

Child and Adolescent Telepsychiatry: Reliability Studies Needed

NORMAN ALESSI, M.D.

ABSTRACT

The progress of child and adolescent psychiatry has been greatly influenced by the development and implementation of diagnostic reliability studies. The key components of these studies have been the use of standardized structured interviews, the use of trained interviewers, and reliability designs that confirm the level of interrater reliability for a diagnosis or a number of diagnoses in a specific clinical environment. Despite the impact of the methodologies associated with these studies and their acceptance as standards, there are no comparable studies in child and adolescent telepsychiatry that use similar research methodologies or technologies. Most of the child and adolescent telepsychiatry literature is represented by program descriptions and patient/practitioner satisfaction surveys evaluating the acceptance of the technology and the care delivered via telepsychiatry. The use of standardized measures and methodologies constitute the essential components of the science of child and adolescent psychiatry. Their use validates studies for practitioners as acceptable and allows the field to grow. Their absence undermines the credibility of any study and decreases its acceptance. Without science to substantiate the enthusiasm often expressed by those implementing systems of telepsychiatry, little progress will occur. Reviewed are structured interviews used in child and adolescent psychiatry research, a critique of current designs, and potential considerations for the development of studies in child and adolescent telepsychiatry research.

INTRODUCTION

TELEPSYCHIATRY IS OFTEN NOTED to be the first medical subspecialty to use telemedicine technology. From publications originating from the 1950s and extending through the 1970s, psychiatry was a pioneer in advancing the application of telecommunication in medicine.¹⁻⁹ Telepsychiatry dealt almost exclusively with adults and rarely with children and adolescents.^{8,10}

Recently, there have been a number of child and adolescent telepsychiatric reports appearing in the literature.¹¹⁻¹⁷ Ermer¹² reports on 100 children and adolescents (aged 3-17) seen both for diagnostic evaluations and treatment via teleconferencing technology. A number had

histories of previous psychiatric hospitalizations; the principle diagnosis was attention deficit hyperactivity disorder. Little information is given concerning co-morbidity or the therapeutic interventions used. The clinical assessments did not utilize structured interview methods for assessment or diagnosis.

Rendon¹³ gives an in-depth report of providing a 16-week course of cognitive behavioral therapy to a 10-year-old diagnosed with an oppositional defiant disorder. These sessions demonstrate the potential of this technology beyond diagnostic consultations alone. Gelber¹⁵ and Blackmon¹⁶ both report on patient satisfaction among those families that received their care via teleconferencing. Hilty¹⁷ describes a one time clinical consultation to a distant

community via teleconferencing with improvement occurring. There was no standardization of the diagnostic methodology. Given that these were either clinical case reports or case series, none of these studies attempted to use a systematic methodology in the evaluation of their populations.

Among recent publications about child and adolescent telepsychiatry, there is one exception.¹⁸ In this article, the authors attempted to do what few others had attempted: a randomized, controlled trial of telepsychiatric assessments of children and adolescents. The authors saw 26 cases of children and adolescents (aged 4–16 years), first face-to-face (FTF), then in a teleconference interview. They concluded via an independent evaluator that teleconferencing was an alternative to FTF interviews. As a clinical study it shows the potential of the technology. Yet, this study had several flaws that decreased its value as a contribution to the scientific literature. First, a standardized diagnostic instrument was not used; second, the raters were not trained and rater reliability as compared to an objective standard was not documented; and, third, an independent rater, rather than quantitative analysis was used to assess the interrater reliability of the diagnoses. These are essential components of a research protocol in child and adolescent psychiatry.

A review of the history of child and adolescent psychiatry demonstrates that well orchestrated studies using standardized interviews and diagnostic criteria, trained raters, and methods that quantitatively determine interreliability of diagnosis have changed the course of the specialty. Most notably have been the studies of Rutter,¹⁹ Herjanic,²⁰ Puig-Antich²¹ and Strober.²² Rutter's¹⁹ Isle of Wight was the first epidemiological study that quantified the frequency of psychopathology in children and adolescents. It revolutionized child and adolescent mental health by using structured and semistructured interviews administered by trained raters to elicit phenomenological criteria in the identification of child and adolescent mental disorders. Prior to this time childhood disorders were thought to be psychodynamic, both, in origin and in conceptualization. There-

fore, a questionnaire-based technique was thought to be of little value in the study of these disorders and most of the literature was clinical case reports.

