-text phrases “facilitated communication”, “supported typing”, and “assisted typing”. Thus, the terms have been stated. It must be assumed that these same key phrases were used to search each database.

In respect of the question “were the search terms appropriate?”, the chosen key phrases seem excessively broad, likely to produce an unwieldy volume of material requiring subsequent filtering.

Appropriate definition of the research question or questions may have suggested key search items suitable for a more targeted and transparent search process, however as noted in Appendix 1, Section 1.4, research questions have not been articulated by the *ad hoc* Committee.

As noted in my response to EVIDAAC Item 2, there is no transparent relationship between the methods described in the “Search for Studies and Reviews” section of the Report and the list of material ultimately considered by the *ad hoc* Committee. Any filtering process should be described and open to scrutiny, in the same way that the overall search protocol should be open to scrutiny. The Report provides no explanation of the process that resulted in the narrow list of materials ultimately considered by the *ad hoc* Committee.

6.2 Were the search items used by Probst (2005) stated and appropriate?

The sources ultimately considered by the *ad hoc* Committee consist of a German-language review (Probst, 2005) and three individual studies that had not been considered by Probst (2005). Probst (2005) is intended to capture data from 37 reviewed studies, and as such is the principal source for the *ad hoc* Committee’s report. It is therefore appropriate to consider whether Probst (2005) met the requirements of EVIDAAC Item 6.

The English language synopsis of Probst (2005) states that the following search items were used in various combinations: facilitated communication; autism; mental retardation; communication disorders; augmentative and alternative communication; communication skills; interpersonal communication; communication aids (for disabled); language skills; special and remedial education; validity; outcomes of treatment; and instructional effectiveness. Other terms were used for German databases (Schlosser and Wendt, 2008). The synopsis goes on to observe:

> *The search terms were stated, although it would have been beneficial to report these for each specific database—keywords tend to vary across databases* (Schlosser and Wendt, 2008, p83).

In a 2009 presentation describing EVIDAAC, these authors and a colleague state:

> *The search terms used should be listed per data base. A listing of search terms in general without cross-referencing it with certain databases would trigger a "no" response* (Schlosser, Wendt and Boesch, 2009).

Thus, the response to Item 6 in respect of Probst (2005) is "no". As Probst (2005) is the principal source relied on by the *ad hoc* Committee, its failure to meet the requirements of Item 6 must also reflect on the *ad hoc* Committee's report.

Appendix 7

Item Seven:

The criteria for inclusion and exclusion of studies are pre-defined (Schlosser, Wendt and Boesch, 2009).

Rating: No.

7.1 What is meant by "pre-defined" inclusion and exclusion criteria?

In 2007, the Chair of the *ad hoc* Committee and colleagues noted that:

> *… the presence of a protocol is essential for the rigorous implementation of a review. … A protocol is developed* a priori *and serves as a 'road map' by outlining the essential procedures for conducting the review* (Schlosser, Wendt and Sigafoos, 2007, p139).

As discussed in relation to EVIDAAC Item 2, there is no evidence of a pre-existing search protocol, nor that the search methods had been approved by any organisation prior to implementation. Similarly, there is no evidence of a formal protocol outlining inclusion and exclusion criteria. Many criteria simply "appeared" after the material had been coded for inclusion, during the analysis stage of the review.

Why should inclusion and exclusion criteria be pre-defined?

> *The criteria for inclusion and exclusion need to be stated upfront or* a priori *so that the reader knows exactly what it takes for a study to qualify for inclusion* (Schlosser, Wendt and Boesch, 2009).
>
> *… if the authors state their inclusion criteria, it is less likely they will (as they are wont to do) preferentially cite studies that support their own prior conclusion* (Oxman, Cook, & Guyatt, 1994, p1368, cited in Schlosser, Wendt and Sigafoos, 2007, p142).
>
> *… it is likely that coders will preferentially include those studies with results that are consistent with their expectations.* (Schlosser, Wendt and Sigafoos, 2007, p145).

As noted by the Chair of the *ad hoc* Committee and colleagues in 2007:

> *Scope and selection biases can be introduced … depending on the types of constraints the authors employ: geographic constraints; temporal constraints; linguistic constraints; study-design restrictions; population constraints; intervention constraints; and outcome constraints* (Schlosser, Wendt and Sigafoos, 2007, p142).

This Appendix examines inclusion and exclusion criteria under these headings to identify the stage at which they were defined. Please see Table Two for a summary of this material.

The question of whether criteria are appropriate is discussed in my response to Item 8 and in Appendix 8.

7.2 Population, Intervention and Outcome constraints

As discussed in response to EVIDAAC Item 1, neither the population nor the intervention has been operationally defined for FC research. For further information, please see Appendix 1, Sections 1.1 and 1.2 respectively. Correspondingly, population constraints have not been reported, and intervention constraints – reliant simply on practices being labelled “facilitated communication” with no attempt to identify the essential elements – are inadequate. Population and intervention constraints cannot, therefore, be said to have been “pre-defined”.

Regarding outcome constraints, only one outcome measure is ultimately considered by the ISAAC *ad hoc* Committee: communicative competence measured as a percentage of correct responses under controlled experimental conditions, used to inform “cause and effect” conclusions regarding authorship. Limitations of this outcome measure are discussed in Appendix 1, Section 1.3. Appendix 8.2 discusses whether the outcome constraint is “appropriate”. The purpose of the current Section is to consider whether it was “pre-defined”.

It may be true that, despite the ISAAC Council’s 2012 directive to broaden the scope of the review, the *ad hoc* Committee always intended to exclude material that did not, by their definition, directly address authorship. However, this decision was not communicated to the ISAAC membership when the changes were announced in 2013.

Nor, apparently, was the constraint made clear to readers coding material for inclusion/ exclusion. The “*analysis level of inclusion considered appropriate given the nature of the material*” (ISAAC, 2014, p3) was identified using a checklist (Appendix A of the *ad hoc* Committee’s report). With the exception of material deemed not to be about FC, this checklist provides no indication of how material is to be treated once coded. The list of materials on pages 4 to 5 of the Report gives the impression that material from all four levels has been included.

As the decision to impose a narrowly defined outcome constraint is made explicit only during the “*Analysis and Synthesis*” stage, it cannot be described as “pre-defined”.

7.3 Study Design Restrictions

The study design criteria that resulted in exclusion of a broad range of the material that ISAAC International had invited through its 2013 “*Adjustments to the framework within which the FC Committee is conducting its efforts*” (ISAAC, 2013a) was not pre-defined. As noted above, the list on pages 4 to 6 of the *ad hoc* Committee’s Report implies that, after coding, four levels of material were identified for inclusion. However, all material not deemed “Level One” – quantitative experimental data addressing the issue of authorship through confrontational testing or message passing – was subsequently excluded during the analysis phase. This exclusion criterion is linked to the outcome constraint discussed in Section 7.2 (above) and in Appendix 8, Section 8.2.

The study design of Level 1 material is described as follows:

> *"*[Q]*uantitative experimental data"* [includes] *studies (or reviews of such studies) that involved an* a priori *controlled manipulation of what knowledge/stimuli was presented to the facilitator and the individual using FC in an attempt to empirically establish who was authoring the messages produced in response to the stimuli* (ISAAC, 2014, p3).

Please see Appendix 8, Section 8.3, which examines whether exclusion of all material not deemed "Level 1" was "*appropriate given the purpose of the review*" (Schlosser, Wendt and Boesch, 2009). The significance of different study designs is also discussed in that Section.

A further study design constraint is described on page 7 of the Report. The *ad hoc* Committee:

> *… relied on the body of studies that included a facilitator-blinded and a facilitator-non-blinded condition (studies without a non-blinded condition are not as convincing from a methodological point of view)* (ISAAC, 2014, p7).

The implications of this further study design constraint are not made clear in the *ad hoc* Committee's report, however it seems unlikely that it has made a significant impact on the outcome of the review. It is not further discussed in Appendix 8.

7.4 Geographic, Temporal and Linguistic Constraints

The *ad hoc* Committee does not appear to have imposed significant geographic constraints on the scope and selection of material for this review.

Temporal constraints were not pre-defined, but instead were applied during the analysis phase of the review. Six reviews were excluded as "dated": Cummins & Prior (1992), Felce (1994), Jacobsen et al. (1995), Kezuka (2002), Mostert (2001), and Simpson & Myles (1995a). No clear definition of "dated" was given.

It may be worth observing that the included review, Probst (2005), discussed studies published between 1993 and 1998, and could itself be considered "dated".

Pre-defined linguistic constraints were implied by the search methods proposed, as only English language databases were used. English language databases may index some studies and reviews published in other languages, however the search for such studies was not systematic.

On the other hand, linguistic constraints were also revealed during the analysis phase of the process. These constraints were due to limitations on languages accessible to the *ad hoc* Committee. Only French, German, Italian and English material was included.

7.5 Inclusion Criterion: Written Material

Regarding the ISAAC International review, both the original announcement of materials to be considered by the *ad hoc* Committee (ISAAC, 2012) and the subsequent "*Adjustments*" (ISAAC 2013a) made it clear that only written material would be considered. This, therefore, is recognised as a pre-defined inclusion criterion.

Appendix 8, Section 8.5 discusses whether the criterion is "appropriate".

7.6 Exclusion Criterion: materials not about FC

Appendix A of the report (ISAAC, 2014, pp.26-27) is an inclusion checklist developed by the *ad hoc* Committee:

> *… to classify the obtained materials and to decide the analysis level of inclusion considered appropriate given the nature of the material* (ISAAC, 2014, p3).

This document does make explicit, for the benefit of those engaged in classifying materials, the decision to exclude *"Materials that were not about FC"*. This decision is also mentioned under the heading "*Develop and Apply Criteria for Inclusion*" in the ISAAC *ad hoc* Committee's Report (ISAAC, 2014, p3).

Therefore, it is recognised that this exclusion criterion was pre-defined.

Appendix 8, Section 8.6 discusses whether this exclusion criterion was "appropriate".

7.7 Exclusion Criterion: materials not peer reviewed

The decision to exclude materials that are not peer reviewed is reported under the heading "*Develop and Apply Criteria for Inclusion*" in the ISAAC *ad hoc* Committee's Report:

> *… we restricted the inclusion of articles, across all four levels of analysis, to those that appeared in the peer-reviewed journal literature (ISAAC, 2014, p3).*

However, the inclusion checklist (Appendix A, ISAAC, 2014, pp.26-27), does not make clear that "peer review" is an inclusion criteria. If the checklist was, in fact, used as described, the decision to include only peer reviewed material appears to have been taken after classification of the material.

It should be noted that ISAAC's announcement of "*Adjustments to the Framework Within Which the FC Committee is Conducting its Efforts*" on February 8th, 2013 contained the following statement:

> *These modifications are consistent with the direction given to the ISAAC EB by the ISAAC Council, which was ratified by a vote at the Council Meeting on July 29th, 2012.*
>
> *The first modification is that the scope of the review will include both quantitative and qualitative studies. The second modification is that both published and unpublished sources will be included in the search and appraisal process in developing the draft* (ISAAC, 2013a).

Thus, the decision to exclude all material that has not been peer reviewed is in breach of a direction given to the ISAAC Executive Board by the ISAAC Council, ratified by a vote at the Council Meeting on July 29th, 2012.

Please see Appendix 8, Section 8.7 for further discussion of this criterion.

7.8 Exclusion Criterion: Reviews that have been subject to synopsis, and individual studies included in those reviews.

As noted in respect of EVIDAAC Item 3, the method the committee has used to navigate Level 1 material is the 5-S hierarchy, adapted for AAC contexts by Schlosser and Sigafoos (2009). In the 5-S approach synopses of systematic reviews are sought before systematic reviews, with individual studies representing the lowest form of evidence.

The *ad hoc* Committee has stated that:

> *… Systematic reviews are preferred sources of evidence because they have systematically aggregated the existing evidence and minimize error that may arise from relying on any one individual study. The pyramid suggests further that consumers seek out synopses (or appraisals) of systematic reviews before the systematic reviews. This preference for appraisals is based on the premise that not all reviews are created equal in terms of quality of methods to address trustworthiness* (ISAAC 2014, p6).

Adoption of this approach by the *ad hoc* Committee is immensely significant, as it requires the reader to accept conclusions of the authors of synopses or of reviews regarding the quality and validity of the original studies. This, therefore, is another opportunity for "experts" to act as gatekeepers. Appendix 8, Section 8.8 discusses why this is inappropriate.

Intention to use the 5-S hierarchy of evidence is announced on Page 6 of the *ad hoc* Committee's report, under the heading of "*Analysis and Synthesis*" (ISAAC 2014). As such, it was neither pre-defined prior to the search process, nor clearly identified in the "methods" section of the Report.

7.9 Reviews deemed "fully consistent" with Probst (2005)

It is noted that, with the exception of Probst (2005), the *ad hoc* Committee has disregarded all reviews from their list that had not already been excluded due to temporal and linguistic constraints (i.e. Mostert, 2010, 2012, Wehrenfennig & Surian, 2008), as their conclusions are deemed "fully consistent" with those of Probst (2005).

