CHAPTER 3

Methodological Issues

3.1 Identifying the Appropriate Conceptual Framework

The focus in this chapter is on methodological issues in research with particular attention to the bilingualism and the childhood language impairment literatures. In the past few decades, the number of studies examining the cognitive control skills (common term: executive functions) of children with different language abilities and proficiencies has shown significant growth in both bodies of literature, largely because most research questions target complex problems that cut across traditional disciplines. Oftentimes, however, the questions of these studies are not conceptually but educationally or clinically driven. Having no or an unclear conceptual framework appears to be related to issues such as the lack of comprehensive theoretical models integrating cognitive control concepts and language processing theories; the mixed use of terminology across disciplinary literatures; and limited familiarity with other disciplines' literature and conceptual frameworks among scholars who work at the intersections of different fields. Moreover, even those who claim to perform interdisciplinary research, frequently just borrow theoretical constructs rather than integrate concepts and theories from two or more disciplines (CohenMiller & Pate, 2019).

A further issue is that the term "conceptual framework" itself is used with at least three different meanings. It may refer to major ideas and arguments of a study; it may be used as a synonym for "theoretical model"; or it may be the guiding principle for the entire research process, from the general statement through the literature review, research questions, methodology, data analysis, and theoretical, as well as practical implications (Ravitch & Riggan, 2017). Thus, depending on one's scientific approach, the same conceptual framework may play different roles in the research process and may be used with different meanings.

To provide perspectives and constructs for multiple disciplines, integrative models are needed. Recent shifts in focus to cultural and linguistic

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 10 Nov 2026 at 05:17:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108992299.004

diversity within both the bilingualism and language impairment literatures has prompted scholars to develop more integrated conceptual frameworks (e.g., Chen & Padilla, 2019; Jiménez-Castellanos & García, 2017; Mistry et al., 2016). A number of these studies included researchers and participants from ethnic and linguistic minority groups, and they added new social constructs to the existing models to integrate such cultural, social-economic, linguistic, and life-experience factors that may help to interrupt educational and social inequality. Traditionally, however, the bilingualism literature focused on children's language background (i.e., level of proficiency, age of acquisition, and frequency of language use), whereas the sociolinguistic literature examined aspects of biculturalism. As of today, when studying cognitive control in bilingual children, demographic factors, such as immigration status, community background, and educational opportunities (e.g., school district), are not typically considered. Thus, specific social constructs are not included in cognitive-linguistic conceptual frameworks, despite having results that link children's cognitive, language, and academic performances to their socioeconomic status (Merz et al., 2019; Sheridan et al., 2017) and cultural background (Grossmann & Varnum, 2011). In future studies, a number of these constructs should be included in conceptual frameworks that are used in the child language impairment literature, especially in studies with bilingual children with developmental language disorder.

3.2 Matching Conceptual Frameworks and Research Designs

Following the selection or development of a comprehensive or integrative conceptual framework, it is important to ensure that it is in line with the study's design. Both the bilingualism and the child language impairment literatures are full of contradictory findings related to methodological issues. One problem is a mismatch between the conceptual framework and the methods. The approaches that researchers select to design their studies reflect the ways they conceptualize the problems they examine. When the methods do not match the conceptual framework, the outcomes are often misinterpreted and explained without any compelling arguments. The most common methodological issues are related to the selection of participants, tasks, and target functions. There is a complex relationship among these variables, and any meaningful research outcome will depend on how they interact within the given research design.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 10 Nov 2026 at 05:17:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108992299.004

3.2.1 Participant Selection

Both language ability and proficiency are continuous variables reflecting a wide range of knowledge and competencies. Despite the nature of these variables, most studies select participants using a group-based approach (e.g., monolingual, bilingual, typical language development, language disorder). This categorization process contributes to the level of heterogeneity of participant groups, which is a critical issue in both the bilingualism and the language impairment literatures. Even when groups are matched on particular variables, such as age, gender, and education level, they differ in many other aspects that might also affect the cognitive-linguistic outcomes. There is, of course, a limit to the number of variables that can be controlled within each design. Furthermore, the use of discrete groups indicates the assumption of representativity, which is typically not met in research with children who are bilingual and/or exhibit developmental language disorder (Kašćelan & De Cat, 2022). The external validity of most studies in both literatures is relatively weak, and scientists, particularly in experimental research, often prioritize internal over external validity. Thus, they increase specificity over generalizability.

