Reliability of Self-Reported Service Use: Test-Retest Consistency of Children's Responses to the Child and Adolescent Services Assessment (CASA)

Elizabeth M. Z. Farmer, Ph.D.,[1,5] **Adrian Angold, MRCPsych,**[2] **Barbara J. Burns, Ph.D.,**[3] **and E. Jane Costello, Ph.D.**[4]

This paper reports on the reliability of children's responses on the Child and Adolescent Services Assessment (CASA)—a self-report instrument for use with 8- to 18-year-olds that gathers information about services used to address mental health problems. Findings were based on interviews completed by 77 children at a one week test-retest interval. Results showed that reports of lifetime service use were as reliable as were reports of service use in the preceding three months. Children reported restrictive and intrusive services more reliably than services that were provided in their natural environment. Reliability appeared to be associated more strongly with characteristics of the type of service than with characteristics of the child. Children also could report reliably on some details about their encounters with service providers (e.g., length of stay, number of visits, and onset of service use).

KEY WORDS: children; mental health; services; reliability; CASA.

[1]Assistant Research Professor, Department of Psychiatry, Duke University Medical Center, Durham, North Carolina.
[2]Assistant Professor, Department of Psychiatry, Duke University Medical Center, Durham, North Carolina.
[3]Professor, Department of Psychiatry, Duke University Medical Center, Durham, North Carolina.
[4]Associate Professor, Department of Psychiatry, Duke University Medical Center, Durham, North Carolina.
[5]Correspondence should be directed to Elizabeth M. Z. Farmer, Duke University Medical Center, Box 3454, Durham, NC 27710-3454.

Can children provide reliable data on services they have received for mental health problems? Is reliability similar for all types of services, or are children more reliable reporters for some types of services than for others? Are demographic and psychiatric characteristics of the child associated with reliability? This paper examines these issues by analyzing the reliability of children's reports on the Child and Adolescent Services Assessment (CASA) (Burns, Magruder-Habib, Costello, & Patrick, 1991).

Researchers interested in children's mental health services have two principal sources of possible information—data from organizations and data from individuals (either self-report or report by a knowledgeable proxy). Organizational data most often come from care-providing facilities or from insurance companies. These sources present substantial limitations for research (Burns, Angold, and Costello, 1992). They tend to exclude services provided outside the mental health system, to emphasize reimbursable or covered services, and often they fail to include information on new or innovative services.

Self-report data can overcome these limitations of organizational data. However, self-report presents its own set of difficulties. One of the primary concerns is whether self-report of service use provides reliable and valid information about services the child has received. This paper focuses on the first of these issues—the reliability of children's self-reports. Reliability refers to the consistency with which service use is reported.

Very little is known about children's ability to reliably report their use of mental health services. Much of the related psychometric research on adults' self-reports of health and mental health service use has focused on the validity of these measures rather than on the reliability (e.g., Cannell, Marquis, and Laurent, 1977; Golding, Gongla, and Brownell, 1988; Jobe et al., 1990). These studies suggest that, with certain caveats, adults can provide fairly accurate accounts of their use of health and mental health services. Comparable research has not been carried out with children. As Golding et al. (1988) discuss, studies of validity of service use are hindered by the lack of adequate data against which to compare self-reports. Given this limitation of validity studies and the fact that reliability is a necessary, though not sufficient, prerequisite for validity, we have chosen to begin the analysis of the CASA by focusing on reliability.

The present paper explores the general issue of reliability and examines three specific questions: (1) Do the characteristics of children themselves (e.g., age, sex, psychiatric symptomatology) affect the reliability of their reports?; (2) Are children's reports similarly reliable across types of services, or do children report some types of services more reliably than other types?; and (3) What levels and types of details about service use can children report reliably?

METHODS AND MEASURES

The Child and Adolescent Services Assessment (CASA)