Herjanic²⁰ set the stage in the use of child and adolescent diagnostic interviews by studying the reliability of self-reporting of children with that of parent reports. Concordance of greater than 75% between parent and child was found concerning factual information and symptoms and descriptions of behavior.

Puig-Antich²¹ continued to move the field forward by developing a structured instrument, the Schedule for Affective Disorders and Schizophrenia for School-Age Children: Kiddie SADS (K-SADS), based on the adult phenomenological structured instrument, the Schedule for Affective Disorders and Schizophrenia (SADS). His childhood mood disorder studies were significant for their use of the adult derived criteria and their methodological rigor.

Similarly, Strober²² demonstrated that adolescents could be interviewed and that the information was highly reliable among raters who independently interviewed and assessed the adolescents. Prior to this study, few thought that major psychopathology could be identified among adolescents due to "adolescent turmoil," especially bipolar disorders and depression.²³

These studies form the scientific foundation of child and adolescent psychiatry. Given the literature to date in child and adolescent telepsychiatry, it appears they have not made their way into these studies. Research in child and adolescent telepsychiatry is in its infancy. Yet, as indicated by the Elford *et al.*¹⁸ report, with its use of a scientific structure, telepsychiatry is taking major steps in achieving credibility. These steps are necessary if the application of teleconferencing is to have a substantial impact within the specialty. One can argue that without them there will be little to no progress. The following will provide an overview of the types of structured and semistructured interview instruments available and the methodological considerations to make when designing a research study concerning child and adolescent telepsychiatry.

ESSENTIAL CONSIDERATIONS FOR THE SCIENCE OF CHILD AND ADOLESCENT TELEPSYCHIATRY

There are a number of advantages that standardized techniques have offered all areas of science, as well as child and adolescent psychiatry. These include: objectivity, quantification, communication, economy, and most importantly scientific generalization.²⁴ All science is based on the idea that any statement of fact made by one scientist can be independently verifiable by another. This requires the objective measurement of all variables needed to substantiate that statement. Quantification allows the reporting of results in finer detail than through personal judgments and it permits the application of mathematical methods of analysis. Communication is facilitated when scientific measures are used. Scientific measures allow open, critical discussion to occur between those working on the same problem and a foundation of knowledge to be built. Although the establishment and implementations of these measures is costly, they ultimately save time and money as compared to subjective evaluations. Personal judgments often cover the same ground and allow little understanding to develop. Scientific generalization is at the heart of scientific work. To elucidate underlying order among events is the ultimate goal. Without measures, there can be little hope that this elucidation will occur.

Structured and semi-structured interviews within child and adolescent psychiatry

Major progress has been made in the study of children and adolescents with mental disorders over the last 30 years. This reflects the progress made in large part because of the development of criteria-based research diagnostic categories that validated the presence of major psychiatric diagnoses, and the development of structured interview instruments that allow the systematic collection of information needed to make diagnoses. Using the Feighner Criteria, then the Research Diagnostic Criteria, adults with major psychiatric disorders were first studied.^{25,26} Subsequently, modified criteria were applied to children and adolescents.²⁷

This had the effect of revolutionizing the psychiatric research of child and adolescent research, and the clinical delivery of care. Research based on these techniques has resulted in the development of epidemiological, treatment, and genetic studies of child and adolescent disorders. This research has benefited all children and adolescents and has helped in the articulation of the needs of this group, as occurred in the recent Surgeon General's Report on Mental Disorders in America.²⁸

These structured interviews offer the following over traditional clinical assessment methods: Reliability, level of measurement, reduction of information and criterion variance, and comprehensiveness.^{27,29} These allow for a stabilization of the assessment beyond the idiosyncratic nature of most clinical evaluations. And, it allows for the establishment of interrater reliability. The level of measurement can be extended beyond the nominal terms of present or absent to the characterization of the degree of severity and impairment. Standard instruments reduce the information variance and the criterion variance found when clinical assessments are used. This allows for greater certainty of the information recorded. Traditional interviews focus on the presenting problems and associated symptoms. This approach may lead to significant symptoms and diagnoses being missed, especially if the diagnosis is not a major diagnosis. Another critical benefit of structured interviews is that a number of comorbid conditions can be diagnosed for each patient, as well as lifetime conditions versus present episode.

