Psychology Faculty Publications

Psychology

Winter 2006

Case Difficulty of Simulation Software

Marcie N. Desrochers
The College at Brockport, mdesroch@brockport.edu

Darlene E. Crone-Todd
Salem State College

Tim J. Conheady
The College at Brockport

Case Difficulty of Simulation Software

Marcie N. Desrochers, SUNY-Brockport
Darlene E. Crone-Todd, Salem State College
Tim J. Conheady, SUNY-Brockport

Marcie Desrochers and Darlene Crone-Todd are both assistant professors in psychology, and Tim Conheady is a Master's student in the psychology program.

Abstract

Preliminary results concerning difficulty levels of client cases in "Simulations in Developmental Disabilities: SIDD" are presented. Participants conducted assessments to identify causes of problem behavior and propose treatments for 10 clients. Although SIDD may teach clinical decision-making skills, providing numerous cases did not guarantee learning for all participants. Exposure to a difficult case early in instruction was associated with better overall performance by participants. Additionally, treatment performance best indicated perceived difficulty level. Further experimental research comparing order of difficulty is recommended.

Introduction

Clinical training in behavioral psychology can involve teaching students the skills required to design effective interventions for clients with developmental disabilities and challenging behaviors. In particular, functional assessment is recognized as best clinical practice (Fox & Davis, 2005) and required by law for children with challenging behaviors in the classroom (IDEA, 1997). Functional assessment entails various assessments used to identify the possible causes of the problem behavior to design effective treatment. To effectively utilize functional assessment, students must master basic competencies such as: weighing the relative merits of various assessment strategies; interpreting assessment findings; determining the cause of behavior; selecting effective treatments based on the case particulars; interpreting graphed data; and evaluating treatment effectiveness. Moreover, these higher-order thinking skills must be applied across varying client characteristics, types of problem behaviors, various situations, and causes of problem behavior. Developing cases and adjusting teaching strategies to establish these skills can sometimes be difficult or complex.

Computer-based client case simulations can support teaching the decision-making skills involved in these complex situations. For instance, representative client cases can be programmed into the instruction to bridge students' learning between the classroom and live field work (Seabury, 2003), and instruction can be tailored to the learners' current conceptual skill levels (Desrochers & Gentry, 2005). Moreover, analyses of the student decision-making skills can be examined from automatically recorded data to evaluate and improve teaching methodology. Learning to solve complex clinical problems requires exposure to multiple cases of varying difficulty level to promote generalization of the concepts taught (Stokes & Osnes, 1989). Irrelevant case features can be presented along with critical features so that students learn to discriminate the critical case features (Foxx & Faw, 2000).

Several methods of exposing learners to different levels of difficulty of case material exist. One approach is to present easier cases early with more complex cases occurring once student mastery is attained (Martin & Pear, 2007). This procedure may reduce student errors and lead to less frustration, quicker acquisition, and less remediation time. Matching instructional material to students' level of reading ability facilitates generalization to new material better than presenting material that is too difficult (Daly, Bonfigio, Mattson, Persampieri, & Foreman-Yates, 2005). Moreover, Chen, Lee, and Chen (2005) used item response theory in web-based instruction to match course difficulty and student learning ability. Although not experimentally evaluated, students rated the system favorably and their performance improved. Another method to

program easier to more difficult material can involve errorless learning where additional cues are provided and then removed as learning progresses. An errorless learning approach was found to be superior to an errorful method (i.e., no additional assistance) when teaching adults with memory deficits (Page, Wilson, Shiel, Carter, & Norris, 2006) and teaching college students academic skills (Heckler, Fuqua, & Pennypacker, 1975).

Presenting difficult cases early in instruction has also been used, with an assumption that learners benefit from making errors. Errors may be necessary for students to learn how to optimally respond to difficult cases. For example, Dormann and Frese (1994) randomly assigned psychology students to a group taught to avoid errors or to an error training group that was given opportunities to make errors. For average and difficult tasks, the error training group performed better than the error avoidant group. In Kalish, Lewandowsky, and Davies' study (2005) students changed their approach to a problem only when errors and information about an alternative strategy were present. Lastly, a mixed order of case difficulty level is a possible strategy, especially if the effect of errors on learning in the instructional context is unknown. Given that in the applied setting the student will be faced with a variety of situations in no particular order, learning may be enhanced if difficulty is likewise randomly arranged across cases. There is a dearth of research comparing the effectiveness of these three approaches (either the most difficult or least difficult cases first, or mixed difficulty level) for problem-based teaching.

