Assessment of the Real and Perceived Effectiveness of Two Educational Models for Students with Learning Disabilities

Philippe Tremblay
Université Laval
philippe.tremblay@fse.ulaval.ca

Abstract

This paper presents a comparative analysis of the real and perceived effectiveness of two education models (inclusion and special education) for students with learning disabilities (LD). The results revealed no significant difference between the two models in terms of target population, objectives, and attributed resources. Significant differences were observed in the effects on students’ performance, as the inclusion model was shown to be globally more effective compared to the special education model. The analysis of the teachers’ perceptions of effectiveness shows that this depends upon three points of view (regular teaching, special education, and student-focused education). The teachers of the two models each believe their own to be effective.

1. Introduction

The evolution of special education models from segregation to a more integrative or inclusive approach is fairly recent in many countries. In the French Community of Belgium, the education of students with learning disabilities (LD) usually takes place in special schools (Communauté française de Belgique [1] [2]; Tremblay [3]), which places the country in the category of segregational special education systems (Euridyce [4]; OCDE [5]). However, since 2007 a few innovative projects based on the inclusive model have emerged. This transition is supported by empirical research that seeks to demonstrate the relative effectiveness of integration/inclusion as opposed to segregation and the effectiveness of specific education models and pedagogical practices. In the process of building student integration, this comparative analysis, in this context, responds to the “fundamental” premise of segregational special education which argues that its specific organisation, in contrast with general education, is more effective in meeting the needs of students with disabilities (Vienneau [6]).

Three significant and well known meta-analyses (Calberg and Kavale [7]; Wang and Baker [8]; Baker [9]) have examined and compared the social integration and academic achievement of integrated students and students in special education. The statistical results emanating from these three meta-analyses all favour integration, reporting that special educational needs (SEN) children integrated into regular classes not only do better academically, but are also more socially integrated than are their special education peers. Size effects vary from .08 to .44 for academic achievement, and on social factors, these fluctuate between .11 and .28, with an average of .195. In a recent literature review, Ruijs & Peetsma [10] showed that in general, the results show neutral to positive effects of inclusive education. The academic achievement of students with and without SEN appears to be comparable to non-inclusive classes or even better in inclusive classes. Lindsay [11], however, advanced that the evidence from the literature does not provide a clear endorsement of the positive effects of inclusion when compared with special education. Lindsay also stated that there is a lack of evidence from appropriate studies and, where evidence does exist, the balance was only marginally positive.

More specifically concerning students with LD, Saint-Laurent, Dionne, Giasson, Royer, Simard and Pierard [12] compared the effect on achievement of a pull-in inclusion program and a pull-out resource class. They observed a positive impact on writing achievement for included students and on reading and math for regular students. The authors also observed that the integrated students did not receive a lesser amount of specialised services compared to the learning disabled students.

Wang [13], who studied the implementation in Pennsylvania of several inclusion models, concurred as to their benefits on learning for both integrated and regular students, as a decrease of 42% was observed in the number of cases referred to special education services. Positive results in terms of behaviour were also obtained by both the integrated and regular students. Moreover, teachers were positively influenced by this implementation. Rea, McLaughlin, and Walter-Thomas [14] compared two integrative models for students with learning disabilities (LD): an integration-inclusion model with co-teaching and an integration-
mainstreaming model with co-intervention (resource class). Compared to the other groups, the outcomes of the included students were superior in first language, mathematics, and science. These authors also examined the social integration of these students, reporting less negative behaviours and greater class attendance.

In meta-analyses focusing on the effectiveness of a co-teaching model in a context of inclusion, Murawski and Swanson [15] concluded that insufficient data prevented a clear determination, as only six out of the 99 studies reviewed met the meta-analysis selection criteria. The latter demonstrated a moderately significant size-effect (+.40). The authors therefore called for further research with experimental groups and control groups with a more defined characterisation of the populations involved to better determine how co-teaching differs from other practices or when no special educational services are provided.