A number of diagnostic interviews have been developed for the assessment of children and adolescents with mental and emotional disorders.^{27,30-32} These are characterized as either structured or semi-structured, based on the manner in which and by whom they are administered. Often, trained lay raters administer the structured interviews for epidemiological studies. This requires that each question be asked in a very specific verbatim manner. Structured interviews include the Diagnostic Interview for Children (DISC), Diagnostic Interview for Children and Adolescents (DICA), the Interview Schedule for Children (ISC), and

the Child and Adolescent Psychiatric Assessment (CAPA).³³⁻³⁸ These have the benefit of having a larger cohort of raters, who usually cost less to administer the interview. Semi-structured interviews allow the rater more flexibility in the choice of questions and the manner in which the questions are answered. Trained clinicians most often administer these interviews, allowing them to use their clinical judgment in deciding when and to what extent to pursue diagnostic areas. Semi-structured interviews included Schedule for Affective Disorders and Schizophrenia for School-Age Children (K-SADS), Children's Interview for Psychiatric Syndromes (ChIPS), and the Child Assessment Schedule (CAS).^{31,39,40,41} Because raters must have clinical training, their use costs more, yet, given their clinical skills, they can make decisions concerning the interview and the disorders that lay raters would not make. This may lead to a shortening of the time of the interview and increase the number of comorbid conditions identified.

A number of these structured and semi-structured interviews have the virtue of all being keyed to the DSM-IV, many have gone through numerous revisions and have been used in a number of studies in which diagnostic interviews are required. The down side of these instruments are that they can sometimes take from 2 to 4 hours to complete and they often require that information be collected from multiple sites including the parents, school, and other clinical settings before final scoring.⁴² Nevertheless, the time taken to administer the interview is outweighed by the quality of information obtained and the fact that the utilization of the instrument allows the duplication of the study design.

Regardless of the instrument chosen, the next step will be to make certain that those using the instrument receive training to guarantee interrater reliability. These can be performed either through training courses offered by those that developed and support the instrument or videotaped training.⁴³

To date there have been no child and adolescent telepsychiatry studies that have used a structured or semi structured interview. Most have used data derived from clinical interviews and assessments, which may provide useful in-

formation, especially of the presenting diagnoses; however, such data provides little other systematic information. The major challenge for one approaching the choice of instrument will be one of time and money, a decision that is usually influenced by the availability of external funds and the nature of the question being asked.

Comparison of designs to evaluate reliability of telepsychiatry for children and adolescents

Once an interview is chosen, the next step is to determine the best design to use for its administration. In the case where reliability is the prime issue, it is important to determine the evaluation situation that would allow for the interview to occur. A classic study design to determine the reliability of psychiatric diagnoses is to use two raters with a patient, one an interviewer and the other an observer.^{22,25,44} This design allows interrater determinations to be made for principle diagnoses, co-morbid conditions, and a symptom-by-symptom analysis. Of course, in specialties such as pathology and radiology, reliability studies have been performed using either pathological specimens or radiological images. Such studies are relatively straightforward, requiring a fixed image to be assessed by multiple raters and then interrater reliability to be made. Isn't psychiatry the same? Couldn't an asynchronous recording of an interview be made, multiple raters rate the recording, then interrater reliability determined? It could and it should during the training of the raters, but it will not suffice for the determination of the reliability of the interview itself. Why? The assessment of children and adolescents requires that we not only assess what is derived from the interview, in this case an asynchronous recording, but it be done by raters both present, face to face, and telepresent via teleconferencing.

Several different methods have been used to deal with the problem of proving reliability. The following are a review of these studies and a critique of their methodology. The lack of structured interviews and their necessity will not be reiterated in the assessment of the studies because their use and necessity have been described.

1. First, a patient is evaluated in a face-to-face evaluation, then independently via a teleconference evaluation. The order of assessments could be randomized and counter-balanced. Then, an analysis could be completed to determine interrater reliability.