The CASA is a semi-structured interview designed to be used in conjunction with the Child and Adolescent Psychiatric Assessment (CAPA) (Angold, Cox, Pendergast, Rutter, & Simonoff, 1991) or other structured psychiatric interviews to gather information about services that a child has received to address mental health problems. The CASA includes several features that make it particularly applicable for research on children's use of services. (1) The interview links service use to mental health/behavior problems reported during the psychiatric interview. Such linking is intended to minimize overreporting of services for non-mental health problems that could occur if respondents were asked to report any services ever used by the child. (2) The CASA obtains information about service use for mental health problems across multiple sectors (e.g., social welfare, juvenile justice, schools). (3) There are versions of the CASA for both parents and children. Parent and child are always interviewed separately. Therefore, the two interviews can be regarded as independent reports which may be used to minimize mis-reporting. In addition, the parent interview gathers information that we have found children are unable to provide on financing, service use by other family members related to the target child's difficulties, and family burden. (4) The CASA asks about lifetime use before asking about service use in the preceding three months. (5) A Detailed Child Services Form is completed for each service that the child used in the preceding three months. This form gathers information about the type of facility, the professional discipline of the provider(s), number of visits for outpatient services, length of stay for inpatient and out-of-home placements, primary problems addressed by the service, and sources of payment (parent interview only). (6) The interview can be conducted by trained lay interviewers following a half-day training. (7) For a child with a considerable history of service use, the CASA can be completed in approximately 20 minutes; for most children, the completion time is considerably shorter.

The Sample

Subjects (ages 8 to 17; mean age of 13.7 years) were recruited from local inpatient and outpatient children's mental health services in Durham, North Carolina. Of the 77 children in the sample, 40 (52%) were either currently attending or were on the waiting list for a university-affiliated child guidance clinic. The remaining 33 children were recruited while they

were inpatients at local private, public, and for-profit hospitals. The ma-
jority (67%) of inpatients were recruited from a university hospital's
adolescent inpatient unit. Forty percent of the respondents were female.
Approximately 50% of the respondents were black, 40% were white, and
10% were of other races. Eighty-six percent of the sample reported at least
one symptom of depression, anxiety, and/or a conduct problem. The ma-
jority of these children reported symptoms of both emotional and
behavioral difficulties. Service use varied widely within the sample, from
children who reported no previous use of any services, to children who
reported extensive involvement with service providers (at the upper end of
this distribution were children who reported having received services from
11 of the 30 listed types of service providers).

Data Collection

The participating clinic and hospitals referred eligible children (i.e.,
those aged 8-18) to the study when they were admitted to an inpatient
facility or were enrolled as a new client at the community guidance clinic.
Each child and family were then contacted by a member of the research
team by telephone (or in person, if the family did not have a telephone).
The family was invited to participate in the study, and both parent and
child provided informed consent. Each respondent completed two inter-
views separated by one week (mean interval = 5.45 days; range = 1-8 days).
Time 1 and Time 2 interviews were conducted by different interviewers for
all children. At each interview the child completed both the CAPA (The
Child and Adolescent Psychiatric Assessment) and the CASA. Each set of
interviews took 1.5 to 3 hours to complete, depending on the number of
symptoms and services a child reported. The CASA itself took 15-20 min-
utes of this total time.

ANALYSIS

The current analysis focuses on the reliability of responses to the
CASA. It examines (a) the reliability of reported services, (b) variation in
reliability associated with characteristics of the children, and (c) reliability
of details about service use (e.g., dates of service commencement, reports
of perceived benefits, and specific information on duration of services and
providers).
Intra-class correlation coefficients (ICCs) were used to assess reliabil-
ity for continuous variables. Such variables include (1) the total number of
services reported, (2) the dates of services commencement, and (3) infor-

mation from the Detailed Child Services Form on the number of visits or length of stay. The total number of services is a general measure of overall reliability. For this measure, the number of services (out of the possible list of 30) that a child reported at each Time 1 and Time 2 were summed. ICCs reflect the degree of consistency in the total number of services reported during the two interviews. This measure does not provide details on the types of services reported. The CASA also collects information on the date a particular service was first used by the child. ICCs were used here to assess whether children could reliably report such dates.

Kappa statistics were used to assess reliability for categorical variables. In the present analysis these include (1) a dichotomous indicator of whether the child reported use of a particular type of service at both interviews and (2) an assessment of the perceived "benefit" resulting from each utilized service.

Even in this sample recruited from clinical settings, many of the services included in the CASA were reported rarely. Therefore, conceptually similar services were combined into "service categories" to produce large enough cells for statistical analysis. The constructed categories were: inpatient; out of home; outpatient; crisis; professionals outside the mental health field; school; non-professionals; juvenile justice; and case management. Case management services were reported by only two children in the study. This reflected the lack of availability of such services in the local service system at the time of the test-retest study. Therefore the present analysis cannot assess the reliability of reports of case management. The note at the bottom of Table 1 provides details on the types of services included in each of these categories.

The relatively small sample size (n = 77) and the infrequency with which some services were reported limited the types and extent of possible analyses. Analyses did explore conceptually interesting comparisons among subgroups of children and types of services. These results should be interpreted carefully in light of the often small cell sizes, but they provide a preliminary assessment of the reliability of children's self-reports of service use.