Heterogeneity is further boosted by the dynamic language practices and diverse life experiences of children in the twenty-first century. Cultural and linguistic diversity are present in every setting of our lives – in research, education, and policy making, among others. Children's language socialization is strongly affected by the growth of diversity within their various language communities (García et al., 2011). Variations in children's language socialization and other life experiences may affect their cognitive control performance to a great extent. These other experiences may include making music or art, frequent use of computer games, and participation in sports (Chaddock-Heyman et al., 2014; Colzato et al., 2013; Diamond & Lee, 2011; Pallesen et al., 2010). Children may practice these activities on a daily basis but only in varying degrees. Thus, even if researchers attempt to control for overall language skills, music ability, artistic talent, gaming experience, and physical activities, children will still exhibit individual differences in these areas, which would have differential effects on their cognitive control performance. Moreover, a given experience may exert a stronger effect on the performance of one task over another (Valian, 2015).

With an individual-differences approach, where the heterogeneity issue becomes part of the research question, scholars could replace the group-based approach with a more continuous one (Kaushanskaya & Prior, 2015; Marton, 2019). A study with an individual-differences focus may allow

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 10 Nov 2026 at 05:17:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108992299.004

researchers to view participants' levels of language, music, art, gaming, and physical skills as a reflection of their own lived experiences (Garivaldo & Fabiano-Smith, 2023). Only a dynamic and flexible approach could capture the cultural and linguistic richness of current communities. Assessing children's abilities along a continuum would allow researchers and educators to evaluate children's cognitive-linguistic skills with more flexibility and greater awareness of potential growth.

3.2.2 Task Selection

In addition to the heterogeneity of most participant groups, task impurity also contributes to the inconsistent findings in the literature. Most of the cognitive control tasks used in the bilingualism and child language impairment literatures measure multiple functions that tap into different underlying mechanisms. This is reflected in the large number of outcomes, including the inconsistent findings (e.g., low replicability), in minimal or no correlation among tasks claiming to target the same functions and processes, and in global scores that are difficult to interpret (Friedman et al., 2008; Marton, 2019; Valian, 2015).

A demonstrative example for the mixed findings from the bilingualism literature is the debate about the existence of a cognitive advantage in bilingual speakers. Some studies reported superior performance in bilingual individuals on specific cognitive control tasks, such as conflict resolution (Costa et al., 2008) or selective attention (Salvatierra & Rosselli, 2011) compared to monolingual controls, whereas others suggested only a general processing advantage for bilingual people across different cognitive control components (Hilchey & Klein, 2011). In contrast to these findings, a number of studies reported no difference in cognitive control between bilingual and monolingual participants (Von Bastian et al., 2016; Yudes et al., 2011). Mixed results related to task impurity can also be seen in the child language impairment literature, for instance, in studies on visuo-spatial working memory. While some authors reported difficulties in visuo-spatial working memory in children with DLD, others showed no evidence for a group difference between children with DLD and their typically developing peers (Archibald & Gathercole, 2006b; Hick et al., 2005; Marton, 2008). These studies used different visuo-spatial working memory tasks, with some targeting the storage and others the control components more heavily.