Elford *et al.*¹⁸ approached their assessment of children and adolescents with this strategy. This design assumes that repeat clinical evaluations will not influence each other. Does a test–retest method influence the subject being evaluated? If so, in what manner? Can it be controlled? The issue of repeat examinations is not dealt with in the study. Preferably, only one structured interview would occur and would form the basis of the data to be analyzed.

2. Second, evaluate a group of patients with two interviewers, face-to-face, and evaluate another group with one interviewer, face-to-face and another remotely via teleconferencing. Compare the interrater reliability or kappa value of each interviewer pair.

Rusken *et al.*⁴⁵ completed such a study on 30 adult psychiatric patients. The interrater reliability calculated for the most common diagnoses, major depression, bipolar disorder, panic disorder, and alcohol dependence showed that the interrater reliability was identical or almost identical for the patients who had two in-person interviews or those who had an in-person and a remote interview. The study had a number of positive attributes. The interviewers were trained, a structured interview was utilized, and the interrater reliability of major diagnoses was used for comparison. Nevertheless, a significant shortcoming is that two different clinical groups were used in the comparison rather than the same population. Could there have been differences in the two groups of patients? Despite this shortcoming, this is a valuable contribution to the telepsychiatry literature. In fact, this is the only true reliability study in the telepsychiatry literature. A similar study does not exist concerning child and adolescent telepsychiatry.

3. Third, two raters are present with the patient face-to-face (FTF₁ and FTF₂), and a

third rater is present via teleconferencing in real time. Two types of structured interviews are performed: (a) by one of the raters present, face-to-face with the other raters observing, and (b) via the teleconferencing rater (TC), with the two face-to-face raters observing.

Patients would be randomly assigned to either interview. The principle interviewer would determine the need for additional information (including interviews with parents). Interrater reliability would then be determined between the three raters, FTF₁ versus FTF₂, FTF₁ versus TC and FTF₂ versus TC in both test situations. This design has the virtue of not requiring the patients be interviewed twice; the interview only involves one group of patients, as opposed to two groups of patients such as that seen in the Ruskin study to assess interrater reliability among examiners, and last it will allow an assessment of the influence of FTF versus TC interview on the interrater reliability. To our knowledge, such a study has not occurred either involving children and adolescents or adults.

SUMMARY

The engine of change for child and adolescent psychiatry has been the development of standardized methodologies for the assessment and characterization of mental and emotional disorders. Significant amongst these methodologies has been the development of structured interviews and the methods, which allowed for the validation of diagnosis and interrater reliability. Fueled with these methods, the quality and quantity of studies involving an enormous range of diagnostic entities has literally exploded in the last 20 years.

In comparison, the area of child and adolescent telepsychiatry has shown little progress. Publications have either been case reports or case series, rarely have controlled studies been conducted. These publications have not taken advantage of the progress that has been made in child and adolescent psychiatry. There are too few scientific studies in child and adolescent telepsychiatry. For child and adolescent

telepsychiatry to mirror the progress of child and adolescent psychiatry, it must adopt these scientific standards. To do so will require the adoption of these techniques by all those who are committed to growth in this area. This will require training and collaboration with established researchers in the field of child and adolescent psychiatry. Further, the implementation of the standards described has been costly and often reflects the work of entire careers. Funding of a substantial nature, usually from the National Institute of Mental Health has been necessary to develop, test, implement, and refine the validity and reliability of these studies.

The need for child and adolescent mental health services nationwide is overwhelming. One way that needed expertise can be more widely distributed will be via telepsychiatry. Before this occurs these standards must be adopted and become a part of every study that involves child and adolescent telepsychiatry. One could speculate that such an adoption will result in the same rapid and prolific growth in telepsychiatry as has occurred in general child and adolescent psychiatry. Without it, the area will flounder.