RESULTS

Reliability for the Full Sample

Table 1 provides kappas and intra-class correlation coefficients (ICCs) for the total sample of 77 children. The ICCs in the bottom row of the table show quite high correlations between the total number of services

reported at Time 1 and Time 2. These coefficients were very similar for reports of service use in the past three months and for lifetime services (ICCs = 0.74 and 0.76, respectively). For the preceding three month period, children reported using an average of 1.94 services. This increased to 3.16 when lifetime history of service use was used as the timeframe. There was no significant difference between the number of services reported at T_1 and T_2.

Coefficients for the "last three months" were very similar to those for "ever." The kappas varied substantially, though, across the service categories. In general, higher reliability was associated with the most

Table 1. Reliability: Total Sample (n = 77)

Service Category	Last 3 Months[a]	Ever[a]
In Patient	kappa = 0.91	kappa = 10
	n = 28 (36%)	n = 34 (44%)
Out of Home	k = 0.92	k = 0.79
	n = 7 (9%)	n = 13 (17%)
Out Patient	k = 0.52	k = 0.51
	n = 36 (47%)	n = 55 (71%)
Crisis Services	k = 0.58	k = 0.60
	n = 11 (14%)	n = 18 (23%)
Non Mental Health	k = 0.58	k = 0.62
Professional	n = 13 (17%)	n = 19 (25%)
School	k = 0.39	k = 0.36
	n = 29 (38%)	n = 51 (66%)
Non Professional	k = 0.43	k = 0.48
	n = 36 (47%)	n = 44 (57%)
Juvenile Justice	k = 0.84	k = 0.94
	n = 8 (10%)	n = 10 (13%)
Number of Service Settings	ICC = 0.74	ICC = 0.76

[a]n = number of children who reported having used the service (at T_1 and/or T_2).
% = "n" expressed as a percentage of the sample.

Definitions of Service Categories

In Patient	Psychiatric Hospital, General Hospital, I.P. Detox/Drug/Alcohol, Medical In Patient
Out of Home	Residential Treatment Center, Boarding School, Group Home, Therapeutic Foster Care
Out Patient	Day Hospital, O.P. Drug/Alcohol, Mental Health Center, In-Home Counseling, Private Professional
Crisis	Emergency Room, Crisis Center, In-Home Crisis Services, Crisis Hotline
Non MH Prof	Social Services, Family MD or Other MD
School	School Guidance Counselor/Psychologist, Special Class
Non-Professional	Peers, Minister, Non Professional Adult, Self-Help Group
Juvenile Justice	Probation Officer, Detention Center

"restrictive" types of services, while lowest reliability was noted among those services that occurred within the child's natural environment.

Kappas larger than 0.80 can be considered to indicate very high reliability. Such high reliability was noted for inpatient services, out of home placements, and involvement with the juvenile justice system. All of these services were very intrusive and/or intensive. Children who had received such services appear to have remembered them and to have reported them consistently.

At the other end of the continuum were services provided within the child's natural environment. These services were concentrated in the non-professional and school categories. For these services, kappas were in the 0.36-0.48 range.

Falling between the very high reliability of the "most restrictive" services and the low reliability of the "least restrictive" services, was a set of services with mid-level restrictiveness. These mid-level services included outpatient, crisis, and non-mental health professional services. Kappas for these service categories fell between 0.51 and 0.62.

In general, then, reliability for self-reported service use was fairly high on the most global measure of such consistency—the ICC for the total number of services used. Looking more specifically at actual types of services, reliability was very good for services involving the "most restrictive" placements (i.e., inpatient, out of home, and juvenile justice). Reliability was low for services that were provided "in vivo" (i.e., school services and help from non-professionals). Services that fell between these two extremes in the intensity of the intervention and in the level of restrictiveness, disruptiveness, and/or intensity also fell between them in reliability.

Does Reliability Vary by Subgroups?

The overall findings presented above may conceal a great deal of heterogeneity in the reliability of children's reports. The following analyses examine whether such variability exists, and if it does, whether it is systematically associated with demographic and/or psychiatric characteristics of the children. For these analyses children were categorized along three dimensions: age, sex, and psychiatric symptoms.

Age

It is possible that younger children were less reliable reporters than were older children, as has been found with some diagnostic interviews (Edelbrock, Costello, Dulcan, Kalas, & Conover, 1985). ICCs at the bottom

of Table 2 begin to address this issue. These coefficients show that children younger than 13 years displayed significantly lower ICCs for total number of services reported than did older children.