A further problem related to task impurity in the literature is that a number of tasks are used interchangeably while showing minimal or no

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 10 Nov 2026 at 05:17:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108992299.004

correlation with one another. Thus, performance on one test may have little or no predictive value for performance on another test (Chan et al., 2008). For instance, measures of resistance to interference often include the Stroop task, the Eriksen flanker task, and the Simon task (Botting et al., 2017; Wang & Gathercole, 2015; Yeung et al., 2020). Stins and colleagues (2005) compared twelve-year-old children's performance on these three tasks and found minimal or no correlation. The speed of processing data across tasks showed some correlations, but the interference effects did not. Similar findings have been reported by Paap and colleagues, who compared the outcomes with the Simon and flanker tasks in college students (Paap & Greenberg, 2013). One reason for the lack of correlation across these tasks may be that they target resistance to interference through the activation of different cognitive control processes and different brain areas. With regards to the Stroop test, there are various suggestions in the literature about its target functions. While Stroop created the task to measure interference (Stroop, 1935, cit. MacLeod, 1991, p. 164), Houben and Wiers (2009) claim that the task measures automatic response inhibition. Mackin and colleagues (2010), however, used the Stroop task to measure speed of processing. Thus, the Stroop effect may be interpreted as an indicator of interference control, automatic response inhibition, or speed of processing. A further reason for the lack of correlation across these interference tasks (i.e., Stroop, Simon, flanker) may be the relatively low reliability of the interference scores.

A third issue related to task impurity is that a number of the widely used cognitive control tasks, particularly the neuropsychological tasks, provide global scores (Marton, 2019; Valian, 2015). In Chapter 2, "Cognitive Control," there is a demonstration of this problem using the Wisconsin Card Sorting Test. Poor performance on that task may be attributed to difficulties in task switching, working memory updating, or inhibition (Marton, 2008). A weakness in any of these functions may result in perseveration. Moreover, many of these tasks have assorted variants that further increase the sources of variability. Even slightly dissimilar tasks may tap into different underlying mechanisms.

Furthermore, these tasks may measure additional processes that are not directly related to the cognitive control system, such as perceptual encoding and response readiness. Hence, these tasks measure complex and broad constructs, and a number of them demonstrate low specificity and/or sensitivity. The Stroop test is considered to be one of the most reliable tasks among the neuropsychological measures used to assess cognitive control. In a meta-analysis based on 33 published articles including the Stroop test with

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 10 Nov 2026 at 05:17:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108992299.004

school-age children with different neurodevelopmental disorders (ADHD, autism, etc.), Homack and Riccio (2004) found medium to good sensitivity but low specificity. Children with different learning disabilities performed more poorly than their typically developing peers, but the Stroop test did not distinguish among the children with different diagnoses. Interestingly, the specificity and sensitivity measures of the Stroop task showed a reversed pattern in adults who were clinically referred for neuropsychological testing (Erdodi et al., 2018). In a sample of 132 participants with psychiatric or neurological disorders, the Stroop test showed high specificity but variable (low to medium) sensitivity with overall low sensitivity for an inverted Stroop effect. Although it is important to note that the children and adult studies used different versions of the Stroop test, they all claimed to target inhibitory functions related to activity in the prefrontal cortex. A similar pattern was found in a study measuring error-related negativity using the flanker task with a large sample (N=457) of adolescent females. Meyer and colleagues (2018) found high specificity but low sensitivity in predicting anxiety disorder among the participants. The findings were in line with outcomes from adults with early frontotemporal dementia, where the flanker task showed fairly high specificity but very low sensitivity (Krueger et al., 2009).

There are plenty of additional examples of issues with specificity and/or sensitivity with these widely used tasks of cognitive control in both the developmental and adult literatures. Thus, when interpreting the mixed outcomes in the bilingualism or in the child language impairment literatures, the specificity and sensitivity values of the cognitive control tasks should always be considered. Unfortunately, only a few studies have examined the psychometric properties of the most widely used neuropsychological tests. Therefore, more research is needed to study the construct validity of these tests, so that we can have a better understanding about the cognitive control components that are measured by the different tasks.