REFERENCES

1. Wittson, C.L., and Benschoter, R.A. (1972). Two-way television: Helping medical center reach-out. *American Journal of Psychiatry*, 129, 624-627.
2. Wittson, C.L., Afflect, D.C., and Johnson, V. (1961). Two-way television in group therapy. *Mental Hospitals*, 12, 22-23.
3. Wittson, C.L., and Dutton, R. (1956). A new tool in psychiatric education. *Mental Hospitals*, 7, 11-14.
4. Solow, C., Weiss, R.J., Bergen, B.J., and Sanborn, C.J. (1971). 24 Hour Psychiatric Consultation via TV. *American Journal of Psychiatry*, 127, 120-123.
5. Menolaxcino, F.J., and Osborne, R.G. (1970). Psychiatric television consultation for the mentally retarded. *American Journal of Psychiatry*, 127, 515-520.
6. MacKinnon, R.A., and Michels, R. (1970). The role of the telephone in the psychiatric interview. *Psychiatry*, 33, 82-93.
7. Benschoter, R.A., Wittson, C.L., and Ingham, C.G. (1965). Teaching and consultation by television. *Hospital and Community Psychiatry*, 16, 99-100.
8. Thomas, F., and Dwyer, T.F. (1973). Telepsychiatry: Psychiatric consultation by interactive television. *American Journal of Psychiatry*, 130(8), 865-869.
9. Fisch, A., and Dwyer, T.F. (1972). Interactive television in the continuing education of foreign-trained psychiatrists. *Journal of Medical Education*, 47(11), 912-914.
10. Straker, N., Mostyn, P., and Marshall, C. (1976). The use of two-way interactive television in bringing mental health services to the inner city. *American Journal of Psychiatry*, 133, 1202-1205.
11. Ermer, D.J. (1999). Experience with a rural telepsychiatry clinic for children and adolescents. *Psychiatric Services*, 50(2), 260-261.
12. Ermer, D. (1999). Child and adolescent telepsychiatry clinics. *Psychiatric Annals*, 29(7), 409-414.
13. Rendon, M. (1998). Telepsychiatric treatment of a schoolchild. *Journal of Telemedicine & Telecare*, 4(3), 179-182.
14. Gelber, H. (1998). The experience of the Royal Children's Hospital Mental Health Service videoconferencing project. *Journal of Telemedicine & Telecare*, 4(Suppl. 1), 71-73.
15. Gelber, H., and Alexander, M. (1999). An evaluation of an Australian videoconferencing project for child and adolescent telepsychiatry. *Journal of Telemedicine & Telecare*, 5(Suppl. 1), S21-S23.
16. Blackmon, L.A., Kaak, H.O., and Ranseen, J. (1997). Consumer satisfaction with telemedicine child psychiatry consultation in rural Kentucky. *Psychiatric Services*, 48(11), 1464-1466.
17. Hilty, D.M., et al. (2000). Telepsychiatric consultation for ADHD in the primary care setting [letter]. *Journal of the American Academy of Child & Adolescent Psychiatry*, 39(1), 15-16.
18. Elford, R., et al. (2000). A randomized, controlled trial of child psychiatric assessments conducted using videoconferencing. *Journal of Telemedicine and Telecare*, 6, 73-82.
19. Rutter, M., and Graham, P. (1968). The reliability and validity of the psychiatric assessment of the child: I. Interview with the child. *British Journal of Psychiatry*, 114, 563-579.
20. Herjanic, B., et al. (1975). Are children reliable reporters? *Journal of Abnormal and Child Psychology*, 3, 41-48.
21. Puig-Antich, J., et al. (1978). Prepubertal major depressive disorders: A pilot study. *Journal of the American Academy of Child Psychiatry*, 17, 695-707.
22. Strober, M., Green, J., and Carlson, G. (1981). Reliability of psychiatric diagnosis in hospitalized adolescents. *Archives of General Psychiatry*, 38, 141-145.
23. Freud, A. (1958). Adolescence. *Psychoanalytic Study of the Child*, 13, 255-273.
24. Nunnally, J.C. (1978). *Psychometric theory*. New York: McGraw-Hill.
25. Robert, L., Spitzer, M.D., Endicott, J., and Robins, E. (1978). Research diagnostic criteria: Rationale and reliability. *Archives of General Psychiatry*, 35(6), 773-782.
26. Spitzer, R.L., and Fleiss, J.L. (1974). A Re-analysis of the reliability of Psychiatric diagnosis. *British Journal of Psychiatry*, 125, 341-347.
27. Rogers, R. (1995). *Diagnostic and structured Interviewing: A handbook for psychologists*. Odessa, FL: Psychological Assessment Resources, Inc.