Decisions to exclude studies and reviews as described in this section appear to have been taken during the analysis stage, not in response to any pre-defined exclusion criterion. However, the exclusions appear unlikely to have a significant effect on the *ad hoc* Committee's findings. Neither of these exclusion criteria will be discussed further in Appendix 8.

Table Two: Summary of Material Included and Excluded

A: Items initially included but subsequently excluded

Category of Material	Initially included	Ultimately Excluded	Comment/ Reason and stage at which exclusion applied	Ultimately Included
Level 1 analysis: **Systematic reviews**	10	 ^6 ^1 2 #1	Excluded during analysis due to: • Temporal constraints • Linguistic constraints • Conclusions consistent with Probst (2005) • Deemed commentary - reassigned to Level 4.	1
Level 1 analysis **Studies - from Probst (2005)** • Blinded/ non-blinded conditions. • Blinded condition only. • Controlled assessment of sexual abuse allegations.	 23 13 6	 *23 13 *6	Excluded during analysis using the 5-S hierarchy of evidence, in which appraisals (e.g. Schlosser and Wendt's 2008 synopsis of Probst, 2005) and reviews allow consumers to "*disregard the actual study …*" (Schlosser and Sigafoos, 2009, pp.232-233)	nil
Level 1 analysis: **Studies - not from Probst (2005)** • Blinded/ non-blinded conditions.	3	nil	Methodological shortcomings see limited support for FC in two studies disregarded, but are ignored in the third study because its conclusions: "*are in accordance with those of numerous other controlled studies*" (ISAAC, 2014, p10).	3
Level 2 Analysis.	11	11	Exclusion during analysis as study design does not include *a priori* test of authorship.	nil
Level 3 Analysis.	4	4	As per Level 2.	nil
Level 4 Analysis:	9	#9	As per Level 2. Note that description of this category provided to coders includes both anecdotal material and commentary on research. See Appendix 11.	nil

B: Other excluded items

Category of Material	Number excluded	Comment/ Reason and stage at which exclusion applied
Listed "excluded materials".	78	*Materials not about FC:* Excluded at time of coding "inclusion checklist". *Materials not peer reviewed:* Excluded after "inclusion checklist" had been coded.
"Disappeared" material submitted by members.	+Est. 224	No reference or explanation in *ad hoc* Committee Report.
Video material.	nil	Exclusion confirmed in 2013 Press release: *Information on new submissions to be considered* (ISAAC, 2013b).

* # ^ One item included in more than one list. + Estimate: Crossley, personal correspondence 2014.

Appendix 8

Item 8:

> *The criteria for inclusion and exclusion are appropriate given the purpose of the review* (Schlosser, Wendt and Boesch, 2009).

Rating: No

8.1 What is meant by "appropriate given the purpose of the review"?

As discussed in respect of EVIDAAC Item 7, it seems exclusions occurred either before initial coding took place (see Appendix 2) or after the remaining material had been coded for inclusion, during the analysis stage of the review. There does not appear to have been a protocol document or "road map" outlining the essential procedures for conducting the review.

The *ad hoc* Committee Chair and colleagues noted in 2007:

> *Protocols reduce the likelihood of a biased selection of studies, because the rules for eligibility are set up* a priori … . *If the authors deviated from the protocol, the reader should first seek a rationale for this deviation. Subsequently, the reader needs to make a judgment about how seriously this deviation affected the authors' control over bias* … (Schlosser, Wendt and Sigafoos, 2007, p139).

Schlosser, Wendt and Sigafoos (2007) therefore make the case for defining the inclusion and exclusion criteria prior to the literature search, but leave room for deviation from these criteria if justified:

> *The important factor … is not a blind adherence to protocol structure (which would ultimately produce 'cookie-cutter' reviews), but a demonstration that the reviewers know what they are doing and are able to substantiate the course of action from a logical methodologic perspective* (Schlosser, Wendt and Sigafoos, 2007, p139-40).

This Appendix examines whether the criteria for inclusion and exclusion have been "*appropriate given the purpose of the review*" (Schlosser, Wendt and Boesch, 2009). The structure of this Appendix mirrors that of Appendix 7.

8.2 Population, Intervention and Outcome Constraints

As noted in Appendix 1, Section 1.1, the population is inadequately defined to support appropriate selection of subjects for experimental investigation of FC. Similarly, as discussed in Appendix 1, Section 1.2 and as elaborated in Section 8.3.1 of this Appendix (and in Section 2.1 and Appendix B of my original submission to the ISAAC review) the intervention is inadequately defined to guide experimental procedures bearing more than superficial resemblance to FC. Under these circumstances, neither the population nor the intervention constraints can be considered "appropriate" to the research question.

This being so, understanding of the population and of the intervention is currently inadequate to support any experiment capable of meeting the outcome restraint imposed by the *ad hoc* Committee: communicative competence measured as a percentage of "correct responses", from which "cause and effect" conclusions about authorship may reasonably be drawn.

Limitations of this outcome measure are discussed in Appendix 1, Section 1.3.

Even ignoring the shortcomings associated with inadequate or inappropriate definitions of population, intervention and outcome, as discussed in Section 8.3.1 most studies deemed "Level 1" by the *ad hoc* Committee are incapable of demonstrating anything other than the widely accepted facts that FC users have difficulty with confrontational testing or message passing in contrived circumstances, and that facilitators can influence communications under those conditions. They are certainly unable to address such important questions as:

> *…how does influence happen, how much of it is there, when and where does it occur, and does the intervention (on balance) infringe or enhance the rights of the client?* (Grayson, 1997, p232)

It seems likely that the ISAAC Council recognised the limitations of existing experimental studies, and that this contributed to its July 2012 directive to the ISAAC Executive Board to broaden the range of evidence to be considered. (Please see Appendix 1, Section 1.4 for further details). However, in apparent contempt of the ISAAC Council's directive, the *ad hoc* Committee's Report has adopted a narrowly defined and inappropriate outcome constraint. This, in turn, has resulted in study design restrictions discussed in the next Section.

8.3 Study-Design Restrictions

As discussed in Appendix 7, Section 7.3, exclusion of a broad range of the material that ISAAC International had invited through its 2013 "*Adjustments to the framework within which the FC Committee is conducting its efforts"* (ISAAC, 2013a) did not occur in response to *a priori* criteria. The *ad hoc* Committee sought to justify these exclusions by reference to the outcome constraint discussed in the previous Section, which had been introduced during the analysis stage of the review. They imply that only Level 1 studies are capable of providing information on cause and effect relationships between the facilitator, the individual with complex communication needs, and the facilitated communication.

This implication is not correct. To quote a respected text on experiments and causation:

> *Many factors are usually required for an effect to occur, but we rarely know all of them and how they relate to each other. This is one reason … why a given causal relationship will occur under some conditions but not universally across time, space, human populations, or other kinds of treatments and outcomes that are more or less related to those studied. To different degrees, all causal relationships are context dependent, so the generalisation of experimental effects is always at issue* (Shadish, Cook and Campbell, 2002, p5).

To illustrate this point, the authors describe research into a potential cancer treatment, Endostatin, during the 1990s:

> *Other respected researchers could not replicate the effect even when using drugs shipped to them from Folkman's lab. Scientists eventually replicated the results after they had travelled to Folkman's lab to learn how to properly manufacture transport, store, and handle the drug and how to inject it in the right location at the right depth and angle. One observer labelled these contingencies the "in-our-hands" phenomenon, meaning "even we don't know which details are important, so it might take you some time to work it out"* (Shadish, Cook and Campbell, 2002, pp4-5).

Shadish, Cook and Campbell make the limitations of experiment-based research very clear:

> *Although experimenting on manipulable causes makes the job of discovering their effects easier, experiments are far from perfect means of investigating causes. Sometimes experiments modify the conditions in which testing occurs in a way that reduces the fit between those conditions and the situation to which the results are to be generalised. Also, knowledge of the effects of manipulable causes tells nothing about how and why those effects occur. Neither do experiments answer many other questions relevant to the real world – for example, which questions are worth asking, how strong the need for treatment is, how a cause is distributed through society, whether the treatment is implemented with theoretical fidelity, and what value should be attached to the experimental results.*
>
> *In addition, in experiments, we first manipulate a treatment and only then observe its effects; but in some other studies we first observe an effect … and then search for its cause, whether manipulable or not. Experiments cannot help us with that search* (Shadish, Cook and Campbell, 2002, p8-9)

In the context of FC research, it is not disputed that there will be a role for experimental studies, however, most of the currently published experiments must be disregarded due to serious inadequacies in their designs. Future research will depend on researchers working together with skilled FC practitioners to identify the elements of “best practice”. It is imperative that the knowledge encompassed by “best practice” is nurtured and supported. In the current climate, there is very real danger that this knowledge will be lost as practitioners change career paths in the face of hostility from agencies and peak bodies such as ISAAC. Similarly, it is essential to individuals who have complex communication needs that the options professionals are able to offer them are not restricted by administrative decisions. Evidence based practice offers a model for matching a client’s highly individual needs with available interventions, using the skills and expertise of individual practitioners informed by available quality research. Administrative edicts restricting the range of interventions that may be offered on the basis of poor quality research must not interfere with this process.

Quality research in the field of AAC is currently very limited. We have seen that the quality of quantitative research targeting FC is particularly poor. The ISAAC review may have been able to provide guidance to future research, if the research questions implied by the ISAAC Council's directive to broaden the range of evidence had been honoured. However the outcome constraint that has been imposed by the *ad hoc* Committee limits the review – like many before it - to uncritical reiteration of the results and conclusions drawn from flawed experimental designs.

Section 8.3.1 will discuss the reasons exclusive inclusion of material deemed "Level 1" by the *ad hoc* Committee is not appropriate given the purpose of the review. Sections 8.3.2, 8.3.3 and 8.3.4 will outline why exclusion of material deemed Levels 2 to 4 is also inappropriate.

8.3.1 Level 1 material

According to the *ad hoc* Committee,

> *Level 1 evidence has the potential to appropriately inform conclusions regarding authorship, and hence, the validity of FC. In order to determine who is authoring the messages, it is required that an experimental design be implemented in which conditions (e.g., blinded facilitator, non-blinded facilitator, facilitated, not facilitated) are established* a priori *and manipulated while assessing the impact of each condition on the output generated* (ISAAC, 2014, p6).

As discussed at some length in my original submission to the ISAAC review (Section 2.1 and Appendix B), experimental designs that fit this description are inadequate to eliminate the obvious rival explanations for their results: that they are artefacts of the intrusive experiment procedures, rather than true reflections of the phenomenon being studied. The consistent conflict between the results of experimental and of observational studies makes it imperative that this possibility is taken seriously and further investigated. As noted by the Australian National Health and Medical Research Council:

> *"Differences in the conclusions reached about effectiveness from studies at differing levels of evidence or within a given level of evidence need to be resolved"* (NHMRC, 2000b, p14).

When experimental evidence is in direct contradiction of clinical observation, BOTH need to be revisited to resolve the contradiction. The indisputable existence of numbers of FC users who achieve independent typing and even ultimately speech demonstrated this point.

Green and Shane (1994, p.163) observed that:

> *Most experiments have some methodological flaws, but when consistent results are obtained in a number of experiments, each using somewhat different methods and flawed in different ways, the evidence has converged and strong conclusions are warranted*.

There is certainly consistency in the results obtained from "Level 1" studies; however there has also been an unfortunate consistency in the design of experiments, with heavy emphasis on confrontational testing or message passing in highly contrived circumstances. As such, it cannot be said that the experiments were *"each using somewhat different methods and flawed in different ways*". Consistent results from such experiments may say more about flaws in the design than about FC.

Conclusions drawn from fundamentally flawed studies must be disregarded. As discussed in Appendix 13, inadequate attention has been paid to the quality of studies included in the ISAAC review.

8.3.2 Level 2 material

The *ad hoc* Committee Report describes "Level 2" material as:

> "[s]*tudies and reviews that included quantitative descriptive data on the output generated through the process of FC without* a priori *testing of authorship*" (ISAAC 2014).

These studies are dismissed because "*numerous alternative explanations cannot be ruled out by using such descriptive designs*" (ISAAC, 2014, p11). However, as noted above, the designs used for "Level 1" studies share this characteristic, given that they are inadequate to eliminate the possibility that the results are mere artefacts of the intrusive experimental design.

Level 2 studies offer legitimate, alternative approaches to examining the question of authorship. For that reason their exclusion from the review is highly inappropriate and clearly indicative of selection bias.

The *ad hoc* Committee's Report states that:

> *Based on the committee's expertise in research design, the committee finds* [Level 2] *evidence as inappropriate in informing the question of authorship, the focus of the committee's work.*
>
> … [T]*hese studies are predicated on the assumption that the participants in their studies are the authors of the messages generated without having engaged in due diligence by verifying that this was indeed the case. In light of the overwhelming Level 1 Evidence for facilitator control this is a tedious assumption at best and an ethically unjustifiable one at worst* (ISAAC, 2014, p11).

Please see my response to EVIDAAC Item 1 and Appendix 1, Section 1.4, which demonstrates that the *ad hoc* Committee's exclusive focus on authorship is in direct contravention with the intentions ISAAC International has expressed in its 2013 "*Adjustments to the Framework Within Which the FC Committee is Conducting its Efforts*" (ISAAC, 2013a).