Concerning the perception of teachers on inclusion’s effectiveness in terms of special education, most of the research shows that they have positive perceptions on the effect inclusion has on students with special educational needs (SEN) (York, Vandercook, MacDonald, Heise-Neff & Caughey [16]; Bang & Lamb [17]; Turner & Traxler [18]; Giangreco, Dennis, Cloninger, Edelman & Schatmann [19]). However, Prom [20] and Langdon [21] found that teachers did not believe that the academic and social needs of the majority of students with SEN could be best met in mainstream classrooms. The teachers also believed that the students with SEN stand to gain with inclusion on a social level, but to the detriment of the regular students’ academic and social achievements (Campbell [22]; Salend [23]).

Teachers’ attitudes would be strongly influenced by the nature and the severity of their students’ disabilities in their classes (child-related variables) and less by teacher-related variables. The variables related to the educational environment and to support are incidentally always associated with inclusive attitudes (Avramidis & Norwich [24]). Finally, the evidence appears to indicate that teachers’ negative or neutral attitudes at the beginning of an inclusive education project may change over time as a function of experience and expertise that develops through the processes of implementation (Sebastian & Mathot-Buckner [25]; LeRoy & Simpson [26]; Avramidis, Bayliss & Burden [27]). There is little research, however, dealing with students with learning disabilities (LD) and with the comparative evaluation of representations of the effectiveness of two concurrent models (inclusion and special education).

2. Objectives

Our study consisted of a comparative analysis of two intervention models for students with learning disabilities (LD). The goal was to compare the real effectiveness of two educational models aimed for students with LD, with relation to the perceived effectiveness by the participating teachers. We sought to determine how both the real and perceived effectiveness of these models were congruent. This also involved combining quantitative and qualitative methodologies to find the complementarity of these approaches with the theme. We characterised and compared the populations in the two sub-samples, followed by an analysis of the human resources attributed to each model, and we finally proposed a comparison of the students’ outcomes to external testing and the teachers’ perceptions in terms of the effectiveness of the models over time.

3. Method

Our comparative analysis of two education models consisted of eight primary school inclusion experiments (12 classes) and nine special education experiments (13 classes).

3.1. Treatment and comparison conditions

Established as a research action study in the Belgian French Community over a two-year period, the first model proposed the implementation and perpetuation of a full-time co-teaching program involving a general education teacher and a remedial teacher and focused on the inclusion of students with LD. This project began in Grade 1 in a regular classroom (from 12 to 22 students) in which 3 to 7 of these students were integrated.

The second model consisted of special education classes at the same academic level for students with learning disabilities. Here, a special education teacher instructed a smaller group of students (4 to 12 students/class). Speech therapy and remedial interventions were also organised during the class day, providing each student, individually, with a minimum of one period/week. On one hand, this type of operation involved separating the students with LD within classes and specialised schools, and on the other, pull-out for speech therapy and other sessions).

On average, the students in the inclusion model received 0.25 Full-Time Equivalent – FTE ($SD = 0.071$), while those in special education received 0.23 FTE ($SD = 0.091$). The t-test showed no significant difference on this point between the
two samples ($p = .475$). These resources varied from 0.14 to 0.44 FTE depending on the classes. We also found that on average, the students in special education received more collective, semi-collective, or individual interventions provided by a remedial teacher, a psychomotor therapist, a kinesiologist, etc., during class. Lastly, the specialised classes distinguished themselves by an appreciable use of speech therapy ($M = 1.34$; $SD = 1.03$ compared to $M = 0.68$; $SD = 0.28$ in inclusion). Thus the level of intervention in speech therapy differed significantly ($p = .042$) between the two groups. The included students benefited mostly and exclusively by resources related to the presence in class of a remedial teacher.

Regarding the teachers’ qualifications (15 teachers in inclusion and 13 in special education), each one possessed a degree in primary education except for two remedial teachers in the inclusive setting who did not have the necessary credentials (kindergarten teachers) and who worked together in the same class. The remedial teachers also differed in terms of their complementary specialised training, and were in fact almost half (3 of 7) to have this degree, compared to 3 of 13 for teachers in special education. Moreover, one teacher had an additional Bachelor’s degree in speech therapy and another, a Masters in Education Sciences. In fact, in the inclusion model, five out of seven teachers had an extra job-related degree. In contrast, only one general education teacher had an additional diploma (Masters in Education Sciences). The remedial teachers in the inclusive setting were shown to be less experienced ($M = 4.43$ years; $SD = 3.74$) than those in the special education classes ($M = 9.5$ years; $SD = 2.49$). The regular teachers working in inclusion had, for their part, an average of 7.75 years of experience ($SD = 6.13$). The previous integrative experiment was very limited, as only one teacher had experience in inclusion.