To overcome some of the problems with task impurity, specificity, and sensitivity, we can develop conceptually driven experimental tasks that are well controlled. These tasks may allow us to disentangle cognitive control processes, manipulate and control various components under different conditions, and compare outcomes to baseline measures. Comparisons across conditions, trial and item types may provide more detailed information about the control processes. It is important to note that these manipulations have to correspond to predetermined processes and components based on the study's conceptual framework. Researchers have to ensure a clear match between the targeted cognitive control components and the measures provided by the experimental task.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 10 Nov 2026 at 05:17:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108992299.004

Furthermore, offline tests that result in "end products" (i.e., global scores) do not reveal information about the processing sequence and strategy use of the participants. However, if they are combined with online methods, such as eye-tracking and pupillometry, electrophysiological measures, the results may reveal details about the decision-making process. Online paradigms may provide more information about individual differences in strategy use, error detection, and post-error behaviors, among others. Even with these combinations of offline and online measures, linking experimental designs with solid conceptual frameworks to answer specific hypotheses is critical because variations in stimulus presentation, in the number of trials, in time intervals, or in response modes may all influence the outcomes.

A paramount criticism of experimental tasks concerns the ecological validity of these measures. Although many studies include a section of educational or clinical implications at the end of their discussion, in most cases, they do not clarify the extent to which performance on experimental tasks predicts everyday functioning. In educational and clinical research, it is essential to consider this issue despite its challenges.

3.2.3 Target Function Selection

As discussed in various chapters in this book, cognitive control represents a broad multifaceted construct comprising a wide range of processes and components. There have been attempts in the literature to determine the functional organization of the control processes by identifying higher order functions (e.g., attention) and core functions (e.g., inhibition). Despite the differences in concepts, all models nevertheless operationalize the control system and the underlying neural structures as componential (Baggetta & Alexander, 2016). While the components are separable, they also interact with each other, which makes the assessment of the cognitive control functions challenging. Friedman and Miyake (2017) described this dynamic relationship as the "unity and diversity" of cognitive control components.

When selecting the target functions for a study on cognitive control, it is important to examine the development and functioning level of the individual components. It is equally critical to understand the interactions among the different components to help describe complex human behaviors. Unfortunately, matching tasks with varying demands to specific control processes and components is quite difficult (Badre, 2011; Luna et al., 2015). This fractionation of the components appears to contribute to the low correlations among the cognitive control tasks even within the same construct.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 10 Nov 2026 at 05:17:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108992299.004

Studies using latent variable analyses of cognitive control functions across age groups revealed differences in the rate of component development, with more unity in younger children and more differentiation in older ones (Brydges et al., 2014; Huizinga et al., 2006; Wiebe et al., 2011). Karr and colleagues (2018) performed a systematic review and a reanalysis of an extensive body of research, including data from 46 samples (N = 9,756) across the human lifespan. Their results showed one- or two-factor models in preschool-age children, three-factor models in school-age children and adolescents, three-factor or nested-factor models in adults, and two-factor models in elderly people. The results across studies were fairly consistent, with some exceptions, such as the findings of a four-factor model in adolescents, including working memory, shifting, inhibition, and planning (Laureys et al., 2022). Furthermore, Karr and colleagues noted a certain degree of publication bias based on their systematic review. Many studies replicated previously highly cited models and accepted those as standards. It is not clear how much this practice has contributed to the consistencies in findings across studies.

In addition to age, task complexity also affects the clustering of the different cognitive control components. The most general latent variables (i.e., inhibition/common executive function, updating, shifting) have been observed only with a medium level of task complexity. In more complex tasks, different combinations and separations of the components were apparent (Friedman & Miyake, 2017). Further, when evaluating the cognitive control abilities of participants, individual differences in behavioral performance, brain activation, and connectivity must be considered in addition to the age-related and task-complexity factors.

3.2.3.1 Mismatch between Tasks and Targeted Cognitive Control Functions

Mismatches between tasks and targeted functions are quite frequent in both the bilingualism and the language impairment literatures. There are in-depth discussions of this problem with respect to language abilities and levels of proficiency in Chapter 1, "The Language Continuum" and in Chapter 5, "Associations among Language Ability, Language Proficiency, and Cognitive Control." Therefore, here, we focus on the measures of cognitive control.