Regardless of this, however, the assertion that the studies deemed "Level 2" by the *ad hoc* Committee do not address authorship is simply fallacious. As discussed in Section 2.3 of my original submission to ISAAC (Davies, 2012), Grayson, Emerson, Howard-Jones and O'Neil (2011), used specialist eye-tracking equipment and fine-grained video analysis to demonstrate that an FC user was making visually guided, intentional movements towards the letters he selected – that is, that he and not his facilitator was in control of his movement. Bernardi & Tuzzi (2011a, 2011b), used statistical analysis of texts generated at FC sessions to demonstrate that those written by FC users were similar to each other and different from those written by their facilitators (which also resembled each other). These studies are significant, not least because of their ability to provide an alternative approach to examining authorship.

Regarding the suggestion that "due diligence" requires *a priori* testing of authorship, please see Appendix 1, Section 1.3 which discusses appropriate outcome measures in AAC contexts.

It may also be observed that "due diligence" should extend to adequate examination of contradictions between quantitative and qualitative research and, as discussed in Appendix 13, the quality and validity of ALL research, not simply studies that do not accord with the reviewers' viewpoints.

8.3.3 Level 3 material

The ad hoc Committee describes "Level 3" material as follows:

> *Written documents that included qualitative descriptive data on the output generated through the process of FC without pre-testing of authorship were deemed appropriate for Level 3 Analysis. Qualitative data were considered those generated through qualitative research methods, such as participant observations and interviews* (ISAAC 2014, p3).

ISAAC International's 2013 "*Adjustments to the Framework Within Which the FC Committee is Conducting its Efforts*" significantly broadened the range of evidence to be considered in line with a direction given to the ISAAC Executive Board by the ISAAC Council, ratified on July 29, 2012. As stated in the "*Adjustments*":

> *The first modification is that the scope of the review will include both quantitative and qualitative studies* (ISAAC, 2013a).

The subsequent "communication to members" titled *Information on New Submissions to be Considered* provided the following list.

- *published peer reviewed reports on FC*
- *unpublished reports on FC*
- *written anecdotal reports on FC*
- ***qualitative reports on FC***
- *any other written evidence on FC* (ISAAC, 2013b, emphasis added).

Despite this, the *ad hoc* Committee asserts that Level 3 evidence can be disregarded because "*the same conclusions apply as for Level 2 evidence*" (ISAAC, 2014, p12).

Qualitative research designs are essential for addressing the kinds of questions that need to be explored before valid quantitative research can be undertaken. This has been recognised by Speech Pathology Australia:

> *… a strong tradition of rigorous qualitative research is building in the field of AAC and multimodal communication for understanding of all the factors and complex relationships that might exist in particular life situations … . Such rigorously conducted qualitative studies provide evaluations of important topic areas, inform future qualitative and quantitative research, and guide changes to policy and practice to improve the lives of people with complex communication needs* (Speech Pathology Australia, 2012, p.15-16).

As noted in my response to EVIDAAC Item 1 and in Appendix 1, a well-formed research question requires clear definition of the population being studied, the intervention being tested, and the outcomes considered. In the context of FC, none of these three factors has yet been understood in a manner adequate for well-designed experimentation. This is not to say that "narrowly defined variables" will not be identifiable – only that significant qualitative research into the dynamics of the support provided through FC will be required before these factors can be adequately defined for the purpose of controlled, quantitative study. Please see Section 2.2 of my original submission to the ISAAC review (Davies, 2012) for elaboration of the value of qualitative research.

Qualitative studies examining the question of authorship generally involve analysis of communications for idiosyncratic word forms or sentence structure (e.g. Niemi and Kaernae Lin, 2002) or for "*unique physical characteristics in typing or pointing, personal themes, recurring phrases, and stylistic qualities*" (e.g. Biklen, Saha, & Kliewer, 1995).

Other studies examine behavioural changes, including improved communication (e.g. Clarkson, 1994; Janzen-Wilde, Duchan & Higginbotham, 1995) and non-verbal behaviour matching typed messages (e.g. Olney, 1995).

Studies by Emerson, Griffiths, Prentice, Cosham, Howard-Jones & Grayson (1998) and Emerson, Grayson & Griffiths (2001) used a range of these methods, along with intensive video analysis, to provide evidence that FC enhances the communication skills of some subjects. Significantly, some subjects who had been successful in more naturalistic activities had previously failed "message passing" tests under controlled conditions, highlighting the mismatch between the findings of naturalistic methods and those of the flawed experimental designs used throughout the 1990s and beyond.

8.3.4 Level 4 material

The direction regarding Level 4 material that was provided to coders in the "Inclusion Checklist" was:

> *Is the document written by individuals using Facilitated Communication, facilitators, former facilitators, and others who share their perspectives on Facilitated Communication? If "yes" include in Level 4 Analysis. Anecdotal reports* (ISAAC, 2014, p27).

It is important to note that the description of papers that "*share* [the author's] *perspectives on Facilitated Communication*" does not apply exclusively to anecdotal material. It may also apply to discussion or appraisal of experimental or observational research evidence. For example, Mostert (2012) is included in both Level 1 and Level 4 lists. It is described as follows:

> *Mostert (2012) reviewed previous reviews of FC and drew broader conclusions from the more recently generated evidence of the pro-FC movement; no additional studies were reviewed* (ISAAC, 2014).

It seems appropriate to question inclusion of commentary on experimental or observational research alongside anecdotal data. The wording of ISAAC's original (2012) call for submissions included appraisal of research under Item 3a:

> *3 What materials may be submitted?*
>
> *a. "peer-reviewed research evidence (i.e., empirical studies, systematic reviews) and/or appraisal of such evidence" (ISAAC, 2012)*

As illustrated by Mostert (2012), there may be a fine line between a review and a commentary. A number of reviews, all of which reflect negatively on FC, are listed as "Level 1" on page 4 of the report. The description of these as "systematic" is, however, questionable. As observed by Wright *et al* (2007):

> *Historically, expert opinion has been presented in narrative reviews which are not evidence-based, and, consequently have limitations. Unsystematic narrative reviews are more likely to include only research selected by the authors, thereby introducing bias; hence, they frequently lag behind and contradict available evidence* (Wright *et al*, 2007, p23).

Even Probst (2005) may more correctly be seen as a narrative review, as "*the conclusions are based more on descriptive data than on inferential analyses"* (Schlosser and Wendt, 2008, p83). Such reviews have, none-the-less, been coded as "Level 1" and are thus eligible for consideration by the *ad hoc* Committee.

On the other hand, it seems that commentary "*written by individuals using Facilitated Communication, facilitators, former facilitators, and others who share their perspectives on Facilitated Communication"* – even if peer reviewed - is relegated to "*Level 4 Analysis, Anecdotal Reports*" (ISAAC, 2014, p27).

Turning now to the explanation given by the *ad hoc* Committee for exclusion of Level 4 material:

> *Because these anecdotal reports are essentially perspectives of individuals they cannot be taken as scientific evidence for (a) supporting a demonstration of authorship or (b) refuting a demonstration of authorship. Therefore, for the purposes of this position statement, a more in-depth analysis of these perspectives was not warranted* (ISAAC, 2014, p12).

Once again, it is important to recognise that the 2013 "*Adjustments to the Framework Within Which the FC Committee is Conducting its Efforts*" specifically invited "*Unpublished, anecdotal … and any other written evidence on FC*" (ISAAC, 2013a). The *ad hoc* Committee's decision to ignore such material is clearly at odds with ISAAC International's intentions for the review.

It might also be noted that "*Unpublished, anecdotal … and any other written evidence on FC*" (ISAAC, 2013a) – may straddle the divide between research and anecdotal evidence. Each has a potential role in EBP.

Lof, in his invited keynote speech at the 2010 Speech Pathology Australia (SPA) national conference, acknowledged that the "*highly regulated circumstances*" demanded for controlled studies "*make it difficult to apply the studies in the real world*", and that:

> …*evidence from real world clinical practice* ["practice based evidence"] *can add important data to the EBP knowledge base*" by testing "*the impact of an intervention in more typical settings and conditions* (Lof, 2011, p.193).

Speech Pathology Australia (SPA) refers to "[r]*ecently developed models of evidence-based practice*" which "*suggest that speech pathologists undertake a defined process when determining appropriate treatment for clients*" (SPA, 2010, page 4). Three such models are mentioned, but only one is discussed in detail: Gillam and Gillam (2006), "*a readily accessible evidence based practice process*" (Speech Pathology Australia, 2010, page 5).

Gillam and Gillam, rate practice based evidence as follows:

> *The data that individual clinicians collect on the children they treat is weighted as Level 4 evidence. We believe this is consistent with Level 4 external* [i.e. research] *evidence from single case studies. Over time, a clinician may collect data on 15 or 20 children who receive the same kind of treatment. This data would provide a good indication of the range of outcomes that an individual clinician obtains. Clinician-generated outcome data from many children who received the same type of treatment would be consistent with studies of multiple cases (Level 3 external evidence)* (Gillam and Gillam, 2006, page 309).

(Note that the "levels of evidence" referred to by Gillam and Gillam do not correspond with those adopted by the *ad hoc* Committee.)

In line with the principles of Evidence Based Practice (EBP), Gillam and Gillam provide parallel hierarchies for external (research) evidence, internal (practice based) evidence, and for client factors. They do not "*Elevat*[e] *research findings to a position of predominance or supremacy over these other factors, or to the exclusion of these factors"* (Prizant, 2011, p46), but instead offer a model in which practice "*rests on the shared integration of the three cornerstones*" (Schlosser, 2004) – that is, on research evidence, clinical expertise, and client needs and preferences.

Lof cautioned that practice based evidence is only valuable if it has been "*adequately obtained*": That "*extensive clinical experience or poorly gathered internal evidence from that experience is not a valid determiner of 'what works' in therapy*" (Lof, 2011, p.193). This is clearly an important observation.

None-the-less, the perspectives of those who may not have the skills or leisure to "adequately obtain" systematic data should also be valued as potentially pointing the way to issues requiring future research. Such material may originate from FC users, parents and carers, and even clinicians who have a hunch but do not yet have the data to back it up. As Isaac Asimov (the well-known US science fiction novelist & scholar) is said to have quipped: *The most exciting phrase to hear in science, the one that heralds new discoveries, is not* "Eureka" *but* "That's funny" ... (Kline, 2008, p236).

8.4 Geographic, Temporal and Linguistic Constraints

As noted in Appendix 7, Section 7.4, the *ad hoc* Committee does not appear to have imposed significant geographic constraints on the scope and selection of material for this review.

Temporal constraints were applied, with six reviews being excluded as "dated". No operational definition of "dated" was given. The *ad hoc* Committee does not appear to have examined each review in detail prior to their exclusion, but note that:

> *The conclusions reached by Mostert (2001, 2010) are fully consistent with Felce (1994), Jacobson et al (1995), Probst (2005), and Wehrenfennig and Surian (2008)* (ISAAC, 2014, p7).

In view of this consistency, exclusion of these dated reviews is unlikely to have had a significant impact on the outcome of the review.

The *ad hoc* Committee Report provides only one example of a paper that was excluded due to linguistic constraints: Kezuka (2002). However this paper had also been excluded on the basis of temporal constraints.

I am unable to comment on the effect linguistic constraints may have had on the outcome of the review, as I am unable to read languages other than English. I note, however, that of the four papers ultimately considered by the *ad hoc* Committee (see Table Two), only one is in English. The absence of English language translations of the three other documents (Probst, 2005; Perini, Rollo and Gazzotti, 2010; and Schiavo, Tressoldi, & Martinez, 2005) makes it extremely difficult for individuals who read only English to form their own opinion regarding the quality of the papers and their interpretation by the authors. The Report gives no indication of how many members of the *ad hoc* Committee were able to read the four included papers, and how many were thus involved in the review and analysis of the studies. More transparency is required in respect of this matter, if ISAAC members are to make a judgment as to the quality of the review.

8.5 Inclusion Criterion: Written Material

ISAAC International's 2013 communication to members titled *Information on New Submissions to be Considered* stated the following:

> *Video evidence is not included in the review. There are two primary reasons for this decision, namely:*
>
> - *Professional video recording and editing capabilities today are widespread, easy to use, and within the financial reach of even the most*

modest budget. This means that video recordings could very easily be edited or recorded in such a way as to promote one conclusion over the other in this review, without the FC Committee's knowledge.

- *Video recordings show only what is in the camera's field of vision. It is possible that important evidence that should be included in the review would be missed, if it took place "off camera". Not having the ability to be present at the time that the video recordings were made means that the FC Committee would not have access to any of this type of additional information if relevant, which could lead to a potential "skewing" of results* (ISAAC, 2013b)

These comments are not only offensive, but also illogical. As other members commented in their correspondence with ISAAC International at the time of the call for new material, written evidence is very much easier to alter than video evidence.

As we now know, only peer reviewed reports of a narrowly defined experimental design were considered by the *ad hoc* Committee. Thus the additional restriction on video evidence was superfluous.