### 3.2. Participants

The experimental group (inclusion) was composed of 12 classes, while the control group (special education) had 13. There were 4.83 learning disabled students/class in inclusion compared to 8.15 in special education. In Grade 1 (8 inclusive classes and 9 special education classes), a total of 228 students participated in the study, including 133 non-disabled students, 37 students with LD in inclusion, and 58 with LD in special education. In Grade 2, there were only four inclusive classes and four special education classes for a total of 125 students: 62 non-disabled students, 21 students with LD in inclusion, and 42 students with LD in special education.

The average I.Q. was 80.95 ($SD = 8.21$) in the inclusion model compared to 80.26 ($SD = 11.24$) in the special education setting. The comparative t-test for independent samples confirmed the absence of significant differences between the two groups ($p = .728$). Tremblay [3] arrived at an average of 82 I.Q. points using the identical methodology ($n = 440$).

We addressed the justifications of placement emanating from the protocol to justify the orientations with 72 keywords (concerns, dropping out, reading, maturity, behaviour, etc.) which were also used by Tremblay [3] in a similar study. These keywords were divided into six justification categories: psycho-pedagogical, pedagogical, psychosocial, psychological, and psycho-affective.

Various statements emerged from this analysis. The two groups showed little difference with regard to these criteria; it was impossible to distinguish one group from another. On the basis of the collected data, it was also impossible for us to articulate the reality of the presence of learning disabilities in the sample students. This term was only rarely referred to (5%); that of “instrumental disabilities”, now obsolete, was more utilized. In fact, we regret the lack of validity and reliability of these diagnoses, much like the quality of the traces present in the protocols justifying the orientation toward special education. Finally, we observed a significant heterogeneity in the justifications evoked. The latter highlighted the possible multi-factorial origins of these disabilities, but also the “catch-all” nature of this type of teaching whereby students not able to follow the general education plan are grouped together.

The analysis of the average age at the beginning of the school year in each sub-sample revealed that this average was higher by approximately one year in special education classes ($M = 7.82$ years; $SD = 307$ days) compared to those in inclusion ($M = 6.88$ years; $SD = 190$ days) in Grade 1. The following year, the difference between the two groups was exactly one year: 6.99 years in inclusion ($SD = 210$) and 7.99 in special education ($SD = 317$ days). The t-test showed significant differences between the two samples ($p = .010$).

In terms of gender, the two samples were composed identically of 2/3 boys. An over-representation of boys was therefore present in the two models. By comparing this proportion with that in the totality of schools for learning disabled students within the Communauté française de Belgique [2], we found this same over-representation (+/- 2/3 boys) to be equally present for this type of special education.

The socio-professional level of the parents (occupation) was divided into two categories: more favourable (management and employee) and less favourable (worker and unemployed). In Grade 1,
82.86% of inclusion students were in the less favourable category compared to 66.66% of students in special education. The $\chi^2$ independence test ($p = .089$) showed no significant difference ($p < 0.05$) between the two groups, although we did observe significant inter-school variations due to the size of the recruitment pool. As for nationality, an average of 7.50% foreign students in Belgium was obtained in inclusion against 10.11% in special education, with no significant difference observed ($p = .356$). Regarding their origin (parents’ place of birth), 42% of the students were of foreign origin in the included model compared to 39% of students in the special education groups, which were very similar and thus showed no significant difference.

Finally, the first language of 17.39% of the students in special education was other than French, compared to 35% of the students in inclusion. The $\chi^2$ independence test revealed that the two groups differed significantly on this point ($p = .027$). However, we found that this factor, as were others, was largely dependent upon the socio-geographical setting of the school where the model was implemented.

### 3.3. Measures

To assess the effectiveness of the two education models, we evaluated the students’ academic achievement. The students were assessed in class by standardized tests (Simonart [28]), by the searchers, in reading/writing and mathematics in October and June of each school year.