28. Satcher, D., (1999). Mental health: A report of the Surgeon General, 1999. Online document: http://www.surgeongeneral.gov/library/mentalhealth/tables_figures.html.
29. McClellan, C., and Werry, J. (2000). Introduction: Research psychiatric diagnostic interviews for children and adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, 39(1), 19–27.
30. Gutterman, E.M., O'Brien, J.D., and Young, J.G. (1987). Structured diagnostic interviews for children and adolescents: Current status and future directions. *Journal of American Academic Child Adolescent Psychiatry*, 26, 621–630.
31. Paul, J., and Ambrosini, M.D. (2000). Historical development and present status of the schedule for affective disorders and schizophreniz for school-age children (K-SADS). *Journal of American Academic Child and Adolescent Psychiatry*, 39(1), 49–58.
32. Hodges, K. (1992). Structured interviews for assessing children. *Journal of Child Psychology & Psychiatry & Allied Disciplines*, 34(1), 49–68.
33. Costello, E., Edelbrock, C., and Costello, A. (1985). Validity of the NIMH Diagnostic Interview Schedule for Children: A comparison between psychiatric and pediatric referrals. *Journal of Abnormal and Child Psychology*, 13, 579–595.
34. Shaffer, D., et al. (2000). NIMH Diagnostic Interview Schedule for Child Version IV (NIMH DISC-IV): Description, differences from previous versions, and reliability of some common diagnoses. *Journal of the American Academy of Child and Adolescent Psychiatry*, 39(1), 28–38.
35. Herjanic, B., and Reich, W. (1983). *Diagnostic interview for children and adolescents (DICA-C): Child version*. St. Louis, MO: Washington University School of Medicine.
36. Reich, W. (2000). Diagnostic Interview for Children and Adolescents (DICA). *Journal of the American Academy of Child and Adolescent Psychiatry*, 39(1), 59–66.
37. Kovacs, M. (1985). The Interview Schedule for Children. *Psychopharmacology Bulletin*, 21, 991–994.
38. Sherrill, J., and Kovacs, M. (2000). Interview Schedule for Children and Adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry*, 39(1), 67–75.
39. Kaufman, J., Birmaher, B., Brent D, et al. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL). *Journal of the American Academy of Child and Adolescent Psychiatry*, 36, 980–988.
40. Weller, E.B., Weller, R.A., Fristad, M.A., Rooney, M.T., and Schecter, J. (2000). Children's Interview for Psychiatric Syndromes (ChIPS). *Journal of American Academic Child and Adolescent Psychiatry*, 39(1), 76–84.
41. Hodges, K., Klein, J., Fitch, P., McKnew, D., and Cytryn, L. (1981). The child assessment schedule. *Catalog of Selected Documents of Psychology*, 11, 56.
42. Reich, W., and Earls, F. (1987). Rules for making psychiatric diagnoses in children on the basis of multiple sources of information: Preliminary strategies. *Journal of Abnormal Child Psychology*, 15, 601–616.
43. Ambrosini, P.J., Metz, C., Prabucki, K., and Lee, J-C. (1989). Videotape reliability of the third revised edition of the K-SADS. *Journal of American Academic Child and Adolescent Psychiatry*, 28, 723–728.
44. Helzer, J.E., Clayton, P.J., Pambakian, R., Reich, T., Woodruff, Jr., R.A., and Reveley, M.A. (1977). Reliability of psychiatric diagnosis: II. The test/retest reliability of diagnostic classification. *Archives of General Psychiatry*, 34(2), 136–141.
45. Ruskin, P.E., et al. (1998). Reliability and acceptability of psychiatric diagnosis via telecommunication and audiovisual technology. *Psychiatric Services*, 49(8), 1086–1088.

Address reprint requests to:

Norman E. Alessi, M.D.

Psychiatric Informatics Program

Department of Psychiatry

University of Michigan Medical School

1500 East Medical Center Drive

Ann Arbor, Michigan 48109

E-mail: nalessi@umich.edu

This article has been cited by:

1. Natasha Greenberg, Katherine M. Boydell, Tiziana Volpe. 2006. Pediatric Telepsychiatry in Ontario: Caregiver and Service Provider Perspectives. *The Journal of Behavioral Health Services & Research* **33**:1, 105-111. [[CrossRef](#)]