

2013

Assessing Intern Impact Factors for Program Evaluation and Improvement

John Brady

Chapman University, jbrady@chapman.edu

Randy T. Busse

Chapman University, busse@chapman.edu

Jeanne Anne Carriere

Chapman University, carriere@chapman.edu

Michael Hass

Chapman University, mhass@chapman.edu

Kelly S. Kennedy

Chapman University, kkennedy@chapman.edu

Follow this and additional works at: https://digitalcommons.chapman.edu/education_articles



Part of the [Child Psychology Commons](#), [Educational Assessment, Evaluation, and Research Commons](#), [Educational Psychology Commons](#), [School Psychology Commons](#), and the [Social and Philosophical Foundations of Education Commons](#)

Recommended Citation

Brady, J. T., Busse, R. T., Carriere, J. A., Hass, M. & K. S. Kennedy. (2013). Assessing Intern Impact Factors for Program Evaluation and Improvement. *Trainers Forum*, 31(2), 62-72.

This Article is brought to you for free and open access by the Attallah College of Educational Studies at Chapman University Digital Commons. It has been accepted for inclusion in Education Faculty Articles and Research by an authorized administrator of Chapman University Digital Commons. For more information, please contact laughtin@chapman.edu.

Assessing Intern Impact Factors for Program Evaluation and Improvement

Comments

This is a pre-copy-editing, author-produced PDF of an article accepted for publication in *Trainers' Forum*, volume 31, issue 2, in 2013. This article may not exactly replicate the [final published version](#).

Copyright

Trainers of School Psychologists

Assessing Intern Impact Factors for Program Evaluation and Improvement

Abstract

We present a program evaluation of school psychology interns' impact on the academic and behavioral functioning of children. Outcomes from a variety of single-case problem-solving interventions conducted from 2008-2012 indicated overall moderate, positive effects. Global supervisor ratings indicated strong perceptions of the interns' positive impact on the children they served.

Assessing Intern Impact Factors for Program Evaluation and Improvement

With the publication of the recently revised National Association of School Psychologists *Standards for Graduate Preparation of School Psychologists* (NASP, 2010) came a call for training programs to assess the impact our students may have on the behavioral and academic functioning of the children we serve. According to the guidelines in section 3.5, *Practica and Internships in School Psychology*, a training program "...assures attainment of competencies by interns as demonstrated by... Effective School Psychology service delivery evidenced by direct, measurable, positive impact on children, families, schools, and other consumers."

Therefore it appears incumbent upon training programs to include impact factors as part of our program evaluation and program improvement systems. To that end, we present two methods for assessing intern impact. The first assessment involved intervention outcomes for academic and behavioral problems. The second was a global assessment of intern performance as perceived by internship site supervisors.

Assessment Methods

The data for the first assessment were gathered from two problem-solving cases that each intern completed in their final year of our program. The guidelines for the cases were intended to provide interns with a method for demonstrating competencies in best practices for data-based problem-solving and report writing. Each case resulted in a report and summary that included a definition of the problem, baseline data, a description of the intervention, progress monitoring data, and an evaluation of the intervention outcomes. The interns submitted reports on an academic intervention case and a behavioral intervention case.

The second impact assessment, the global rating of intern performance, was accomplished by site supervisors rating their intern's performance on a 4-point scale: 4) Intern

made a dramatic positive difference in the lives of students with who he or she worked; 3) Intern made a significant positive difference in the lives of students with who he or she worked; 2) Intern made some positive difference with the students with who he or she worked; 1) Intern made little or no positive difference with the students with who he or she worked.

The global impact ratings were completed on two areas of performance: a) Using data to facilitate decisions about service delivery and educational placements to improve academic achievement or mental health; b) Developing, implementing, and monitoring individual and/or group interventions to improve academic achievement or mental health. These ratings were gathered at the culmination of the internship experience. Though gross assessments, we deemed site supervisor ratings as potentially important for evaluating the impact our interns had in practice settings.