To learn more regarding the teacher representations in relation to the model’s effectiveness, we incorporated a question relative to this theme in the interviews held with the teachers at the end of the study. Following the re-transcription of these interviews, we conducted an analysis of this data’s descriptive and systematic content using a methodology from the Grounded Theory (Glaser and Strauss [29]). The data from the re-transcriptions of the semi-directed interviews with the teachers was coded by identifying different items (themes). This coding system helps to categorise similar concepts and enables a theorisation. In carrying out an analysis of structural content, however, in which the member responses of each member of the co-teaching dyad were tested for consistency, we were able to study various concomitances between teachers on certain points. The teacher responses concerning their representations were analysed and a synthesis of their answers is presented in the following section.

### 4. Results

#### 4.1 Real effectiveness

Our analysis of the student outcomes (pre- and post-test) shows that the almost nil average deviation between ($t = .110, p = .913$ in reading/writing; $t = -.67, p = .097$ in math) the two groups in reading/writing went from (0.12) at the beginning of Grade 1 to 3.51 (out of a possible 20) end of year. The t-test showed that these differences were significant ($t = 3.271, p = .002$) in June. In mathematics, the progression went from -1.13 to -0.32 but was not considered significant ($t = -.363, p = .718$).

Our effectiveness analysis was completed with a study of the progress between the beginning and the end of the school year (Wilcoxon test) with all of the students (students with LD in inclusion, students with LD in special education, and non-disabled students of the same academic level). For the included students, a decrease in rank was observed between the beginning and the end of Grade 1, with the exception of the students with the lowest scores in reading/writing, which would suggest that the students with lower outcomes progressed better than the others did and explains the difference in average between the two groups in this subject in Grade 1. However, the Wilcoxon test results were not significant for the two subjects (reading/writing and math) evaluated ($p = .147$ and $p = .326$, respectively). For the special needs Grade 1 students, we noted a significant drop in rank between the beginning and the end of the year in both subjects under study. The difference in the negative rankings was superior to that in the positive rankings and the results were considered significant ($p = .000$ and $p = .000$, respectively).

In the beginning of Grade 2, the special education students fared better in reading and math compared to their included peers. Indeed, the learning disabled students produced an average deviation of 0.67 in reading and -0.82 in math. The t-test showed that in October, the outcomes of the two groups were not significantly different in reading/writing nor in math ($t = -.883, p = .382; t = -.1132, p = .263$, respectively). At the end of Grade 2 (June), the average deviation became positive in favour of the included students, as the difference went from -0.67 to 1.35 in reading/writing and from -0.82 to 1.95 in math. The t-test did reveal, however, that these differences were not significant for these two subjects ($t = 1.802, p = .091; t = 1.726, p = .079$, respectively).

For the students with LD in inclusion, our performance analysis revealed clear gains in both subjects for all of the levels. The Wilcoxon test was considered significant in reading/writing ($p = .000$).
0.035), but not in math (\( p = 0.086 \)). For the students with LD in special education, we again observed a significant decrease in achievement in math (\( p = 0.021 \)), but not in reading/writing (\( p = 0.070 \)).

4.2 Perceived effectiveness

In analysing the interviews with the teachers, it was evident that they attributed three definitions to the models by using different points of reference: effectiveness relative to regular teaching, effectiveness relative to special education, and effectiveness relative to the student.

As a first point of reference, the students with LD must have a level of achievement similar to regular students for the model to be considered as effective. For teachers working in inclusion (regular and remedial teachers), this means “real” success. Students with LD are directly compared with regular students in the inclusive model framework. This type of success involves approximately 50% of the students, according to the teachers.

For teachers in special education, the relation to rhythm and requirements is very different for inclusion teachers. They acknowledge progressing more slowly in regular classes. There is even a fear, in some way, of moving too quickly when teaching. The level of requirement is more malleable. Special education teachers are convinced that they fare better than regular teachers, as teachers in a regular setting do have their challenges in terms of classroom management. Moreover, special education would be effective in terms of aspects other than strictly academic achievement. The teachers stated that it took time to manage other problems, such as those pertaining to behavioural issues. The special education model would fare better in managing these problems as there would be more resources (human, temporal, etc.) to do so. The model’s effectiveness would thus be more global, covering other aspects than just academic achievement (behaviours, self-esteem, etc.). Co-intervention, or the work between the teacher and paramedical staff (e.g., speech pathologist) on the student’s difficulties, is advanced to explain this effectiveness. These two elements (a paramedical case management integrated to school life and a smaller group of students) appear to be fundamental to this special education model.