As to whether it was "*appropriate given the purpose of the review*", as discussed in Section 8.2 (above), in the absence of good quality research, "best practice" provides the most detailed information available regarding the supports encompassed by FC; the characteristics of FC users; and the most meaningful outcomes measures that may be expected from use of the strategy. Video footage of "best practice" may therefore have been informative in a review of the nature implied by the revised list of material to be considered (ISAAC, 2013a and 2013b).

It is significant that no member of the *ad hoc* Committee sought opportunities to meet and interact with FC users face-to-face. This may have avoided the issue of edited video material. Discourse analysis of any such meetings may be informative on a number of levels.

8.6 Exclusion Criterion: Materials not about FC

Exclusion of papers "*because they did not focus on FC or mentioned FC only in a tangential manner" (ISAAC, 2014, p5)* is inappropriate.

In accordance with Lof's description of science based practice (Lof, 2011), in the absence of quality research specifically targeting FC users, research on overlapping populations - including research on sensory and movement differences - informs practitioners who work with clients who use FC. Please see Section 4 of my original submission to the review (Davies, 2012) for further information about the relevance of research that "*did not focus on FC*".

Regrettably, dogmatism around FC has meant that the relevance of new research in areas such as neurology and sensory and movement differences have been ignored or denied by "experts", who act as gate-keepers for the *status quo* in research organizations and peer-reviewed journals. The effect this stonewalling of FC research has had on attempts to re-examine the "evidence" **must** be acknowledged and addressed.

8.7 Exclusion Criterion: materials not peer reviewed

As noted by the Chair of the *ad hoc* Committee and colleagues in 2007, reviews may legitimately encompass a wide range of materials:

> *Sources may include general-purpose databases (such as MEDLINE and PsycINFO), search engines (e.g. Google), meta-search engines (e.g. Dogpile), journals, personal or published bibliographies, trials registers (e.g. Current Controlled Trials, The Cochrane Central Register of Controlled Trials), conference proceedings, books or book chapters, unpublished literature, etc.* (Schlosser, Wendt and Sigafoos, 2007, p140).

In July 2012 the ISAAC Council directed the ISAAC Executive Board that both published and unpublished sources were to be included in the search and appraisal of material by the FC *ad hoc* Committee (ISAAC, 2013a). When unpublished sources are deemed acceptable, it seems entirely reasonable to assume that the term "published sources" should include both peer reviewed and other material. Despite this, the *ad hoc* Committee has excluded from consideration all materials that have not been peer reviewed.

Exclusion of material that has not been peer reviewed is explained as follows:

> *Given that our committee includes researchers/scholars from both qualitative and quantitative traditions in which peer-reviewed articles enjoy the highest regard, we restricted the inclusion of articles, across all four levels of analysis, to those that appeared in the peer-reviewed journal literature.*

The relevance of the Committee's expertise to the decision to exclude material that has not been peer reviewed is unclear, however please see my response to Item 4 and Appendix 4 regarding the importance of unpublished studies. Please also see my response to Item 5 and Appendix 5 regarding the value and importance of a broad approach to searching published materials, in order to avoid Database bias. The central points are that:

- There has been a long disconnect between research and practice that demands explanation
- The gate-keeper role of researchers and academics in peer-reviewed literature, and in determining the direction of research in Universities and other research organisations must be recognised
- The need to actively guard against publication and database bias is widely acknowledged.

It should be recognised as significant that exclusion of material that has not been peer reviewed results in exclusion of personal submissions from members responding to ISAAC International's request for input.

8.8 Exclusion Criterion: Reviews that have been subject to synopsis, and individual studies included in those reviews.

The method used by the *ad hoc* Committee to navigate Level 1 material is the 5-S hierarchy, adapted for AAC by Schlosser and Sigafoos (2009). This hierarchy ranks synopses of systematic reviews above the reviews themselves, and ranks individual studies as the lowest form of evidence. Synopses and reviews are preferred: "*because someone other than the practitioner who intends to use the evidence has appraised the reviewed studies*" (Schlosser and Sigafoos, 2009, p226).

In the 5-S approach, reliance on a synopsis of a review:

> … *may save time and relieve the practitioner of the need to judge the soundness of a systematic review. Instead, he or she could rely on the appraisal and consult the original review only for gleaning the major conclusions that bear relevance to the question at hand. Similarly, by identifying an appraisal of a pertinent study first … , a practitioner would be able to disregard the actual study from his or her decision-making without having to read the original study* (Schlosser and Sigafoos, 2009, pp.232-233).

The 5-S hierarchy is clearly an approach intended to assist busy practitioners in their efforts to integrate research into their evidence-based clinical decisions. However, suitability of the approach for the purpose of research – and especially for research intended to address areas of controversy - is not discussed in the literature cited by the committee.

The approach is in significant contrast with other hierarchies of evidence such as NHMRC (2000b), Gillam and Gillam (2006), and – the example given by Schlosser and Sigafoos (2009) – the American Speech–Language–Hearing Association (ASHA, undated). As discussed in relation to EVIDAAC Item 3, neither the synopsis (Schlosser and Wendt, 2008) nor the review (Probst, 2005) used by the *ad hoc* Committee would be admitted as evidence under the ASHA hierarchy, which places "*meta-analysis of randomized controlled trials*" as the highest level of evidence (ASHA, undated) but makes no reference to systematic or narrative reviews of non-randomised trials.

The danger in accepting "authority" rather than re-examining individual studies for quality and threats to validity is that errors can, all too easily, be replicated. Most widely used "hierarchies of evidence" either exclude "expert opinion" (e.g. NHMRC, 2000b) or list it as the lowest form of evidence (e.g. Gillam and Gillam, 2006, ASHA, undated). Further, a reliance on expert opinion is not consistent with Evidence Based Practice, which:

> … *intentionally deemphasizes the role of expert authority and instead promotes a transparent, rational decision-making process that can be taught, refined, and applied by all clinicians* (Satterfield *et al,* 2009, p371).

My original submission to the ISAAC review stressed that it was of the utmost importance that ISAAC's readers examine original reports of experimental studies, rather than relying on "reviews" that collect and collate findings without providing detail of the flawed processes they have been drawn from. Please see Appendix B of my submission (Davies, 2012) for further information.

As a result of the study design constraint that restricted the ad hoc Committee's considerations to Level 1 material, combined with the 5-S approach which excludes reviews that have been subject to synopsis and individual studies covered by those reviews, the *ad hoc* Committee Report has been based entirely on the findings of four papers: the synopsis of Probst (2005) undertaken by the Chair and a colleague, and three individual studies of dubious quality. (Please see my response to Item 13 and Appendix 13 for a discussion of the quality of these studies). The impact this approach has had on source selection bias in the Review is discussed in my response to EVIDAAC Item 3 and in Appendix 3.

In short, exclusion of studies and reviews through the 5-S approach is not appropriate for the purpose of the review.

Appendix 9

Item 9:

A log of rejected studies is reported/available upon request (Schlosser, Wendt and Boesch, 2009).

Rating: No

As noted by the *ad hoc* Committee Chairman and colleagues in 2007:

> *Although included studies are needed to conduct the systematic review, the role of excluded studies is not to be underestimated. After all, listing excluded studies allows discerning readers to see for themselves whether or not the inclusion and exclusion criteria have been properly applied. In fact, solid reviews provide examples of studies that were excluded on the grounds of particular criteria* (Schlosser, Wendt and Sigafoos, 2007, p144)

The *ad hoc* Committee report does provide lists of material included at different levels, and a list of excluded material. However, these are incomplete and misleading.

It is apparent that some exclusions occurred before initial coding took place (see Appendix 2). These include an estimated 224 items submitted by members in response to ISAAC International's invitation (ISAAC, 2013a and 2013b), but never mentioned in the Report. There is no log of these omitted studies, nor any explanation given for their omission.

Only exclusion of "materials not about FC" occurred at the time of coding for inclusion/ exclusion, while exclusion of "materials not peer reviewed" appears to have occurred after coding but before further analysis. On the basis of these two exclusion criteria, pages 4 to 6 of the *ad hoc* Committee's Report imply that 76 items have been included and 78 excluded from consideration.

However, as discussed in relation to EVIDAAC Items 7 and 8, the *ad hoc* Committee ultimately included only four items, with 72 items being excluded during analysis. Table Two summarises these exclusions.

Appendix 10

Item 10:

> *A reasonable percentage of studies (≥ 20%) is evaluated reliably for inclusion by more than one rater*. (Schlosser, Wendt and Boesch, 2009).

Rating: **No**

As stated by the Chairman of the *ad hoc* Committee and colleagues in 2007:

> *As part of the appraisal, the reader should assess whether the inclusion and exclusion criteria were applied in a reliable manner. In other words, the reader must appraise the systematic review for selection bias. Specifically, a reasonable and minimum percentage (20–30%) of randomly selected studies under consideration for inclusion should be evaluated by at least two independent raters. Sound systematic reviews will offer inter-rater agreement percentages or correlation coefficients as estimates that the process was handled reliably* (Schlosser, Wendt and Sigafoos, 2007, p144).

The process undertaken by the *ad hoc* Committee in determining the "level" of different material is described as follows:

> *The Chair of the* ad hoc *Committee coded all potential written documents found through the search or submitted to the ISAAC office for inclusion. Exceptions are those articles that were written in languages beyond the competence of the Chair which was restricted to German and English. In these cases two other members did the reading. A large percentage (approximately 60%) of written documents was independently coded by a second member of the* ad hoc *Committee. Any disagreements between the two coders were resolved through consensus-building* (ISAAC, 2014, pp3-4).

However, as discussed in Appendix 2, Section 2.4 there is reason to doubt adherence to the "pre-defined" methodology. It seems apparent that some exclusions occurred before coding took place and, as discussed in respect of EVIDAAC Items 7 and 8, others occurred during the analysis phase of the review after coding. Thus, confirmation that a large percentage of documents was independently coded by a second member of the *ad hoc* Committee provides no information regarding the decision-making process at other stages of the Review.

Further, it should be noted that no inter-rater agreement percentages are made available.

Taking an alternative approach, as the ISAAC review has been based almost exclusively (with the exception of three individual studies) on Probst (2005), it is apposite to examine Probst in terms of EVIDAAC Item 10. Regretably, it appears that Probst (2005) does not meet the required standard:

> *The author seems to have made decisions regarding the inclusion or exclusion of studies … . The involvement of an independent second rater and the gathering of inter-rater agreement data would have allowed for a better assessment of the reliability of the data generated by this review* (Schlosser and Wendt, 2008, p83).

Appendix 11

Item 11

*The coding categories refer to the types of information or data that are being extracted from each included study. These categories need to be stated upfront/*a priori (Schlosser, Wendt and Boesch, 2009).

Rating: No.

11.1 Were coding categories stated a priori*?*

In 2007, Dr Schlosser and colleagues stated that:

> *Readers of systematic reviews want a clear outline of the process by which data was extracted from the original studies. Such a summary should state which studies were coded and who coded them, how they were coded, what coding categories were applied, and how they were defined. It is helpful if the authors reference a coding protocol and perhaps even make it available upon request …* (Schlosser, Wendt and Sigafoos, 2007, p145).

In the ISAAC review, data extraction using *a priori* coding categories appears to have been limited to the coding of material against "criteria for inclusion". Other decisions regarding inclusion/ exclusion were taken before coding (that is, items submitted by members or identified through the search process that were excluded without mention) or during the later, analysis phase of the review. Please see my comments in respect to EVIDAAC Items 7 to 10, which address this process.

As shown in Table Two (Appendix 7), only one review (Probst, 2005) and three individual studies (Olney, 2001; Schiavo, Tressoldi, & Martinez, 2005; and Perini, Rollo and Gazzotti, 2010) were ultimately identified for further analysis by the *ad hoc* Committee. Items 11 and 12 address the process by which data was extracted from these papers.

11.2 Data extraction bias

The Chair of the *ad hoc* Committee and colleagues have noted that outcomes may bias a reviewer's selection and treatment of individual studies:

> *Some workers in the field of research synthesis have argued that the outcomes presented in the original studies may bias the coders … . If, for example, an outcome of an intervention contradicts the coder's expectation, he or she may not code other variables of this study accurately* (Schlosser, Wendt and Sigafoos, 2007, p145).

Schlosser, Wendt and Sigafoos (2007) suggest that one way such outcome bias may be avoided is to have readers view the introduction and methods sections of a study separately from the results and discussion sections. Such an approach, however, will not protect against outcomes bias if a study's results are artefacts of the study design, as I have argued is the case in respect of studies deemed "Level 1" evidence by the *ad hoc* Committee. For more details see Appendix 8, Section 8.3.1 and – in my original submission to the ISAAC review (Davies 2012), Section 2.1 and Appendix B.

Remarkably, despite their flaws, some of the controlled studies reviewed by Probst (2005) produce some results that are potentially supportive of Facilitated Communication. These included Calculator and Singer (1992); Calculator and Hatch (1995); Bebko, Perry and Bryson (1996); Cardinal, Hanson, Wakeham (1996); Heckler (1994); Sheehan & Matuozzi (1996); Vazquez, (1995); and Weiss, Wagner, & Bauman (1996).