Issues related to task impurity and to the complexity of different cognitive control constructs may contribute to the mismatches between both task selection and identification of target functions (see Figure 3.1). Mismatches have been reflected by findings from studies using different tasks (that may not correlate with one another) to test the same cognitive control function (e.g., resisting distractor interference assessed with the Simon, flanker, and matching-to-sample tasks) or by outcomes from studies using the same

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 10 Nov 2026 at 05:17:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108992299.004

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 10 Nov 2026 at 05:17:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108992299.004

	Participant selection	Task selection	Target function selection
Methodological issues	• Group-based approach • Minimal consideration to cultural and linguistic diversity	• Task impurity • Global scores • Interchangeably used tasks show minimal or no correlation with each other • Low sensitivity and/or specificity	• Difficulty with matching tasks to specific control processes and components
Suggestions for future research	• More frequent use of individual-differences approaches	• Develop conceptually driven, well-controlled experimental tasks • Combine off-line tests with online methods	• Determine the targeted mechanisms for each task • Consider individual differences in behavioral performance, brain activation and connectivity, as well as age-related and task-complexity factors

Figure 3.1 Methodological issues and suggestions

task to assess different cognitive functions (e.g., the Wisconsin card sorting task to assess task switching, implicit learning, response inhibition, working memory, and conceptual thinking). Variations in age and task complexity further confound the ability to match between task and target function. For instance, in a developmental study using complex linguistic span tasks to measure verbal working memory, Marton and colleagues (2006) found that younger children (7–8 years old) relied more on storage resources and exhibited difficulty in attention switching and flexible adaptation to changing task requirements compared to older children (10–11 years old) and young adults. The latter two groups relied more on their monitoring and working memory updating skills while performing the same tasks. These differences were also reflected in variations in error types. While younger children produced many interference errors, older participants not only made fewer errors, but their errors were simple omissions rather than interference. Thus, to measure the same cognitive control functions, different tasks should be used with different age groups.

The effect of task complexity was clearly demonstrated in a study on processing speed by Cepeda and colleagues (2013). A large variety of tasks were used to test processing speed across the lifespan, including box completions, digit copying, color naming, and others. Additionally, a series of cognitive control tasks were administered to target participants' working memory, inhibitory control, and task switching abilities. This study revealed that the more complex processing speed tasks also target cognitive control functions, particularly in children and older adults. Thus, a number of tasks used in the literature as pure measures of processing speed assess complex cognitive control skills as well. Therefore, these processing speed tasks should not be used as baseline measures in complex cognitive control tasks, which is the current practice, because of their overlapping constructs. Instead, Friedman and Robbins (2022) suggest assessing cognitive control skills across multiple contexts and statistically extracting the common variables to receive purer measures with no or minimal random measurement errors.

3.3 Data Analysis

The first question we address in this section is how much the selected scoring method matters. Beyond the basic measures of overall accuracy and reaction time, many behavioral studies have involved more detailed scoring to better describe their groups' performance profiles, including the examination of the effect of set size, item type, trial order, the application of Z-scores, and partial credit.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 10 Nov 2026 at 05:17:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108992299.004

Considering the complexity and the popularity of working memory span tasks, we use these tasks to demonstrate how the findings may differ, depending on the scoring method. In an earlier study, Gillam and colleagues (1995) scored each participant's responses with three different methods. With the conservative method, answers were counted as correct only when they were also recalled in the correct serial position (first place, second place, etc.). With the liberal scoring method, credit was given for each correctly recalled item in the proper order, even if they were not recalled in the correct serial position. For example, if the first two items were omitted but the following three items were in the correct order, then the serial positions were all incorrect, but the order of the three words was correct. Therefore, the child received credit for those three items. With the free scoring method, credit was given for each correct item, regardless of its serial position or the order of recall. Although all participants received the highest scores with the liberal scoring method, the presence of the primacy effect (i.e., recalling the first few items with higher accuracy than the following items) was not impacted by the different scoring styles. In contrast, when examining the recency effect (i.e., recalling the last few items with higher accuracy than the prior items), significant interactions with *group* × *task* × *serial* position emerged. Some researchers have incorporated this serial position idea into their working memory model, so they could examine both the content and the binding between the content and its context (e.g., serial or temporal position) when evaluating the outcomes (Cowan, 2005; Oberauer, 2009).