Sites and Participants

Chapman University is a small, private school located in Orange County, CA. The student population is approximately 5500. Our School Psychology Education Specialist graduate program accepts about 20 students each year. Our internship sites are varied, ranging from urban low SES areas to affluent suburban schools.

The interns for whom data were available for this evaluation were predominately female (90%; N = 73). Their ethnicity/race was 55% European American, 21% Hispanic/Latino American, and 24% Asian American. (We did not have specific data on supervisor demographics, although the majority was female.) For the available case data, the interns provided services to 154 children (110 boys, 44 girls), ranging from PreK to high school. The children's ethnicity/race were varied, with approximately 44% Hispanic/Latino American, 35%

European American, 12% African American, and 9% Asian American. (These percentages were based on available ethnicity/race data from 2010-2012).

Target Behaviors and Interventions

The target behaviors and problem-solving interventions were varied. For the academic cases, the target behaviors included on-task behavior, work completion, reading fluency, early literacy skills, and math skills. The academic interventions included token systems, flashcards, guided practice, and programs such as Read Naturally. For the behavior cases, the target behaviors included school elopement, social skills, aggression, anxiety, and disruptive behaviors. The behavior interventions included token systems, social skills training, behavioral contracts, and counseling.

Intervention Outcome Assessment Methods

We originally began data collection in 2008 when we included formalized problem-solving cases as part of our internship requirements, and as inclusion for our internal program review, and for NASP approval program review. The intervention outcomes for Year One (2008) were based solely on Goal Attainment Scaling (GAS). Outcomes for Year Two were GAS ratings and treatment integrity assessment. In Year Three we expanded the outcome assessment to include Convergent Evidence Scaling (CES). In Years Four and Five, we added the site supervisor global impact ratings.

Goal Attainment Scaling (GAS) is a criterion-referenced rating scale approach that can be readily applied in school settings. Kiresuk and Sherman (1968) developed the GAS method to evaluate the effectiveness of mental health services. When applied in educational contexts, multiple sources (i.e., teachers, aides, parents, consultants, or students) can complete GAS

ratings of academic or social behavior intervention outcomes at individual or group levels (Coffee & Ray-Subramanian, 2009)

The GAS method involves: (a) selecting a target behavior, (b) defining the behavioral or academic outcomes in objective terms, and (c) ranking performance from negative to positive outcomes (Elliott & Busse, 2004; Roach & Elliott, 2005). The outcomes are operationally defined in successive levels of progress on a five-point or six-point rating scale, e.g., -2 to +2, wherein -2 indicates that a problem is much worse, 0 is baseline, and +2 indicates an intervention goal is attained. Raters can provide hourly, daily, or weekly reports of student progress, depending on the problem behavior. These ratings are derived from direct indicators of progress (i.e., direct observations or permanent products) and/or from the raters' perceptions of progress.

Convergent Evidence Scaling (CES) is a method originally developed for single-case school consultation outcomes by Busse, Kratochwill, and Elliott (1995; see also Busse, Elliott, & Kratochwill, 2010). CES is a framework for synthesizing multiple single-case assessment indicators into a common interpretive framework utilizing the logic of goal attainment scaling (GAS). CES requires that levels of progress be operationally defined using a numerical 5 point scale that ranges from -2 to +2. Each outcome indicator is transformed into a CES value, which is then averaged to form an overall CES value to summarize an overall treatment effect. Potential outcome indicators can include GAS ratings, single-case effect sizes, and visual inspection.

Impact Factor Outcomes

Rather than providing a somewhat overwhelming amount of outcome data across the years, we provide a summary of outcomes from 2008 to 2012 (see Table 1), and we present a

table of the 2012 outcomes as an exemplar of our methods. As shown in Table 2, brief contextual information is presented on the child, target behavior, intervention, and outcomes.

- insert Tables 1 and 2 about here -

Across the five years, the overall impact outcomes resulted in moderate, positive effect sizes. For the first two years, as measured by GAS, the overall mean ratings were 1.32 and .96, indicating a moderate, positive effect. For years three through five, the mean CES outcomes were 1.14, .77, and 1.35, indicating a moderate, positive effect. When the data were disaggregated on the moderator variables of intervention type, the overall mean effect for the academic interventions was 1.25 (range 1.00 to 1.47), indicating a positive, moderate effect. For behavior interventions, the overall mean was .91 (range .56 to 1.17), indicating a moderate, positive effect.