So, too, do two of the three individual studies ultimately included by the *ad hoc* Committee (Olney, 2001 and Schiavo, Tressoldi, & Martinez, 2005).

It is recognised that the supportive evidence provided in these studies is quite limited. While perhaps adequate to "falsify" bald assertions that "*FC does not work*" (e.g. Lof, 2011. p192), it certainly does not match the impressive results reported from observational studies. None-the-less, the fact that any supportive evidence at all was produced in the face of severely flawed experimental designs must be acknowledged as quite remarkable.

Results potentially supportive of FC are not elaborated in the synopsis of Probst (2005), although they are hinted at under the heading "Are there individual cases in which 'unexpected communicative performances' have been observed under FC conditions?":

> *None of the studies unequivocally demonstrated unexpected literacy gains due to the levels attained corresponding to previously observed levels under non-FC conditions, or to poor internal validity of the study* (Schlosser and Wendt, 2008, p82).

It is significant to note that, according to the synopsis, the only studies coded by Probst (2005) for "threats to internal validity" were the seven that fitted the criteria concerning "unexpected communicative performance". Data extracted in relation to these "threats" were coded as "*inconsistency of performance, facilitator effect, consistency of the performance with individual developmental norms, social-psychological influences*" (Schlosser and Wendt, 2008, p82-83).

No mention is made of threats to validity posed by inappropriate selection of subjects, intervention that bears only superficial resemblance to FC, controls that interfere with the support needs of the subjects and the support that can be provided by the facilitators, or any of the other issues I have discussed in Appendix 8, Section 8.3.1 and in my original submission to the ISAAC review (Section 2.1 and Appendix B).

It is clear that this selective approach to analysis – detailed examination only of those seven studies that fitted the criteria concerning "unexpected communicative performance" – is inappropriate. However it is typical of the approach taken in the reviews cited as "Level 1 material" by the *ad hoc* Committee. As noted in a literature review commissioned by Disability Services Queensland in 2006:

> *There is bias evident in the reviews to the extent that the reviews critically evaluate only the studies in which positive effects of FC have been obtained; very*

little attention is paid to the methodology or the validity of the results of the studies where negative effects have been found. When one examines the methodology in these studies very many of them would fail on a test of "best practice" in experimental procedures. (Tehan & Senior, 2006, p12).

11.3 Who extracted the data for the ISAAC review?

The reader must also consider who was involved in the coding of included studies, what their qualifications are, and what training they received in the coding procedures (Schlosser, Wendt and Sigafoos, 2007, p145).

Consideration of Probst (2005) by the *ad hoc* Committee is intended to capture data from 37 individual studies. It is difficult to assess the reliability of data generated by the review, however, as there are questions regarding the coding process and the lack of an independent second coder to generate a measure of inter-rater agreement. Please Appendix 12, Section 12.1, for more details.

Regarding the *ad hoc* Committee's review of the three individual studies (Olney, 2001; Schiavo, Tressoldi, & Martinez, 2005; and Perini, Rollo and Gazzotti, 2010), The Report does not indicate which members of the *ad hoc* Committee were involved in reviewing these studies. As two of the studies were in languages other than English, it would be interesting to know how many members of the Committee had the capacity to participate.

The issue of "*who was involved in the coding of included studies"* (Schlosser, Wendt and Sigafoos, 2007, p145) is significant in any review of an issue as polarised as FC. In my email to ISAAC International dated 29 April 2013, I expressed the following concern:

> ***Transparency:*** *It is difficult to have faith in secret deliberations by an anonymous committee whose chair has, recently and very publicly, denounced FC. Such a lack of transparency threatens to discredit, not only the review process, but also the organization that has sanctioned it* (Cathie Davies, email to ISAAC via Franklin Smith, 29 April 2013.)

This comment was further to my earlier email, in which I made the following points:

> ***Anonymity of the committee:*** *The press release has confirmed that the committee will remain anonymous until release of a "draft final document".*
>
> *Any process whereby submissions are considered by anonymous individuals behind closed doors cannot be conducive to the kind of conversation that allows ideas to be tested, shaped and challenged – the kind of conversation we should be having.*
>
> *I have put a great deal of effort, in good faith, into assisting ISAAC in their investigation of FC. It is difficult to feel confident in the process when the only committee member whose name has been released goes to press – more than two months after his appointment as chairman of the ISAAC review committee! - with the statement:*

> *Despite overwhelming evidence negating the validity of Facilitated Communication (FC), this thoroughly discredited technique continues to be promoted and practiced, with sadly tragic consequences* [Schlosser and Sigafoos, 2012. P.1]

My son, our family, and many, many others now and in the future are counting on ISAAC's integrity in this review. A more transparent process would go a long way towards reassuring us that our trust is not misplaced.

When the membership of the *ad hoc* Committee and the names of outside reviewers became public, it was clear that the Chairman was not the only one who had publicly denounced FC. Conversely, no FCT practitioners or FC users were represented.

It must be remembered that:

> "*In science, the importance of our viewpoint when examining information of any kind cannot be too strongly emphasized, for how we look at a thing determines what we see"* (Rothchild, 2006, p.5).

In an issue where viewpoints are so polarized, the question of who should review the evidence is fraught. As noted by renowned twentieth century philosopher of science, Thomas Kuhn:

> *"When paradigms enter, as they must, into a debate about paradigm choice, their role is necessarily circular. Each group uses its own paradigm to argue in that paradigm's defense. …* [T]*he circular argument … can not be made logical or even probabilistically compelling for those who refuse to step into the circle. The premises and values shared by the two parties to a debate over paradigms are not sufficiently extensive for that" (Kuhn, 1962, p94).*

A fresh look at the evidence by individuals who have not previously taken a position on FC is essential. Kuhn noted that revolutions could only be resolved:

> *… by men so young or so new to the crisis-ridden field that practice has committed them less deeply than most of their contemporaries to the world view and rules determined by the old paradigm* (Kuhn, 1962, p144).

However, it is essential that such individuals are given the best opportunity to assess the evidence placed before them, and for this they must hear voices from both sides of the debate. Even if some members of the *ad hoc* Committee had been uncommitted, the highly prejudicial selection of material for the review allowed only one side to be heard.

A "stacked" Committee and the silencing of one side of the debate will resolve nothing.

11.4 Interpretation of the included studies

The synopsis summarized the conclusions of Probst (2005) as follows:

> *On the basis of this systematic review of 37 studies involving 343 subjects, the author concluded that there is overwhelming evidence for a facilitator effect, or facilitator control: the results regarding pseudo-correct responses are consistent*

> *with the hypothesis of direct physical control by the facilitator. Further, communicative performance under controlled FC conditions was shown to be poor. Finally, none of the subjects yielded unexpected literacy gains. These results reveal FC to be an invalid treatment with great potential to have harmful psychological and social side effects* (Schlosser and Wendt, 2008, p82-3).

As noted in Section 11.2, some individual studies reviewed by Probst (2005) produced results that may be interpreted as supportive of FC. These have been selected for particular scrutiny, unlike studies where only negative effects have been found. Potentially positive findings are not reported in the synopsis and are clearly not reflected in Schlosser and Wendt's strong statement (above).

Has similar selective treatment occurred in the ISAAC report?

The *ad hoc* Committee Report stated that:

> *The three additional studies identified at Level 1 support the conclusions reached by Probst (2005) and subsequent reviews"* (ISAAC, 2014, p11).

As previously noted, two of these three studies are in languages other than English, and I am therefore unable to examine them in detail. However, English language abstracts were provided as part of the *ad hoc* Committee's Report. These make it apparent that two of the three individual studies do provide some limited support for FC.

In Olney (2001), for example:

> *FC users were asked to respond to multiple-choice, vocabulary-based computer game items in both nonblind and blind conditions. … Although none of the participants had revealed literacy in previous assessments, four responded to game items at a greater-than-chance level* (ISAAC, 2014, p8)

Similarly, the English abstract of Schiavo *et al* (2005) states that, in confrontational testing, subjects:

> *… were able to give some independent answers, in despite* (sic) *of the influence of the facilitator: this is noteworthy for people who often communicate just by behavioural problems* (Schiavo et al, 2005, quoted in ISAAC, 2014, p11).

The *ad hoc* Committee reviewer or reviewers dismiss Olney's findings by claiming that:

> *The internal validity of this study is fundamentally flawed, in particular, by (a) lack of pretesting, (b) lack of control for unspecific factors that may confound the independent variable ("blind-nonblind") with "participant training" and "assessment implementation" variables, and (c) selective consideration of outcomes in favor of the FC-is-valid-claim* (ISAAC, 2014, p8).

In respect of Schiavo *et al* they find that:

> *The results along with the appraised shortcomings do not support FC as a valid method* (ISAAC, 2014, p11).

The “*appraised shortcomings*” referred to are:

> *It was not reported (a) how the tasks were assigned to the participants, (b) whether the facilitators knew the item pool, (c) whether one or more facilitators were the same for more than one participant, (d) whether the facilitators knew each other. It is unclear who decided on the correctness of a response; there were no controls in place to minimize the possibility of false point assignments (e.g., no inter-rater agreement of any kind)* (ISAAC, 2014, p11).

These may be legitimate concerns, but once again it must be recognized that studies that provide no support for FC have not been subjected to this level of scrutiny by the reviewers.

The third study included by the *ad hoc* Committee, Perini *et al* (2010), provides an example of this unequal treatment. **S**erious threats to internal validity are recognized by the *ad hoc* Committee, but disregarded and the conclusions deemed sound because: “[i]*n spite of the methodological deficits described, conclusions are in accordance with those of numerous other controlled studies*” and *“are concordant with conclusions presented both in systematic reviews and in several position statements published by academic and professional groups* (ISAAC, 2014, p10).

The differential treatment of these three studies on the basis of their conclusions appears to be straightforward evidence of bias based on the researchers’ expectations and personal viewpoint.

The value that the reviewers apparently placed on concordance of Perini *et al* (2010) with “*several position statements published by academic and professional groups”* (ISAAC, 2014, p10) is worth further examination. The NHMRC hierarchy of evidence (National Health and Medical Research Council, 2000b, page 10) explicitly states that:

> *Current levels of evidence exclude expert opinion and consensus from an expert committee as they do not arise directly from scientific investigation* (Mazerole and Legosz, 2012, page 35).

Other hierarchies do include *“expert committee reports, conference proceedings, and/or opinions of respected authorities”* (Gillam and Gillam, 2006, page 308), but only as the lowest form of evidence.

Accepting the findings of a flawed study because they agree with expert opinion turns conventional hierarchies of evidence on their head.

11.5 Are the ad hoc *Committee’s interpretations valid?*

Olney (2001) acknowledges that:

> *Participants in this study responded more accurately to computer game items in the nonblind condition than in the blind condition. This pattern of responses is consistent with other controlled studies of facilitated communication* (Olney, 2001).

This finding does demonstrate a level of facilitator influence – a finding that is not difficult for FC supporters to accept. However, the *ad hoc* Committee has interpreted evidence of facilitator influence as "*lack of evidence of validity*" (ISAAC, 2014, p8). Is this a reasonable conclusion to draw?

As noted in Appendix B, Section B.7.5, of my 2012 submission to ISAAC (Davies, 2012), the conclusion most commonly drawn from findings of facilitator influence is that FC users are passive intermediaries in the communication process. This interpretation is emphasized in the term used by the *ad hoc* Committee – "*facilitator control*". However, this interpretation ignores the possibility that at times – for example, when unsure what is expected of them - FC users may actively seek cues from their facilitators. As suggested by one commentator:

> *Facilitator influence must go on, since quite apart from the physical contact, influencing one's communication partners is a feature of any interaction. It is important to recognise this, for it reminds us to take one step back from the 'all-or-nothing' feel to some of the FC literature. The really difficult questions that need answering are not as simple as 'does facilitator influence go on'? Of course it does. What we need to know is how does influence happen, how much of it is there, when and where does it occur, and does the intervention (on balance) infringe or enhance the rights of the client?* (Grayson, 1997, p232)

The findings of Schiavo *et al* (2005), that subjects "*… were able to give some independent answers, in despite* (sic) *of the influence of the facilitator*" (Schiavo et al, 2005, quoted in ISAAC, 2014, p11) makes it clear that facilitator influence should not be interpreted as exclusive control of the communication by the facilitator.

Research shows that facilitator influence may be significantly reduced through best practice and appropriate facilitator training (e.g. Cardinal, Hanson, Wakeham, 1996). Similarly, awareness that facilitator influence can happen has significantly informed the development of FC "best practice" and is reflected in:

- rigorous facilitator training;
- emphasis placed on independent communication as the ultimate goal of Facilitated Communication Training (FCT);
- development of protocols to (for example) collect ongoing validation data; and
- development of a rigorous protocol for validating controversial or life-changing communications.

11.6 The implications of statistical weakness

The *ad hoc* Committee Report states, in respect of Olney:

> *While some of the participants were literate and were the authors of the letter completing task, the outcomes of these subjects were not better with FC than without* (ISAAC, 2014, p8).