In addition to sequencing cases, identification of the difficulty level of the content is a major consideration in an instructional situation (Crone-Todd, Pear, & Read, 2000). Methods to determine difficulty level of case material include use of a formula based on complexity (e.g., number of variables represented); expert judgments; student ratings; and student performance. Litchfield, Driscoll, and Dempsey (1990) found no significant differences between using a formula compared to expert ratings. Ratings, however, are not always reliable, as when ratings from teachers indicate a more difficult level than ratings from students on the same material (Macaulay & Pantazi, 2006). A direct and individualized method is to examine student correct performance relative to the number of errors during exposure to the problem situation (e.g., Chen et al., 2005). The more errors students make, the more difficult the instructional material.

Purpose of this Study

We examine the relationship between order of client cases, ratings of difficulty of client cases, and performance outcomes for 15 participants who completed computer-based simulation instruction with 10 client cases.

Participant Characteristics and Research Methodology

Participants (N = 15) included ten graduate psychology students in an Applied Behavior Analysis course at a medium-sized, comprehensive New York State college, and five on- and off-site Bachelor-level behavior specialists from a local, nonprofit agency. The average age of the three male and twelve female participants was 26 years. Ten of the fifteen participants had previous experience working with people with developmental disabilities.

A computer-based program (SIDD, or Simulation in Developmental Disabilities, see Desrochers, House, and Seth, 2001) that teaches a functional assessment approach to treatment was used to present participants with 10 computer-based clinical cases. Client characteristics, type and cause of the client's behavior problem, situation, effective treatments, and relevant people involved varied across cases. Participants played the role of a clinician and conducted assessments to determine why the client's problem behavior was occurring (functional hypothesis), and selected the most effective treatment based on that information. Dependent measures include participants' decisions recorded objectively to disk during use of SIDD and subjective evaluations of difficulty level. At the completion of each of the ten client cases, participants rated

25

difficulty level using a five-point Likert scale (1 = Extremely Easy to 5 = Extremely Difficult). Performance measures on functional hypotheses and treatment decisions were analyzed in terms of frequency of first-attempt correct functional hypothesis and treatment selection. Since the difficulty levels of the 10 client cases presented in SIDD were previously unknown, the cases were presented in a mixed fashion in an attempt to identify difficulty and to investigate any difficulty level sequence on learning. A randomized block design was used to determine order of presentation of the 10 client cases. Each participant received two-hour long sessions held across 2-5 days.

Results
Given the relatively small sample size for this study, a decision was made to evaluate statistical significance for all tests using a criterion of alpha = 0.10.

Prior Experience and Performance
An analysis of case level difficulty and participants' field experience reveals that experience was neither associated with participants' correct treatment selection on the first attempt (r_{pb} = -.03, p = ns), nor did it result in different with ratings of difficulty ($t(13)$ = 0.79, p = 0.22). This establishes that participant performance in this study is not confounded by any prior experience in a related area.

Overall Performance Using SIDD
Participants more often identified the correct case-related functional hypothesis for the client's problem behavior (M = 75%, SD = 22) than selecting effective treatments (M = 61%, SD = 18.5, $t(14)$ = 2.28, p<.01).

Difficulty of Client Cases
Performance outcomes and subjective ratings of difficulty level of the client cases indicate that some cases are more difficult than others (see Figure 1). See issue website http://rapidintellect.com/AEQweb/win2006.htm

Student performance indicators. Participants were least accurate in identifying the functional hypotheses for Aaron, Helen, Barbara, and Manuel's problem behavior (67% or less) and most accurate with Danielle and Adam (87%). Participants were least accurate in selecting effective treatments for Helen and Alan (27% or less) and most accurate with Arlis and James (80% or more). In both functional hypothesis and treatment measures, Helen is the most difficult case as measured by student performance.