We observed that the inclusion teachers were unanimous in saying that all students would not have progressed as well in special education. According to these teachers, the inclusion model appears to be more effective than is special education, and that his superior effectiveness is a result of the co-teaching and the regular school context (group, program, requirements, etc.).

Thirdly, the inclusive model is also considered to be effective because it makes the students progress and helps them advance despite the presence, at times, of serious problems; they progress according to their own personal objectives (intervention plan). In this regard, both models have a “universal” effectiveness, with every student. However, in spite of everything, the teachers remain disappointed with some of the results. With some students, this effectiveness is less visible. The Response to Intervention (RTI) is a key criterion for teachers in inclusion, not only in utilising educational practices but also in the orientation toward another model (special education). In special education, the student’s rhythm takes precedence with all of the teachers. By following the child’s rhythm, they seek to push them to the maximum of their possibilities.

5. Discussion

The two groups presented generally similar characteristics on the majority of sub-indicators. Indeed, on such factors as I.Q., social environment, justification for the orientation, gender, nationality, and origin, no significant difference was observed between the two groups. In addition, our test outcomes at the beginning of the year also revealed no palpable difference between the two groups in both reading/writing and mathematics. Only in average age and language spoken in the home was there any distinction. Moreover, the students in these two groups appeared to be representative of the reality observed elsewhere in special education with regard to previous studies on the subject (Tremblay [3]), yet not the case in general education because of numerous over-representations. Albeit a relatively limited sample, it is difficult to extrapolate further; it would thus appear that the composition of the groups was not dependent on the model (justification, gender, age, etc.) but rather on the context in which their schools evolved (language, origin, etc.).

Indeed, on the basis of the available data, we argue that the two sub-samples are considered as being globally comparable. We do point out, however, that due to the lack of reliability of the acquired data, the similarity between the samples remains mostly formal. In light of these data and the selected indicators, it is difficult to predict who should be referred and who should not be; even armed with an evaluation report, the gravity of these disabilities is impossible to measure.

The wide range of justifications evoked as well as the tools in place do not allow for an analysis of this nature, as the lack of validity, reliability, and diagnostic sensitivity – in terms of the analysed protocols – is significant, which presents an obvious methodological challenge in the
interpretation of our findings. These results confirm the considerable vagueness surrounding both the definition and diagnosis of these disabilities (Zuriff [30]; Fletcher [31]; Kavale [32]; Lyon, Fletcher, Shaywitz, Shaywitz, Torgensen, Wood, Schulte & Olsen [33]).

This is supported by some of the teacher representations on a population having “its place” within the model; we found that the perceived effectiveness of the educational treatment holds an important place in the actors’ decisions. Whether in special education or inclusive education, some students, with LD, may thus be maladjusted in terms of the model which may not be suitable for them. This does not in any way signify that the teachers have not tried different approaches with these students, but that the students have not sufficiently evolved to maintain the model’s short term effectiveness acts as a type of evaluation of the validity and reliability of the diagnosis a posteriori. This limited validity and reliability would also even affect the model’s effectiveness.

The impact of these education models on student achievement demonstrates that those in the inclusive setting noticeably progressed in the external evaluations in reading/writing at both the beginning and the end of Grade 1. In Grade 2, a similar tendency was observed, yet no significant differences were recorded between the included students and learning disabled students. In mathematics, no difference was observed the first year, yet in the following year, the average deviation between the two cohorts was quite different, but again not significant.

Further to the performance analysis, we also examined student ranking and found that in Grade 1, the significant progress made by the included students in reading/writing pertained to those with the lowest outcomes. In Grade 2, the gains in reading/writing were both generalised and significant. As for the special education students, their performance decreased significantly between the beginning and the end of their first year, regardless of rank, and in Grade 2, the same tendency was observed, thus the gap between groups tended to grow over time. These results also appear to correlate with those of Rea & al. [14] who reported superior outcomes by students in a pull-in setting over those in a pull-out setting.

In analysing the sessions held with the teachers to assess the model’s effectiveness, we observed that both models were viewed as being effective, which is not paradoxical. While one model may outperform the other, this does not signify that the second model does not produce good results. We found that to assess and compare student achievement, teachers count on different points of reference: regular teaching, special education, and the students themselves. The first point of reference implies that the included students must succeed like the students in regular teaching situations. For those teachers interviewed, these inclusion students, with performances similar to regular students, constitute the “real successes” of the model.