Four individuals in this study achieved statistically different scores under "facilitator blind" conditions. It is not correct that the outcomes of these individuals were not better with FC than without. They were consistently better – however the difference was not statistically significant "*probably due to the small sample size*" (Olney, 2001).

This issue of statistical weakness is likely to affect many FC studies, as the number of subjects is generally small and the sample has not been chosen randomly from a clearly defined population. It is important to note that acknowledging that a finding is not statistically significant and cannot therefore be taken as evidence in support of FC, is very different from the suggestion that it supports the findings of Probst (2005), i.e. that "*results reveal FC to be an invalid treatment"* (Schlosser and Wendt, 2008, p83).

Olney's own conclusion is very much more pertinent: *Clearly, the dialogue about the usefulness and validity of FC is not finished* (Olney, 2001).

Appendix 12

Item 12

> *At least a 20% sample of the data are extracted reliably by more than one rater (blinded to the treatments, if applicable)* (Schlosser, Wendt and Boesch, 2009).

Rating: No.

12.1 Coding and Inter-rater agreement

We are advised that:

> *The Chair of the* ad hoc *Committee coded all potential written documents found through the search or submitted to the ISAAC office for inclusion. Exceptions are those articles that were written in languages beyond the competence of the Chair which was restricted to German and English. In these cases two other members did the reading. A large percentage (approximately 60%) of written documents was independently coded by a second member of the* ad hoc *Committee. Any disagreements between the two coders were resolved through consensus-building* (ISAAC, 2014, p3-4).

Although the described process appears to be in accordance with Item 12, it appears to have applied only to the coding of material against criteria for inclusion/ exclusion, not to data extraction from included papers. No coding categories were provided to guide data extraction from the three individual studies identified for inclusion in the ISAAC review (Olney, 2001; Schiavo, Tressoldi, & Martinez, 2005; and Perini, Rollo and Gazzotti, 2010). The Report does not indicate which members of the *ad hoc* Committee were involved in reviewing those papers, nor what process was followed for the review.

More information is available in respect of Probst (2005) which, under the 5-S hierarchy of evidence is intended to capture data from 37 individual studies. As previously noted, the review itself is in German and I am unable to read it – however the 5-S hierarchy allows the original paper to be disregarded in favour of the English language synopsis (Schlosser and Sigafoos, 2009).

While the synopsis of (Probst (2005) notes that the review: "*makes explicit the coding categories used for extracting data from each study*", it goes on to note that:

> *The author seems to have made decisions regarding the inclusion or exclusion of studies as well as about the extraction of data from included studies. The involvement of an independent second rater and the gathering of inter-rater agreement data would have allowed for a better assessment of the reliability of the data generated by this review* (Schlosser and Wendt, 2008, p83).

In view of inadequate information regarding inter-rater agreement; the apparent lack of process for extracting data from material once it has been coded for "Level of Inclusion"; and questions regarding process and the lack of an independent second rater in the principal source used by the *ad hoc* Committee (Probst, 2005), it is clear that the response to Item 12 must be negative.

12.2 Dispute resolution

In 2007, the Chairman of the *ad hoc* Committee and colleagues noted that:

> *In addition to the primary coder, an independent coder should rate a certain percentage of the studies. Thus, inter-rater agreement data and an estimate of the reliability of the coding process can be reported. …* [I]*f disagreements arose in this process, the reader should find a section* [of the review] *delineating some sort of consensus-building process whereby the coders discussed these discrepancies and tried to resolve them* (Schlosser, Wendt and Sigafoos, 2007, p145).

In the *ad hoc* Committee's report the only reference to a process for resolving disagreement is found at the commencement of the "Methods" section, under the heading "Overall Process":

> *A democratic process was adopted throughout the review process. On most issues we were able reach consensus. Where the committee could not reach agreement on matters we voted and this was documented on e-mail. By voting we agreed to this process (if we had not considered a vote appropriate we could have abstained)* (ISAAC, 2014, p2)

As discussed in Appendix 11, Section 11.3, the Chairman and other members of the *ad hoc* Committee have made their anti-FC views public over many years. It is not possible to avoid the impression of a "stacked" committee. Under those circumstances, abandonment of the attempt to reach consensus and reversion to voting is simply one more means of silencing dissenting voices.

In an earlier paper, the Chair of the Committee and colleagues noted:

> *Readers may come across some reviews that report a consensus building process, but no inter-rater agreement data … . Sound reviews will report both types of information, because the latter is needed to estimate the reliability of the data extraction process* (Schlosser, Wendt and Sigafoos, 2007, p145)

As no such data is made available in the *ad hoc* Committee report, one can only speculate on the nature of any disputes or efforts to resolve disputes. It is not clear whether the resignation of one member of the *ad hoc* Committee on 9 March, 2014 - less than two weeks before the "Final Report" was submitted to ISAAC on 18 March 2014 – related to such a dispute.

Appendix 13

Item 13

> *Criteria used to arrive at judgments of quality are pre-defined and appropriate for the types of included designs* (Schlosser, Wendt and Boesch, 2009).

Rating: No

In the 2009 presentation quoted above, the Chairman of the *ad hoc* committee and colleagues elaborate on Item 13 as follows:

> *This refers to criteria used for appraising the quality … of the included studies, including study design, reliability, treatment integrity, blinding allocation, etc. In rating this item, consider whether the criteria are stated upfront and whether they are appropriate/sufficient to the included treatment designs* (Schlosser, Wendt and Boesch, 2009).

Starting with the principal source used by the *ad hoc* Committee, the synopsis of Probst (2005) states that:

> *… there are several things that could have enhanced the quality of this review further. For example, the criteria used to assess study quality are listed only for some of the subgroups of studies, and the reader is left to wonder whether the same criteria applied to other subgroups* (Schlosser and Wendt, 2008, p83).

Which "*subgroups of studies*" were examined for quality?

> *Although it was stated that data concerning quality of the studies would be extracted, it is unclear what criteria were employed. For the studies related to accusations of sexual abuse, no data categories were reported. For studies concerning unexpected communicative performance, the following data were extracted: communicative performance (i.e. percentage of correct responses); threats to internal validity (inconsistency of performance, facilitator effect, consistency of the performance with individual developmental norms, social-psychological influences); and concluding evaluation of the hypothesis of unexpected communicative performances.* (Schlosser and Wendt, 2008, p81-2).

It would appear, therefore, that the only studies examined for quality by Probst (2005) were those that potentially provided some support for FC. Please see my discussion of data extraction bias in Appendix 11, Section 11.2.

Clearly, Probst (2005) does not meet the requirements of Item 13. This fatally compromises the *ad hoc* Committee's principal source, which is intended to capture the data from 37 individual studies. As stressed in my 2012 submission to ISAAC (particularly Appendix B), reviews that merely collate conclusions from fundamentally flawed studies are as invalid as the studies they have been based on.

The *ad hoc* Committee's Report does comment on the quality of the three individual studies they include alongside Probst (2005), concluding that they are fraught with methodological shortcomings. (see Appendix 11, Section 11.4.) Remarkably, the

Committee decides that the shortcomings serve to negate only the conclusions of the two studies that demonstrate some support for FC. The conclusions of the third study are deemed sound because they: "*are in accordance with those of numerous other controlled studies*" and "*are concordant with conclusions presented both in systematic reviews and in several position statements published by academic and professional groups* (ISAAC, 2014, 10).

Conclusions drawn from fundamentally flawed studies must be disregarded. However, it is important to recognise that disregarded conclusions cannot be taken as support for the opposite conclusion. It is therefore highly inappropriate for the *ad hoc* Committee's "*overall appraisal*" of these additional studies to state that they "*support the conclusions reached by Probst (2005) and subsequent reviews*" (ISAAC, 2014, 11),

It is surprising that the ISAAC review has repeated the error of earlier reviews in collating conclusions without any consideration of study quality. The Chairman of the *ad hoc* Committee and colleagues warned against this inappropriate practice in 2007:

> *Because a systematic review can only be as sound as the included studies, an investigation of the quality of those studies is an important component of the data extraction and appraisal process* (Schlosser, Wendt and Sigafoos, 2007, p145-6).

Several approaches to assessing the quality of studies are explored by Schlosser, Wendt and Sigafoos (2007), however none has explicitly been adopted in the ISAAC review.

Appendix 14

Item 14

Methods used to arrive at judgments of effectiveness for each study are pre-defined and operationalized (Schlosser, Wendt and Boesch, 2009).

Rating: No

Schlosser, Wendt and Boesh elaborate on Item 14 in their 2009 presentation, as follows:

The criteria used to determine effectiveness from the original studies such as "effect size", standard mean difference, percentage of non-overlapping data etc., need to be stated upfront. The criteria need to be operationalized in a way that allows someone to replicate the judgment. If the criteria are stated upfront but not operationalized mark "no." (Schlosser, Wendt and Boesch, 2009).

To be clear:

Operational definitions are the specifications of how variables will be defined and measured (or assessed) in a study (Creswell, 2002, p624).

14.1 What is "Effectiveness"?

In health and medical research, the term "effectiveness" is taken to mean "*The extent to which an intervention produces favourable outcomes under usual or everyday conditions"* (NHMRC, 2000a, p98). This is in contrast to the term "efficacy", which means "[t]*he extent to which an intervention produces favourable outcomes under ideally controlled conditions such as in a randomised controlled trial* (NHMRC, 2000a, p98).

Whatever the intention of EVIDAAC Item 14, it is not possible to consider effectiveness as "*outcomes under usual or everyday conditions*" in respect of the ISAAC review, as it includes no studies undertaken in such conditions. Neither, as I have discussed elsewhere, is it possible to determine efficacy, due to the severely flawed experimental designs used by the only studies that have been included by the *ad hoc* Committee.

Reference to *"… 'effect size', standard mean difference, percentage of non-overlapping data etc. …*" in EVIDAAC Item 14 (Schlosser, Wendt and Boesch, 2009) make it apparent that quantitative outcome measures are being sought. This appendix will consider the question of whether outcome measures can be said to have been operationally defined in the studies that were included by Probst (2005) and by the *ad hoc* Committee.

14.2 Quantitative Research Questions

In quantitative research, questions must be narrow, specific, measurable, and related to observable phenomena. This is because the research is aimed at explaining or predicting precise relationships between narrowly defined variables. According to the NHMRC:

A well-formulated question generally has three parts:

- *The study factor (e.g. the intervention, diagnostic test, or exposure)*

- *The population (the disease group or a spectrum of the well population); and*
- *The outcomes.* (NHMRC, 2000a, p.13).

As discussed in relation to EVIDAAC Item 1, in the context of FC none of these three factors is adequately defined to support the kind of study that has been reviewed by Probst (2005) or by the *ad hoc* Committee.

According to a checklist of "Criteria for Appraising Reviews" that formed part of the 2009 presentation by the Chairman of the *ad hoc* Committee and colleagues:

> *The degree to which the studies* [included in a review] *share the same dependent and independent variables must be reported* (Schlosser, Wendt and Boesch, 2009).

The *ad hoc* Committee offers no comment on this issue, however, given that the variables have not been operationally defined, it is apparent that they cannot be compared across studies.

14.3 Dependent variables – Outcome measures

"[J]*udgements of effectiveness*" (Schlosser, Wendt and Boesch, 2009) are derived from outcome measures – the "dependent variables" in experimental studies. Quantitative studies are designed to test hypotheses regarding the relationship between "independent variables" and these "dependent variables". Thus, operational definition of "*judgements of effectiveness*" requires clear, operational definition of all variables.

Outcome measures in the studies targeted by the *ad hoc* Committee are stated as "communicative competence", measured as a percentage of "correct responses". Setting aside the question of whether this is a legitimate measure of "communicative competence" for the population (see Appendix 1, Section 1.3), it must be noted that the instruments used and the circumstances under which they are applied vary significantly between studies. For example, subjects may have been asked to name pictures, words, letters or objects, or to respond to auditory stimuli. They may have been asked to spell, or simply to point to objects, pictures etc. They may have been asked for straightforward facts, for responses that required interpretation, or for expression of a preference or opinion. They may have been subjected to standardised testing (for example the Peabody Picture Vocabulary Test) or to a series of questions or stimuli that may or may not have been tested for validity or reliability. (Please see my comments in Appendix B of my submission to the committee.)

By definition, it cannot, therefore, be said that the studies "*share the same dependent … variable*" (Schlosser, Wendt and Boesch, 2009).

14.4 Manipulated independent variable

Independent variables are attributes or characteristics that influence or affect the outcome (or dependent variable) (Creswell, 2002, p621).

The main focus in experimental designs is on the manipulated or treatment variable. The synopsis of Probst (2005) states that, in respect of studies of the validity of FC:

> *The independent variable was often manipulated under 'facilitator non-blind', 'facilitator-blind', 'facilitator double-blind', or 'communication without FC' (non-FC) conditions* (Schlosser and Wendt, 2008, p82).