Subjective ratings. Consistent with the performance measures, participants rated Helen as the most difficult (M = 3.5, SD = 1.3) while Adam was rated the easiest case (M = 2.1, SD = .9). On average, overall client cases were rated between Somewhat easy and Neither easy nor difficult (M = 2.5, SD = .42).

Performance and ratings considered together. While some of the functional hypotheses were difficult, designing an effective treatment is the most difficult aspect of the simulation. For instance, 29% of cases rated as most difficult had functional hypothesis errors compared to an error rate of 92% for treatment selection. Figure 1 shows that a point biserial correlation between subjective difficulty rating and correct selection of functional hypothesis exists for only one case, while for 5 out of 10 cases difficulty ratings correlate with treatment selection accuracy. Thus, designing effective treatment is a more difficult aspect of the simulation than identifying causes. Although average participant case difficulty ratings were not significantly correlated with identification of functional hypothesis (r_s = -.41, p = ns), difficulty ratings were significantly correlated with treatment selections (r_s = -.8, p<.001, 1-tailed) across cases.

Further analyses reveal that errors on the most difficult case were more often due to treatment selection not being tied to the functional hypothesis (11 of 13 participants or 85%). Moreover, many participants (53%) made a treatment selection error with Helen even though they specified the correct functional hypothesis. Upon examination of the characteristics of this case (e.g., type of cause, treatment implications), there were no obvious factors distinguishing it

26

from other cases rated by participants as rated less difficult and for which the correct treatments were selected.

Order of Cases

The identification of correct functional hypotheses is negatively correlated with the order in which Helen appears (rs = -.45, p = .05, 1-tailed). Hence, when students were presented with Helen early on, they were more likely to identify the correct functional hypothesis across the 10 cases. For treatment selections, 5 of 15 participants (42%) performed better on the last five client cases compared to the first five cases. For 4 of these 5 participants, the most difficult case (Helen) was presented during the first half of the cases. For 6 of the 10 participants who performed worse during the second half of the cases, the most difficult case occurred during the second half of the learning sequence. This result held true even when the difficult case was controlled for when comparing the mean correct treatment selections across the remaining 9 cases. More accurate treatment selections were negatively correlated with the order in which Helen occurred (rs = -.407, p = .06, 1-tailed). Participants who were presented with Helen early on were more likely to provide the correct treatment across cases. Hence, these results suggest that the order in which a difficult case is presented is associated with learners' overall selection of the correct functional hypothesis and treatment.

Discussion

The SIDD application may help some students learn how to correctly identify the cause of, and design treatment for, client problem behavior depicted in these simulations. The data also suggest, however, that providing participants with a large number of cases at varying difficulty levels does not guarantee that learners will generalize the concept to new cases. Our results provide preliminary evidence that order of difficulty is related to performance outcomes in problem-based teaching in this study. Sequencing difficult cases early in instruction is associated with better overall functional hypothesis and treatment selection. Although presenting difficult cases first is contrary to general recommendations found in the behavioral literature (e.g., Martin & Pear, 2007), other research findings suggest that this approach, and the accompanying errors, may facilitate learning in the long run (Dormann & Frese, 1994; Kalish et al., 2005).

Arranging difficult cases early in the instructional sequence may be beneficial for a number of reasons. Perhaps when difficult cases are presented at the start of the instructional sequence, learners who struggle with the material and find a solution are better able to confirm the effectiveness of their strategy and find less difficult cases easier to solve. Alternatively, difficult cases at the end of the learning sequence may establish faulty decision-making patterns if no further remediation is provided. Conversely, it may be that functional hypothesis and treatment selections are affected by factors other than order of difficulty. For instance, participants' assumptions about causes of problem behavior, variations in learning style, or differences in interpretation of the instructions may contribute to our findings. Experimental research is needed to more clearly address whether decreasing order of difficulty is important for problem-based learning using this application. A further consideration is how to determine the difficulty level of case simulations. In this study, subjective ratings of difficulty were highly related to designing effective treatment. Both subjective ratings and performance indicated that Helen was the most difficult case.