Other modifications to overall scoring were implemented by Leeser et al. (2016), who explored how recall correlated with accuracy, reaction time, and set size using five scoring methods: (1) *recall*, (2) *recall* + *accuracy*, (3) *recall* + *accuracy* + *reaction time*, (4) *set size*, and (5) composite *Z-score*. The five scoring methods were correlated within and across working memory tasks with *set size* showing the weakest correlation, which is an interesting finding, given that set-size manipulations are common in the working memory literature. Set size manipulations are often used to reveal limitations in working memory or attentional capacity (Baddeley et al., 1975; Oberauer et al., 2016; van den Berg et al., 2014). Moreover, when Leeser et al. (2016) compared the working memory outcomes with sentence interpretation, they found interactions across working memory *task type* × *scoring method* × *type of analysis*. Their conclusions included two important points: (1) even if the scoring methods show high correlation, there may still be relevant differences among them, and (2) depending on the manner tasks are scored and how the variables are modelled, completely different results may emerge. Thus, in answer to our first question, we may conclude that scoring methods do matter.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 10 Nov 2026 at 05:17:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108992299.004

In addition to questions about scoring, a number of concerns have been raised regarding manipulations with statistical analyses. Here, we highlight only a few problems from the literature, specifically cognitive biases and issues related to replicability. Two cognitive biases, "dichotomania" and "nullism" (Greenland, 2017), are common issues in the bilingualism and language impairment literatures. Dichotomania refers to the perception of quantities as dichotomous, even when they are not. We have touched upon this problem in our discussion of the measures of language proficiency. When individuals are grouped as monolingual or bilingual speakers, regardless of their level of language proficiency, then a continuous variable is treated as dichotomous. Greenland (2017) demonstrated this issue with regards to the *p* value and its categorization as "significant" or "nonsignificant." Significancy, as a dichotomous indicator is determined by the use of a fixed cut-off point, instead of reporting the real value of the *p*. Another cognitive bias, nullism, is about the assumption of "no difference," even when there is no theoretical or conceptual ground for it. This bias is seen in both frequentist and Bayesian analyses. To avoid this issue, Amrhein and Greenland (2022) suggest that the *p* value should be calculated for multiple values for the effect size and not just for the null hypothesis that is posited with the assumption of no difference between the inspected values.

Replicability is a further problem in the bilingualism and language impairment literatures. According to Baker (2016), more than 70 percent of active researchers have tried and failed to replicate someone else's experiments. The question is whether replicability is really important and whether it can tell us anything about the quality of research. Given the number of issues related to the selection of participants, tasks, and target functions, it seems inevitable that two studies would differ in a number of variables. Baker (2016), however, found some other answers based on a survey. More than 60 percent of the scientists surveyed reported pressure to publish and selective reporting in the literature as the two main reasons for the low replicability rate. With current funding and promotion systems, the pressure of publishing has not vanished, despite reports of scientific misconduct linked to such pressure (e.g., John et al., 2012). Selective reporting or publication bias was identified as one of the factors behind the inconsistent findings on the bilingual cognitive advantage. De Bruin and colleagues (2015) found that the chances for having one's paper published was much higher (68 percent) if the results supported superior performance in the bilingual group compared to less than half that publication rate (29 percent) if the outcomes showed no group difference in cognitive control between monolingual and bilingual speakers.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 10 Nov 2026 at 05:17:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108992299.004

These few examples demonstrate that there are many issues that we have been unable to solve yet. Advancements in methodology and research techniques also come with new challenges. It is important, however, to be aware of these issues and incorporate these limitations into the interpretations of the results.