Manipulation of this variable is intended to isolate the contribution of "facilitator influence" on communicative competence (that is, the dependent variable). However, methods used to manipulate the variables were commonly highly intrusive and likely to alter both the support needs of the subjects and the capacity of facilitators to provide support. For example, to achieve "blind" and "double blind" conditions, facilitators may have been asked to avert their eyes; or to wear glasses adapted to block vision. Alternatively, the facilitator's view of the stimulus material may have been blocked by physical barriers. In some studies, facilitators left the room while the stimulus was being presented, adding a significant delay and other possible confounding factors to the test conditions. When auditory stimuli were presented, some facilitators were asked to wear headphones which played either speech noise or white noise to block sound, or alternatively were delivered auditory stimuli that may or may not match that delivered to the FC user. In one study, facilitator support was replaced by a mechanical device supposed to provide physical support. (Please see Appendix 8, Section 8.3.1 and Appendix B of my 2012 submission to ISAAC).

It is thus not possible to say that "*the studies share the same … independent variables"* (Schlosser, Wendt and Boesch, 2009).

Further, the experimental designs used in the studies under review are inadequate to eliminate the obvious rival explanations for their results: that the control procedures themselves interfered with the phenomenon being studied.

14.5 Secondary independent variables

Extraneous factors are any influences in the selection of participants, procedures, statistics, or design likely to affect the outcome and provide an alternative explanation for the results (Creswell, 2002, p621). Researchers seek to neutralise these extraneous factors – identified as secondary independent variables - through statistical or design procedures. However, as noted in my comments in respect of EVIDAAC Item 1, the characteristics of the population from which subjects may legitimately be selected, and the nature of the supports and accommodations encompassed by the term "Facilitated Communication", are not yet adequately defined. Statistical and design procedures cannot address that which is not understood.

Other factors are also significant:

> [I]*t is widely accepted that many different factors influence the use and effectiveness of AAC. It is also well known that it is important to consider individual factors, communication partners, the social context and the purpose of the interaction as well as the requirements imposed on the person who operates the device. Limitations in the technology and the specific demands related to specific tasks are other important*

aspects related to the usefulness and effectiveness of AAC systems (Hedvall & Ryderman, 2010, p230, referenced in Mazerole and Legotsz, 2012, p82.)

It is difficult to imagine that such a complex phenomenon as person-centred communication support in a dynamic social environment may be reduced to a controlled experiment. Given the variety of secondary independent variables likely to impact differently on each study, it is once again impossible to say that "*the studies share the same … independent variables"* (Schlosser, Wendt and Boesch, 2009).

*14.6 Are the "***judgements of effectiveness***" made in the included studies sound?*

In Probst (2005), mean communicative performance in each study, measured as a percentage of "correct responses", was weighted and corrected for chance "as appropriate". Overall means under facilitator-blind and facilitator-non-blind conditions were generated across included studies. (Schlosser and Wendt, 2008, p82). The synopsis notes:

> *It is unclear to this reviewer why the author did not proceed with the calculation of effect sizes and the subsequent statistical analysis of the differences in mean percent correct of the facilitator-blind vs. facilitator-non-blind conditions. Perhaps the nature of the primary data did not permit this. As such, the conclusions are based more on descriptive data than on inferential analyses* (Schlosser and Wendt, 2008, p83).

As previous discussion has shown, it cannot be said that the studies "*share the same dependent and independent variables …*" (Schlosser, Wendt and Boesch, 2009). It would therefore be highly inappropriate to attempt further statistical analysis by combining the primary data from different studies.

It is equally inappropriate to ignore the flaws in the research and conclude that:

> *While there are some areas where improvements could have been made to this review, the overall difference in mean aggregated data of 27% cannot be ignored and does support the conclusions drawn by the author. Given the consistency of results across so many studies, a very rare occurrence in our field, FC should indeed be classified as an invalid or empirically non-supported treatment. Hence, clinicians and other relevant stakeholders alike have ample reason not to consider FC as a viable treatment choice for individuals with little or no functional speech* (Schlosser and Wendt, 2008, p83).

The dangers of drawing conclusions from such flawed and incomplete data were outlined by the Chairman of the *ad hoc* Committee and colleagues in 2007:

> *In systematic reviews on treatment efficacy, the criteria used to arrive at judgments of effectiveness need to be clearly stated and reported for each study … . Often, reviews that are narrative and unsystematic rely on the conclusions reached by the authors of the original studies, use a vote-count procedure (yes or no) or apply some undisclosed 'effectiveness criterion' to come to a conclusion about efficacy. This is, in part, one reason why narrative reviews often arrive at*

different conclusions— even those that use the same data set … (Schlosser, Wendt and Sigafoos, 2007, P147).

As discussed in relation to EVIDAAC Item 13 and in Appendix 13, a major stumbling block for systematic reviews relates to the quality of individual studies:

> *The purpose of a systematic literature review is to evaluate and interpret all available research evidence relevant to a particular question. In this approach a concerted attempt is made to identify all relevant primary research, a standardised appraisal of study quality is made and the studies of acceptable quality are systematically (and sometimes quantitatively) synthesised. This differs from a traditional review in which previous work is described but not systematically identified, assessed for quality and synthesised* (National Health and Medical Research Council, 2000a, p2.)

Review articles that merely collate conclusions from fundamentally flawed studies of FC are as invalid as the studies they have been based on. To reiterate, reports of individual studies (as opposed to review articles that merely collate "findings") show that experimental studies have overwhelmingly been based on poor understandings of what facilitated communication is (and isn't), resulting in poor practice. Facilitators were poorly trained, "best practice" was ignored, and inappropriate subjects were selected. Experiments featured intrusive "controls" that were highly likely to alter both the support needs of the subjects and the nature of support that could be provided. Please see Appendix B of my 2012 submission to ISAAC, Davies 2012, for further comment.

In the absence of good quality qualitative research, "best practice" provides the most detailed information available regarding the supports encompassed by FC; the characteristics of FC users; and the outcomes that may be expected from use of the strategy. At present, these factors are not adequately understood to provide the precise definitions that would be required to support quantitative studies.

14.7 Direction for future research

Descriptive summaries of research that embrace a variety of study designs do not need to be unsystematic. As recognised by one commentator:

> *… if studies are dissimilar, precluding a meta-analysis, a descriptive summary of the studies in a systematic review should be performed* (Wright et al, 2007, p24).

Such an approach is clearly preferable to an attempt to conjure consistency where there is none:

> *Reviewers often narrow inclusion criteria to deal with heterogeneity by including only those studies reporting a particular outcome, or by limiting the review to specific study designs. The disadvantage of this approach is it biases the review against potentially valuable studies not reporting an outcome in a specific manner* (Wright et al, 2007, p24).

I have argued that the *ad hoc* Committee has taken an inappropriately narrow approach to selection and inclusion of material, resulting in a highly biased review. It is important to heed ASHA's *Facilitated Communication* [Technical Report], which states:

> *Researchers in the experimental and qualitative traditions are encouraged to collaborate in order to design valid, mutually agreed-on procedures for probing validation* (ASHA 1994).

As discussed in Appendix 1, Section 1.1.1, the validity of group-based experimental design is highly questionable in AAC research. While the approach may be appropriate for medical research, where carefully defined treatments can be randomly allocated to large groups of subjects drawn randomly from a well-defined study population, this is not generally the case in therapy or education research and, as discussed in relation to EVIDAAC Item 1, is certainly not the case in AAC research:

> *One cannot generalise results from a group to any specific individual who uses AAC; rather initial trials should be conducted with the individuals to determine whether or not the individual has the necessary sensory and motor skills to use the methods in question. …* [There is] *tremendous heterogeneity across the population of individuals who use AAC as well as within specific individuals at different times of the day and in different situations.* (Dowden and Cook, 2002, quoted in Mazerole and Legosz, 2012, p79-80).

Professor Pat Mirenda has suggested:

> [W]*e need to be bold – not conservative – in formulating research questions and executing studies that push existing boundaries and test hypotheses that may be unconventional but may also lead to new insights and applications.*
>
> *One way to approach this task is to identify people with ASD who have become competent (and independent) communicators through the use of AAC (including FC), and then to work backwards to answer the question "Are there common factors that appear to have contributed to these good outcomes?" If we find any (and I believe we will), we can then design longitudinal hypothesis-driven studies to examine these factors in natural contexts* (Mirenda 2008, p229).

References

American Speech-Language-Hearing Association, 1994, *Facilitated Communication [Technical Report]*, http://www.asha.org/policy/TR1994-00139.htm, accessed on 31.5.12.

American Speech-Language-Hearing Association, 2004, *Roles and responsibilities of speech language pathologists with respect to augmentative and alternative communication: technical report*. At http://www.asha.org/docs/html/TR2004-00262.html.

American Speech-Language-Hearing Association, undated, *Steps in the Process of Evidence-Based Practice. Step 3: Assessing the Evidence*. Accessed 25 April 2014 at http://www.asha.org/members/ebp/assessing.htm .

Australian Institute of Health and Welfare (AIHW), 2003, *ICF Australian User Guide.* Version 1.0. Disability Series. AIHW Cat. No. DIS 33. Canberra: AIHW.

Bebko, JM Perry, A & Bryson, S, 1996, 'Multiple method validation study of facilitated communication: II. Individual differences and subgroups results.' *Journal of Autism and Developmental Disorders*, Vol. 26, pp. 19–42.

Beck, AR & Pirovano, CM, 1996, 'Facilitated communicators' performance on a task of receptive language', *Journal of Autism and Developmental Disorders*, Vol. 26, pp. 497–512.

Bernardi, L. & Tuzzi, A, 2011a, 'Statistical Analysis of Textual Data from Corpora of Written Communication – New Results from an Italian Interdisciplinary Research Program (EASIEST)', in MR Mohammadi (ed), *A comprehensive book on Autism Spectrum Disorder*, InTech, Rijeke, Croatia.

Bernardi, L. & Tuzzi, A, 2011b, 'Analyzing Written Communication in AAC Contexts: A Statistical Perspective', *Augmentative and Alternative Communication*, Vol. 27(3), pp. 183-194.

Beukelman DR. & Mirenda P, 2005, Augmentative and alternative communication: Supporting children and adults with complex communication needs. 3rd edition. Baltimore (MD): Paul H Brookes Publishing Company.

Biklen, D Saha, N & Kliewer, C, 1995, 'How teachers confirm the authorship of facilitated communication: A portfolio approach', *Journal of the Association for People with Severe Handicaps*, Vol. 20, pp. 45–56.

Bomba, C O'Donnell, L Markowitz, C & Holmes, DL, 1996, 'Evaluating the impact of facilitated communication on the communicative competence of fourteen students with autism', *Journal of Autism and Developmental Disorders*, Vol. 26, pp. 43–58.

Cabay, M, 1994, 'A controlled evaluation of facilitated communication with four autistic children', *Journal of Autism and Developmental Disorders*, Vol. 24, pp. 517-527.

Calculator, SN & Hatch, ER, 1995, 'Validation of facilitated communication: A case study and beyond.' *American Journal of Speech-Language Pathology*, Vol. 4, pp. 49–58.

Calculator, S & Singer, K, 1992, 'Letter to the editor: Preliminary validation of facilitated communication', *Topics in Language Disorders,* Vol. 13(1), pp. ix-xvi,

Cardinal, DN Hanson, D & Wakeham, J, 1996, 'Investigation of authorship in facilitated communication', *Mental Retardation*, Vol. 34, pp. 231–242.

Clarkson, G, 1994, 'Creative music therapy and facilitated communication: New ways of reaching students with autism'. *Preventing School Failure*, Vol. 28, pp. 31–33.

Creswell, J, 2002, *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research,* Pearson, Boston [Fourth Edition].

Crews, WD Sanders, EC Hensley, LG Johnson, YM Bonaventura, S Rhodes, RD & Garren, MP, 1995, 'An evaluation of facilitated communication in a group of nonverbal individuals with mental retardation', *Journal of Autism and Developmental Disorders*, Vol. 25, pp.205–213.

Cummins, R. A., & Prior, M. P, 1992, Autism and assisted communication: A response to Biklen. *Harvard Educational Review, 62,* 228–241.

Davies, CA, 2012, *Submission to ISAAC Committee to Develop and Official Position Statement on Facilitated Communication (FC).* Submitted by email to Franklin Smith on 13 June, 2012, and again (with minor corrections for readability) on 4 April, 2013.

Dowden P and Cook A, 2002, *Choosing effective selection techniques for beginning communicators.* In J Reichle, D Beukelman, J Light, Implementing an augmentative communication system: exemplary strategies for beginning communicators. Baltimore, MD: Paul H. Brookes Publishing Co.

Eberlin, M McConnachie, G Ibel, S & Volpe, L, 1993, 'Facilitated communication: A failure to replicate the phenomenon', *Journal of Autism and Developmental Disorders*, Vol. 23, pp. 507-530.

Emerson, A Griffiths, A Prentice, A Cosham, T Howard-Jones, P Grayson, A, 1998, 'Evaluation of Facilitated Communication' *International Journal of Language & Communication Disorders,* Vol. 33, pp. 397-402

Emerson, A Grayson, A & Griffiths, A, 2001, 'Can't or won't? Evidence relating to authorship in facilitated communication', *International Journal of Language & Communication Disorders*, Vol. 36, pp.98–103.

Felce, D, 1994, Facilitated communication: Results from a number of recently published evaluations. *British Journal of Learning Disabilities, 22*(4), 122-126.