Future research is surely needed, but it appears that valid measures of "difficulty" require converging evidence of the type presented here. Determining instructional difficulty based on student performance can be readily accomplished during a pre-instruction assessment to allow appropriate individualized computer-based instructional sequencing (e.g., Rittle-Johnson & Keodinger, 2005). Since prior field experience was not correlated with treatment selection accuracy, more training on this important topic is needed. While students may master identification of a correct functional hypothesis, they need to learn how the cause of problem behavior relates to treatment. Although classroom instruction was provided prior to use of the SIDD software, some participants may have lacked a basic understanding of functional assessment. A pre-test would have identified these gaps in understanding.

27

It is critical to develop effective instructional tools to overcome deficiencies in the complex clinical decision-making skills essential to deliver "best practices" for clients with challenging behavior. Examining the relationships between pertinent instructional variables—subjective evaluations, decision-making choices, and case difficulty—is a first step. Future experimental research examining the effect of order of instructional material on the learner's performance and difficulty ratings is also necessary.

References

Chen, C. M., Lee, H. M., & Chen, Y. H. (2005). Personalized e-learning system using item response theory. Computers & Education, 44, 237-255.

Crone-Todd, D. E., Pear, J. J., & Read, C. N. (2000). Operational definitions for higher-order thinking objectives at the post-secondary level. Academic Exchange Quarterly, 4(3), 99-106.

Daly, E. J., Bonfiglio, C. M., Mattson, T., Persampieri, M., & Foreman-Yates, K. (2005). Refining the experimental analysis of academic skills deficits: Part I. An investigation of variables that affect generalized oral reading performance. Journal of Applied Behavior Analysis, 38, 485-497.

Desrochers, M.N., & Gentry, D. (2004). Effective use of computers in instruction. In D.J. Moran & R.W. Malott's (Eds.) Empirically supported education methods: Advances from the behavioral sciences. NY: Elsevier Science/Academic Press.

Desrochers, M. N., House, A. M., & Seth, P. (2001). Supplementing lecture with Simulations in Developmental Disabilities: SIDD software. Teaching of Psychology, 28, 227-230.

Dormann, T., & Frese, M. (1994). Error training: Replication and the function of exploratory behavior. International Journal of Human-Computer Interaction, 6, 365-372.

Fox, J., & Davis, C. (2005). Functional behavior assessment in schools: Current research findings and future directions. Journal of Behavioral Education, 14, 1-4.

Foxx, R. M. & Faw, G. D. (2000). The pursuit of actual problem-solving behavior: An opportunity for behavior analysis. Behavior and Social Issues, 10, 71–81.

Heckler, J. B., Fuqua, R. W., & Pennypacker, H. S. (1975). Errorless differentiation of academic responses by college students. Teaching of Psychology, 2, 103-107.

IDEA '97. (2002 August 26). IDEA '97 Regulations. Retrieved November 7, 2003, from http://www.ed.gov/offices/OSERS/Policy/IDEA/index.html.

Kalish, M. L., Lewandowsky, S., & Davies, M. (2005). Error-driven knowledge restructuring in categorization. Journal of Experimental Psychology: Learning, Memory, and Cognition, 31, 846-861.

Litchfield, B. C., Driscoll, M. P., & Dempsey, J. V. (1990). Presentation sequence and example difficulty: Their effect on concept and rule learning in computer-based instruction. Journal of Computer-Based Instruction, 17, 35-40.

Macaulay, M., & Pantazi, I. (2006). Material difficulty and the effectiveness of multimedia in learning. International Journal of Instructional Media, 33, 187-195.

Martin, G., & Pear, J. (2007). Behavior modification: What it is and how to do it (8th Ed.). Upper Saddle River, NJ: Prentice Hall.

Page, M., Wilson, B. A., Shiel, A., Carter, G., & Norris, D. (2006). What is the locus of the errorless-learning advantage? Neuropsychologia, 44, 90-100.

Rittle-Johnson, B. & Koedinger, K. R. (2005). Designing better learning environments: Knowledge scaffolding supports mathematical problem solving. Cognition and Instruction. 23(3), 313-349

Seabury, B. (2003). On-line, computer-based, interactive simulations: Bridging classroom and field. Journal of Technology in Human Services, 22, 29-48.

Stokes, T. F., & Osnes, P. G. (1989). An operant pursuit of generalization. Behavior Therapy, 20, 337-355

28