Supervisor ratings of intern impact resulted in a mean of 3.72 for item one (impact on using data to facilitate decisions about service delivery and educational placements to improve academic achievement or mental health), and a mean of 3.64 for item two (developing, implementing, and monitoring individual and/or group interventions to improve academic achievement or mental health).

Discussion and Conclusion

Based on the intervention data, it appears our interns had an overall moderate, positive impact on children's academic and behavioral functioning. The overall academic intervention outcomes were stronger than the behavioral intervention outcomes (see Table 2). Supervisor ratings indicated that they perceived their interns as having had a strong, overall positive impact on children. Taken together, the impact factor data indicated that the overall impact of our interns was moderate. This finding is not surprising and may reflect actual school-based

outcomes. In practice, sometimes school psychologists have a strong impact on the children we serve, and sometimes that impact is more measured or non-existent. This conclusion is based on several factors. For example, intervention outcomes are reliant on the social situations and behaviors of multiple people, including the child, teacher, parent/caregiver, and others who are involved in the child's life, such as aides, administrators, and peers. It is difficult to control for personal and social factors beyond our immediate sphere.

One outcome of our approach has been to increase the data-based decision-making and accountability component of our program. Rather than a haphazard approach to identifying impact factors, we have attempted to provide a systematic method to evaluating our interns' relative impact on the children who we serve. A second strength is that, as we have learned from our data collection efforts, treatment integrity is of paramount importance in the work we do. As such, in the last four years of our study, our students have monitored integrity (in one case the intervention was focused on increasing integrity). Another strength is the use of single-case effect sizes for assessing and evaluating intervention outcomes.

Although we have begun to assess impact factors on children, the NASP guidelines call for evaluating interns' impact on families, the school, and the community, none of which we directly assessed. One of our capstone projects is a systems change project, wherein interns are required to engage in changing a system level intervention (e.g., student study teams). This project traditionally has been slated for the final semester of training, and builds upon our consultation and program evaluation course from the previous semester. Our data have shown us that many of our students, although attempting to implement change projects, are not able to follow through with the process, typically because of time constraints. Based on these data, we

have changed our program sequence so that the system change project aligns with our program evaluation project in a year long enterprise. We will be gathering data on the outcomes.

As for family and community impact factors, these may be the most difficult charges. Potential outcomes indices could include GAS ratings of parent/caregiver perceptions of interns' impact on their families' lives. Community impact is a much more difficult variable to assess and, in many ways, may be beyond the capabilities of a typical internship program. We are currently pursuing methods to accomplish each of these training goals and invite any advice from our school psychology training colleagues.

References

- Busse, R.T., Elliott, S.N., & Kratochwill, T.K. (2010). Convergent evidence scaling for multiple assessment indicators: Conceptual issues, application, and technical challenges. *Journal of Applied School Psychology, 26*, 149-161.
- Busse, R.T., Kratochwill, T.R., & Elliott, S.N. (1995). Meta-analysis for single-case consultation outcomes: Applications to research and practice. *Journal of School Psychology, 33*, 269-285.
- Coffee, G., & Ray-Subramanian, C.E. (2009). Goal attainment scaling: A progress-monitoring tool for behavioral interventionists. *School Psychology Forum, 3*, 1-12.
- Elliott, S.N., & Busse, R.T. (2004). Assessment and evaluation of students' behavior and intervention outcomes: The utility of rating scales methods. In R. Rutherford (Ed.), *Handbook of research in emotional and behavioral disorders* (pp. 123-142). New York: Guilford.
- Kiresuk, T.J., & Sherman, R.E. (1968). Goal attainment scaling: A general method for evaluating community mental health programs. *Community Mental Health Journal, 4*, 443-453.
- National Association of School Psychologists (2010). *Standards for graduate preparation of school psychologists*. Bethesda, MD: Author.
- Roach, A.T., & Elliott, S.N. (2005). Goal attainment scaling: An efficient and effective approach to monitoring student progress. *Teaching Exceptional Children, 37*(4), 8-17.