Does this mean that younger children cannot report service use as reliably as older children? The upper portion of Table 3 explores this issue by presenting kappas for each of the service categories for younger (8-12 years old) and older (13-17 years old) children. Three main findings emerge from these frequencies and kappas. First, younger children reported many fewer services than did older children. Second, the services that were reported by the younger children tended to be within the less restrictive categories. Third, younger children did not appear to be significantly less reliable reporters for services that they actually reported.

The striking difference between the younger and older groups appears to be that younger children, for the most part, did not report use of the most restrictive services. No younger children reported any involvement with the juvenile justice system. Only one child reported ever having been in an out-of-home placement. Very few younger children reported inpatient services. However, younger children who did report such restrictive services seemed to report them as reliably as did older children. Therefore, it does

Table 2. Reliability By Age

Service Category	Last 3 Months[a]		Ever[a]	
	8 - 12 y.o. (n = 26)	13 - 17 y.o. (n = 51)	8 - 12 y.o. (n = 26)	13 - 17 y.o. (n = 51)
In Patient	k = 1.0	k = 0.88	k = 1.0	k = 1.0
	n = 2 (8%)	n = 26 (51%)	n = 3 (12%)	n = 31 (61%)
Out of Home	n/a	k = 0.91	k = 0	k = 0.82
	n = 0	n = 7 (14%)	n = 1 (4%)	n = 12 (24%)
Out Patient	k = 0.66	k = 0.45	k = 0.54	k = 0.47
	n = 11 (42%)	n = 25 (49%)	n = 15 (58%)	n = 40 (78%)
Crisis Services	k = 1.0	k = 0.50	k = 0.78	k = 0.54
	n = 1 (4%)	n = 10 (20%)	n = 3 (12%)	n = 15 (4%)
Non Mental Health	k = 0	k = 0.65	k = 0.26	k = 0.72
Professional	n = 2 (8%)	n = 11 (22%)	n = 5 (19%)	n = 14 (27%)
School	k = 0.34	k = 0.42	k = 0.32	k = 0.38
	n = 11 (42%)	n = 18 (35%)	n = 18 (69%)	n = 33 (65%)
Non Professional	k = 0.33	k = 0.42	k = 0.33	k = 0.40
	n = 7 (27%)	n = 29 (57%)	n = 7 (27%)	n = 37 (72%)
Juvenile Justice	n/a	k = 0.83	n/a	k = 0.94
	n = 0	n= 8 (16%)	n =0	n = 10 (16%)
Number of Service Settings	ICC = 0.50	ICC = 0.70	ICC = 0.41	ICC = 0.74

[a] n = number of children who reported having used the service (at T_1 and/or T_2).
% = "n" expressed as a percentage of the sample.

not appear that younger children were actually less reliable reporters of services that they have received. Rather, it appears that they had not experienced the types of restrictive services that showed high levels of test-retest reliability.

Sex

The coefficients in Table 3 confirm the anticipated finding that girls and boys experienced somewhat different types of services within the system. Girls did not report less extensive involvement with the system (mean number of services reported "ever" by girls is 3.55, compared to 2.93 for boys). However, they did report a different pattern of service categories.

Girls were more likely than boys to report both inpatient and outpatient services. (These comparisons of service use patterns were based on reports of lifetime use). Girls also were more likely to report having received school services at some time during their lives. As would be expected, girls were unlikely to report any contact with the criminal justice

Table 3. Reliability By Sex

	Last 3 Months[a]		Ever[a]	
Service Category	Female (n = 31)	Male (n = 46)	Female (n = 31)	Male (n = 46)
In Parent	k = 0.93	k = 0.90	k = 1.0	k = 1.0
	n = 12 (39%)	n = 16 (35%)	n = 16 (52%)	n = 18 (39%)
Out of Home	k = 1.0	k = 0.88	k = 0.71	k = 0.83
	n = 2 (6%)	n = 5 (11%)	n = 5 (16%)	n = 8 (17%)
Out Patient	k = 0.46	k = 0.56	k = 0.36	k = 0.61
	n = 16 (52%)	n = 20 (43%)	n = 25 (81%)	n = 30 (65%)
Crisis Services	k = 0.37	k = 0.69	k = 0.60	k = 0.60
	n = 4 (13%)	n = 7 (15%)	n = 8 (26%)	n = 10 (22%)
Non Mental Health	k = 0.53	k = 0.62	k = 0.82	k = 0.44
Professional	n = 5 (16%)	n = 8 (17%)	n = 8 (26%)	n = 11 (24%)
School	k = 0.33	k = 0.44	k = 0.26	k = 0.37
	n = 12 (39%)	n = 17 (37%)	n = 25 (81%)	n = 26 (56%)
Non Professional	k = 0.44	k = 0.40	k = 0.68	k = 0.32
	n = 18 (58%)	n = 18 (39%)	n = 18 (58%)	n= 26 (56%)
Juvenile Justice	n/a	k = 0.91	k = 0.65	k = 1.0
	n = 1 (3%)	n = 7 (15%)	n = 2 (6%)	n = 8 (17%)
Number of Service Settings	ICC = 0.66	ICC = 0.78	ICC = 0.74	ICC = 0.76