3.4 Bridging the Gap between Scientific Research and Clinical and Educational Approaches

Even though the majority of papers published in applied research journals devote a section to the educational or clinical implications of their research, or to the description of their study's broader impact on the society, there is a critical need for more systematic translational and clinical research, and for greater collaboration among educators, clinicians, and scientists. Cognitive control is a favored topic, not only in research but in education and clinical practice as well (in the latter two fields, the term "executive function" is used). Multiple hypotheses have been suggested to explain the existence of this gap between research and practice, including the differences between the kind of knowledge that educators and clinicians need during their daily practice and the kind that researchers produce and share. Researchers and educators or clinicians also seem to speak "different languages"; they use different terminology; create and present arguments within different conceptual frameworks; and prioritize different ideas and goals (Garivaldo & Fabiano-Smith, 2023; McIntyre, 2005; Nye, Jr., 2008; Snyder et al., 2015). This, of course, is not an exhaustive list of issues that contribute to the largely independent paths that researchers and educators, as well as clinicians follow.

When examining the ways this gap could be narrowed, one of the first steps should undoubtably be the expression of mutual respect toward each other's expertise and the acknowledgement of the value of complementarity. It is crucial that professionals from different disciplines and subject fields demonstrate interest and willingness to work toward intellectual coherence, scholarly merit, and true commitment to teamwork (Hruby, 2012; McIntyre, 2005). An openness toward real dialogues and the expansion of everyone's own knowledge base may further strengthen collaborations and interprofessional practice. These examples are general steps that provide a foundation for more specific strategies related to individual areas of research and practice, such as the investigation and clinical application of cognitive control. Snyder and colleagues (2015) recommended the utilization of different cognitive control models in clinical research and practice.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 10 Nov 2026 at 05:17:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108992299.004

Their suggestions included the careful selection of specific cognitive control components, the use of multiple tasks for each component, and the extraction of the most sensitive and reliable measures from each task. To further strengthen the partnership and enhance practical relevance, these goals and strategies may be complemented by specific, clinically oriented research questions and approaches, as well as with insider perspectives. A recent exploratory study from the bilingualism literature provides an excellent example of how the quality of research may benefit from the involvement of an insider's perspective (Garivaldo & Fabiano-Smith, 2023). The data from three studies were reanalyzed and reinterpreted by a researcher from within the target community. The combination of a novel theoretical framework and the addition of an insider's perspective resulted in the reinterpretation of the language data that reflected some previously hidden sociolinguistic aspects of the participants' natural language practice.

An in-depth review of the literature shows that the gap between researchers and educators and clinicians is a long-standing one, despite continuous efforts to identify the problems and find solutions. It is a complex problem that involves conceptual, methodological, and ethical issues (Hruby, 2012). Surely in the current digital era, however, we can aspire to find new ways to initiate lively dialogues more easily among professionals from different disciplines.

3.5 Conclusions

In this chapter, we reviewed a few examples from a burgeoning number of methodological problems. Without trying to solve all of these issues, the goal was to demonstrate the complexity of the problems and to call attention to the importance of researchers', educators', and clinicians' awareness of the variety of intricate questions. It is essential that we allow researchers and practitioners to report freely even unexpected and diverse results and findings without having to worry about not getting published or funded because of interpretations that may diverge from the mainstream literature. This is particularly important in research, education, and clinical practice involving disadvantaged communities, groups of speakers of minority languages, and participants from ethnically and culturally diverse backgrounds. Both research and education have to respond more sensitively using individual-difference approaches to the dynamic and distinct practices of the communities in the twenty-first century.

Downloaded from https://www.cambridge.org/core. UNIVERSIDAD AUTONOMA METROPOLITANA, on 10 Nov 2026 at 05:17:34, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781108992299.004