To evaluate the model’s effectiveness, teachers working in special education base themselves equally on regular schooling, but not by comparison with a group of same-aged children, but rather on the child’s past schooling and his evolution. The teachers firmly believe that the students “are better off” in special education rather than in regular classrooms, which in itself constitutes a source of effectiveness. They nonetheless recognise the need to work more slowly than in regular teaching, that they may not cover all of the content, and that this is as much because of the classroom population as it is a lack of time. This time could be used for achieving other objectives than purely “educational” ones. Even though special education and inclusion appear to act as a diagnostic tool (with the treatment defining the pathology), the lower achievement level does not appear to have the same consequences in the two models. A student who makes little progress or regresses in terms of the average at the beginning of the year will be more likely to be reoriented in the inclusion model, while it is easier to maintain students making little progress in the special education model. In summary, more liberties are taken in a regular classroom setting to create a new one. To put it bluntly, the special education model does not acknowledge academic failure.

We observed that the inclusion model teachers had a tendency to compare the effectiveness of their model with that of the special education model. This point of reference is used to highlight the effectiveness and superiority of the inclusion model. Overall, the students could have learned more and learned better in the inclusion model setting. The remedial teachers, who were familiar with the two models, for the most part favoured this option. It goes without saying that special education teachers are not familiar with the inclusion model, and therefore, do not use this point of reference in the reverse.

These three points of reference are not without contradictions, according to the inclusion teachers. They stated with certainty that a student learns better and learns more in an inclusive classroom than in a special education setting, but that they could no longer keep that same student because of his/her slow learning rate, the increasing age-grade gaps, loss of self-esteem, etc. The differences between students within a classroom must be limited, as they will be used to decide upon a student’s reorientation to an inclusive setting. In this concept, the student “deserves” moving up to
the higher class. Based on teachers’ assessments of each class, this norm fluctuates from one situation to another. Some students considered as failures in one class would be successes in the other. This relation to regular teaching therefore leads to reorientations toward special education. More than 25% of inclusion students were reoriented in the inclusive groups under study compared to less than 10% in special education.

The inclusive model is visibly subjected to various pressures (program, parents, rhythm, reading method, etc.) that positively affect its effectiveness, but it also visibly has negative impacts (failures, etc.). Regular education is largely characterised by an educational “format” (Vincent, [34]) that puts constant pressure on both teachers and principals. This places the inclusion teachers in a very uncomfortable position, obliging them to choose between the student’s progression and the sense they give to the model.

6. Conclusion

Two fundamental limitations prevent us from providing a clear opinion as to the superiority of one education model over the other. Despite being representative of certain aspects of the population in this type of special education in the Belgian French Community, the sampling was relatively limited in both groups. As we studied only twelve classes in inclusion and thirteen in special education, it is difficult to argue that these classes are representative of what is being done (or can be done) in each setting. In addition, the choice and nature of certain data taken into consideration (i.e., orientation protocols, interviews, etc.) add nuance to any possible conclusion. A hasty generalisation at this point would only be presumptuous.

In addition, the inclusive, action research approach presented a novel experience that was highly welcomed by the majority of the groups. This type of enthusiastic response is often observed in projects calling for volunteer participation. Perhaps the teachers’ personal involvement was the key to the success of the project. Furthermore, most of the teachers committed to a maximum of two years knowing that the study was of limited duration. It is possible that over time, part of these novel effects would wear off. The presence of the research team may have also been an additional source of motivation. Furthermore, various studies (Sebastian & al. [25]; LeRoy & al. [26]; Avramidis & al. [27]) show that the perceptions were more positive at the end of the school year. In our assessment of the first two years of implementation of this inclusive model, we essentially studied their “start-up” period, as the classes were prototypes of the inclusive model.

Our second focus was to evaluate whether the inclusive model constituted a credible alternative over the other. Conscious of the confines of this study, we conclude with relative confidence that the inclusive model had no negative effect on the participating students, as compared to the other model. In response to the second research question, we also conclude that the choice of the inclusion model as an opportunity for change presents a credible alternative to the special education model by allowing for the emergence of an original and viable configuration of resources for the inclusion of students with learning disabilities.

7. References