Filipek, PA, Accardo, PJ Baranek, GT Cook, EH Dawson, G Gordon, B Gravel, JS, & Volkmar, F, 1999, The screening and diagnosis of autistic spectrum disorders. *Journal of Autism and Developmental Disorders*, *29*(6), 439-484.

Gillam SL & Gillam RB, 2006, Making evidence-based decisions about child language intervention in schools. *Lang Speech Hear Serv Sc*., 37(4), 304-15.

Grayson, A, 1997, In Autism Research Unit (eds), *Living and Learning with Autism: The Individual, the Family and the Professional*, pp. 231-242. Autism Research Unit/National Autistic Society.

Grayson A. Emerson A. Howard-Jones P. & O'Neil L, 2011, Hidden communicative competence: Case study evidence using eye-tracking and video analysis. *Autism*. EPub ahead of print 28 April 2011.

Green, G & Shane, H, 1994, 'Science, reason, and facilitated communication' *Journal of the Association for Persons with Severe Handicaps*, Vol 19(3), pp. 151-172.

Hasson & Joffe, 2007, *The case for Dynamic Assessment in speech and language therapy*. Open access, retrieved from http://clt.sagepub.com/content/23/1/9.full.pdf+html.

Heckler, S, 1994, Facilitated communication: A response by child protection, *Child Abuse and Neglect*, Vol. 18, pp. 495–503.

Hedvall P & Rydeman B, 2010, An activity systemic approach to augmented and alternative communication, *Augmented and Alternative Communication*, December 2010, 26(4) 230-241.

Higginbotham, J., & Bedrosian, J, 1995, Subject selection in AAC research: Decision points. *Augmentative and Alternative Communication*, 11, 11–13.

Iacono, T. & Caithness, T, 2009, Assessment issues. In P. Mirenda & T. Iacono (Eds.), *Autism spectrum disorders and AAC* (pp. 23-48). Baltimore: Paul H. Brookes.

ISAAC, 2012, ISAAC Announces the Formation of a Committee to Develop an Official Position Statement on Facilitated Communication (FC) Woodbridge, Ontario, CANADA – May 11th, 2012. Accessed on 26 April from https://www.isaac-online.org/wp-content/uploads/Position-Statement-on-Facilitated-Communication.pdf

ISAAC, 2013a, *ISAAC Announces Adjustments to the Framework Within Which the FC Committee is Conducting its Efforts*. Woodbridge, Ontario, CANADA – February 8th, 2013. Accessed on 26 April from https://www.isaac-online.org/wp-content/uploads/ISAAC-Facilitated-Communication-FC-Committee.pdf .

ISAAC, 2013b, *Information on New Submissions to be Considered*. Accessed on 26 April from https://www.isaac-online.org/wp-content/uploads/FC-Communication-Revised-23-09-2013.pdf .

ISAAC, 2014, *Report on which the FC Position Statement is based*. Accessed on 28 April 2014 from https://www.isaac-online.org/wp-content/uploads/Report_FC_March18.pdf .

Jacobson, J. W., Mulick, J. A., & Schwartz, A. A, 1995, A history of facilitated communication: Science, pseudoscience, and antiscience. (Science Working Group on facilitated communication). *American Psychologist, 50,* 750-765.

Janzen-Wilde, ML Duchan, JF & Higginbotham, DJ, 1995, 'Successful use of facilitated communication with an oral child', *Journal of Speech and Hearing Research*, Vol. 38, pp. 658–676.

Kezuka, E, 2002, A history of the facilitated communication controversy. *Japanese Journal Of Child And Adolescent Psychiatry*, *43*(3), 312-327.

Kline, RB, 2008, Becoming a Behavioral Science Researcher: A Guide to Producing Research That Matters, Guilford Press, New York.

Kuhn, T, 1996, *The Structure of Scientific Revolutions*, 3rd edn, University of Chicago Press, Chicago and London.

Lof, G, 2011, 'Science-based practice and the speech-language pathologist', *International Journal of Speech-Language Pathology*, Vol. 13(3): pp. 189–196

Mazerolle, P & Legosz, M, 2012, *Facilitated Communication and Augmented and Alternative Communication: A review*. Accessed 8 August 2013 at http://www.communities.qld.gov.au/resources/rti/disclosure-log-dl-16-file01.pdf

Mirenda P & Iacono T, 2009, *Autism Spectrum disorder and AAC*. Paul H Brookes Publishing Co, Baltimore, London Sydney.

Mostert, MP, 2001, Facilitated communication since 1995: A review of published studies. *Journal Of Autism And Developmental Disorders*, *31*(3), 287-313.

Mostert, MP, 2010, Facilitated communication and its legitimacy—twenty-first centurydevelopments. *Exceptionality*, *18*(1), 31-41.

Mostert, P, 2012, Facilitated communication: The empirical imperative to prevent further professional malpractice. *Evidence-Based Communication Assessment and Intervention, 6(1),* 18-27.

Myles, BS Simpson, RL & Smith, SM, 1996, 'Impact of facilitated communication combined with direct instruction on academic performance of individuals with autism' *Focus on Autism and Other Developmental Disabilities*, Vol. 11, pp. 37–44.

National Health and Medical Research Council, 2000a, *How to review the evidence: systematic identification and review of the scientific literature*. Accessed 11 August at http://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/cp65.pdf

National Health and Medical Research Council, 2000b, *How to use the evidence: assessment and application of scientific evidence*. accessed 11 August at http://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/cp69.pdf

National Joint Committee for the Communication Needs of Persons With Severe Disabilities, 2002, *Access to communication services and supports: Concerns regarding the application of restrictive "eligibility" policies* [Technical Report]. Available from www.asha.org/policy or www.asha.org/njc. doi:10.1044/policy.TR2002-00233

National Joint Committee for the Communication Needs of Persons With Severe Disabilities, 2003, *Position statement on access to communication services and supports: Concerns regarding the application of restrictive "eligibility" policies*

[Position Statement]. Available from www.asha.org/policy or www.asha.org/njc. doi:10.1044/policy.PS2003-00227

National Research Council, 2001, *Educating Children with Autism. Committee on Educational Interventions for Children with Autism*. C. Lord., & J.P. McGee. (Eds). Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.

Niemi,J and Karna-Lin, E, 2002, Grammar and lexicon in Facilitated Communication: A linguistic authorship analysis of a Finnish case. *Mental Retardation*, 40, 347-357.

Olney, M, 1995, 'Reading between the lines: A case study on facilitated communication', *Journal of the Association for People with Severe Handicaps*, Vol. 20, pp. 57–65.

Olney, MF, 2001, Evidence of literacy in individuals labelled with mental retardation. *Disability Studies Quarterly, 21(2*), no pagination available.

Oxman, AD., Cook, DJ., & Guyatt, GH, 1994, Users' guide to the medical literature. VI. How to use an overview. Evidence-Based Medicine Working Group. *Journal of American Medical Association*, 272, 1367–1371.

Perini, S., Rollo, D. & Gazzotti, R, 2010, Strategie comunicative nell'interazione conun bambino autistico: Dalla comunicazione facilitata all'intervento comportamentale. [Communicative strategies in the interaction with an autistic child: From facilitated communication to behavioural treatment]. *Psicoterapia Cognitiva e Comportamentale, 16(1),* 103-117.

Prizant, BM, 2011, 'The Use and Misuse of Evidence-Based Pratice: Implications for Persons with ASD', *Autism Spectrum Quarterly*, Fall 2011, pp. 43-49.

Prizant, B Wetherby, A & Rydell, P (2000) 'Communication intervention issues for children with autism spectrum disorders', in AM Wetherby & BM Prizant (eds.), *Autism Spectrum Disorders: A Transactional Developmental Perspective*, Paul H. Brookes, Baltimore, MD.

Probst, P, 2005, Communication unbound--or unfound?--Ein integratives Literatur-Review zur Wirksamkeit der 'Gestützten Kommunikation' ('Facilitated Communication/FC') bei nichtsprechenden autistischen und intelligenzgeminderten Personen. Zeitschrift Für Klinische Psychologie, Psychiatrie Und Psychotherapie, 53(2), 93-128.

Rothchild, I, 2006, Induction, deduction, and the scientific method. In *The society for the study of reproduction*. Accessed 22.12.2012 from http://www.ssr.org/Documents/2006-01-04Induction2.pdf?origin=publication_detail.

Rothstein, H., Sutton, A. J., & Borenstein, M, 2005, *Publication bias in meta-analysis: Prevention, assessment and adjustments*, UK: John Wiley & Sons.

Satterfield, JM. Spring, B. Brownson, RC. Mullen, EJ. Newhouse, RP. Walker, BB and Whitlock, EP, 2009, Towards a Transdisciplinary Model of Evidence-Based Practice. *The Milbank Quarterly*, 87(2), 368-390)

Schiavo, P., Tressoldi, P., & Martinez, E. M, 2005, Autismo e comunicazione facilitata: Prove di verifica dell'autenticità. [Autism and facilitated communication: The results of an authorship test]. *Giornale Italiano delle Disabilita, 5(2)*, 3-17.

Schlosser, RW, 2004, Evidence-Based Practice in AAC. The ASHA Leader, June 22.

Schlosser, RW & Sigafoos, J, 2009, Navigating evidence-based information sources in augmentative and alternative communication. *Augmentative and Alternative Communication*, 25, 225-235.

Schlosser, RW & Sigafoos, J, 2012, An experiential account of facilitated communication. *Evidence-Based Communication Assessment and Intervention, 6(1)*, 1-2.

Schlosser, RW, & Wendt, O, 2008, Facilitated communication is contraindicated as a treatment choice; a meta-analysis is still to be done. *Evidence-based Communication Assessment and Intervention*, 2, 81-83.

Schlosser, RW. Wendt, O. & Boesch, MS, 2009, *Not all reviews are created equal: An overview of appraisal methods*, Texas Tech U HSC. (Powerpoint of a presentation).

Schlosser, RW Wendt, O & Sigafoos, J, 2007, Not all systematic reviews are created equal: Considerations for appraisal. *Evidence-Based Communication Assessment and Intervention*, 1, 138–150.

Shadish, W., Cook, T., and Campbell, D, 2002, *Experimental and Quasi-Experimental Designs for Generilized Causal Inference* Boston: Houghton Mifflin.

Sheehan, CM & Matuozzi, RT, 1996, Investigation of the validity of facilitated communication through disclosure of unknown information *Mental Retardation*, Vol. 34, pp. 94–107.

Sigafoos, J, Arthur, M. & O'Reilly, 2003, *Challenging behavior and developmental disability*. Baltimore: Brookes Publishing.

Simpson, RL & Myles, BS, 1995, 'Facilitated communication and children with disabilities: An enigma in search of a perspective', *Focus on Exceptional Children*, Vol. 27, pp. 1–16.

Smith, MD & Belcher, RG, 1993, 'Brief report: Facilitated communication with adults with autism', *Journal of Autism & Developmental Disorders*, Vol. 23(1), pp. 175-183

Speech Pathology Australia, 2009, *Position Paper: Evidence Based Speech Pathology Practice for Individuals with Autism Spectrum Disorder*. http://www.speechpathologyaustralia.org.au/library/Clinical_Guidelines/ASD_EBP.pdf, accessed 4.9.13.

Speech Pathology Australia, 2010, *Position Paper: Evidence Based Practice in Speech Pathology*. http://www.speechpathologyaustralia.org.au/library/position_statements/EBP_in_SP.pdf accessed 4.9.13.

Speech Pathology Australia, 2012, *Augmentative and Alternative Communication Clinical Guideline*. Melbourne: Speech Pathology Australia.

Tehan, G & Senior, G, 2006, *Final report on Facilitated Communication project*, University of Southern Queensland.

Tuzzi, A, 2009, 'Grammar and lexicon in individuals with autism: A quantitative analysis of a large Italian corpus', *Intellectual and Developmental Disabilities*, Vol. 47(5), pp. 373-385.

Vazquez, CA, 1995, 'Failure to confirm the word-retrieval problem hypothesis in facilitated communication', *Journal of Autism and Developmental Disorders,* Vol. 25, pp. 597–610.

Wehrenfennig, A., & Surian, L, 2008, Autismo e comunicazione facilitata: Una rassegna degli studi sperimentali. *Psicologia Clinica Dello Sviluppo*, *12*(3), 437-464.

Weiss, MJ Wagner, SH & Bauman, ML, 1996, 'A validated case study of facilitated communication', *Mental Retardation*, Vol. 34, pp. 220–230.

Wetherby, A Prizant B & Schuler, A, 2000, 'Understanding the nature of communication and language impairments' in AM Wetherby & BM Prizant (eds.), *Autism Spectrum Disorders: A Transactional Developmental Perspective*, Paul H. Brookes, Baltimore, MD.

White, HD, 1994, Scientific communication and literature retrieval. In H Cooper &LV Hedges (Eds), *The handbook of research synthesis*. New York: Russell Sage Foundation. pp. 41–55.

Wright, RW. Brand, RA. Dunn, W, Spindler, KP, 2007, How to write a systematic review, *Clinical orthopaedics and related research*, 455 (2007): 23-29.