Table 1. Summary of Effect Sizes for the Academic and Behavioral Intervention Outcomes

Year	Academic Effect Size	Behavioral Effect Size	Overall Effect Size
2008	GAS = 1.38; Moderate Effect	GAS = 1.10; Moderate Effect	GAS = 1.32; Moderate Effect
2009	GAS = 1.20; Moderate Effect	GAS = .56; No Effect	GAS = .96; Moderate Effect
2010	CES = 1.19; Moderate Effect	CES = 1.15; Moderate Effect	CES = 1.14; Moderate Effect
2011	CES = 1.00; Moderate Effect	CES = .58; No Effect	CES = .77; Moderate Effect
2012	CES = 1.47; Moderate Effect	CES = 1.17; Moderate Effect	CES = 1.35; Moderate Effect
	Mean Academic Outcome = 1.25; Moderate Effect	Mean Behavioral Outcome = .91; Moderate Effect	Overall CES Outcome = 1; Moderate Effect

Note: Goal Attainment Scaling [GAS] ratings are on a 5-point scale ranging from +2 [goal attained; behavior significantly improved] to -2 [goal not attained; behavior significantly worse]; Convergent Evidence Scaling [CES] is a form of effect size based on combining multiple outcome measures. CES ratings are on a 5-point scale from +2 [strong positive effect] to -2 [strong negative effect].

Table 2. Intervention Impact Factor Outcomes

Child	Target Behavior	Intervention	Outcome
6 th grade boy	Reading fluency	Reading instruction	CES = 1
6 th grade boy	Disruptive behaviors	Token system	CES = 1
6 th grade girl	Reading fluency	Repeated reading	CES = 2
2 nd grade boy	Disruptive behaviors	Behavior contract	CES = 2
9 th grade boy	Site words	Flashcards	CES = 2
5 th grade boy	Disruptive behaviors	Skillstreaming	CES = 0
4 th grade boy	Writing skills	Writing instruction	CES = 0
1 st grade boy	Disruptive behaviors	Individual counseling	CES = 2
1 st grade boy	Letter identification	Phonics/flashcards	CES = 2
8 th grade boy	Inappropriate touching	Social story	CES = 1
1 st grade boy	Phonological skills	Lindamood Bell	CES = 1
3 rd grade boy	Inappropriate verbalizations	Token system/counseling	CES = 2
11 th grade girl	Grammar structure	Writing skills	CES = 1
11 th grade girl	Homework completion	Study skills	CES = 1
10 th grade boy	Work completion	After school program	CES = 2
9 th grade boy	Work completion	Study skills	CES = 2
6 th grade boy	Reading fluency	Reading skills	CES = 2
2 nd grade girl	On-task behavior	Token system	CES = 2
9 th grade girl	Spanish skills	Direct instruction	CES = 2
3 3 rd grade boys/girl	Social skills	Peer mentoring	CES = 2
2 nd grade boy	Reading fluency	Read naturally	CES = 2
2 nd grade boy	Disruptive behaviors	Token system	CES = 2
11 th grade boy	Homework completion	Behavior contract	CES = 2
9 th grade girl	Test anxiety	Counseling	CES = 2
9 th grade boy	Decoding skills	Direct instruction	CES = 1
9 th grade girl	Off-task behavior	Self-monitoring	CES = 0
4 th grade boy	Homework completion	Token system	CES = 1
4 th grade boy	Off-task behavior	Behavioral contract	CES = 1
1 st grade boy	Number recognition	Flashcards	CES = 1
9 th grade boy	Truancy	SFBT counseling	CES = -1
5 th grade girl	Math skills	Fact dash program	CES = 1
			Mean CES Rating = 1.35; Moderate Effect

Note: Convergent Evidence Scaling [CES] is a form of effect size based on combining multiple outcome measures. CES ratings are on a 5-point scale from +2 [strong positive effect] to -2 [strong negative effect].