[a] n = number of children who reported having used the service (at T_1 and/or T_2).
% = "n" expressed as a percentage of the sample.

system—only 6% of girls (compared with 17% of boys) reported that they had had a probation officer or had been placed in a detention center.

These findings suggest that girls and boys received somewhat different services within the system. They do not suggest, however, that reliability differed consistently by sex. ICCs for total number of services were not significantly different for boys and girls. Kappas for some service categories appear to be different for the two groups. However, these results do not show any consistent pattern in direction among the various service categories nor across the two time frames of recall ("ever" and "last three months").

Psychiatric Symptoms

The type(s) of psychiatric problems that a child displays also could affect the reliability of responses. In operationalizing psychiatric problems, we have focused on symptoms included in the DSM-IIIR criteria for depression, anxiety, and/or conduct problems. Symptoms that would contribute to a diagnosis of major depression or an anxiety disorder were designated here as "emotional problems." Symptoms drawn from the criteria for conduct disorder or oppositional disorder were designated here as "conduct problems."

Eighty-six percent of the children in the sample displayed at least one symptom in one or both of these areas. Twenty-three percent of the sample (n = 18) reported at least one symptom of an emotional problem with no corresponding conduct problem. Four percent (n = 3) of the sample displayed only conduct symptom(s) with no corresponding emotional problem. Fifty-eight percent of the sample displayed at least one symptom in both the emotional and conduct spheres. Given the few children who reported either no symptoms or isolated conduct symptoms, the following comparisons focus on the service use and reliability of two subsets of children: (1) those with emotional problems only, and (2) those with both emotional and conduct problems.

Table 4 reports kappas and intraclass correlation coefficients for these two subsets. The most global assessment of reliability—the intraclass correlations for total number of services—suggests that children with both emotional and conduct problems reported lifetime service use somewhat more reliably than did children with isolated emotional problems. This difference, though, is not replicated in the more recent recall period.

Several interesting patterns emerge in the kappas for the service categories in Table 4. First, children with both emotional and conduct problems were much more likely to report the most restrictive types of services (e.g.,

inpatient, out of home, and juvenile justice) than were children whose problems included only emotional problems. Second, children with both types of problems reported outpatient and crisis services with greater reliability than did the subset of children whose problems were strictly emotional. Both of these patterns may contribute to the somewhat higher overall reliability for lifetime service use for the combined-problem group. From these findings, though, it is not clear why this difference would not also appear in reports of service use in the previous three months.

Do Children Report Details Reliably?

This final section of results examines the reliability with which children can report on more specific details of their service histories. A significant caveat applies to the following results. These results may suggest higher levels of reliability for these details than are warranted. Reliability of these details could be calculated *only* for children who provided data on the same service at both the T_1 and T_2 interviews. Hence, reliability on

Table 4. Reliability By Psychiatric Symptoms

	Last 3 Months[a]		Ever[a]	
Service Category	Emotional Only (n = 18)	Emotional and Conduct (n = 45)	Emotional Only (n = 18)	Emotional and Conduct (n = 45)
In Patient	k = 1.0	k = 0.86	k = 1.0	k = 1.0
	n= 4 (22%)	n = 21 (47%)	n= 6 (33%)	n= 24 (53%)
Out of Home	n/a	k = 0.91	k = 0	k = 0.87
	n = 0	n= 7 (16%)	n= 1 (6%)	n= 11 (24%)
Out Patient	k = 0.22	k = 0.53	k = 0.11	k = 0.57
	n = 12 (67%)	n = 18 (40%)	n = 13 (72%)	n = 33 (73%)
Crisis Services	k = 0.34	k = 0.78	k = 0.49	k = 0.70
	n = 4 (22%)	n = 6 (13%)	n = 5 (28%)	n = 10 (22%)
Non Mental Health	k = 0.45	k = 0.73	k = 0.85	k = 0.64
Professional	n = 3 (17%)	n = 8 (18%)	n = 5 (28%)	n = 11 (24%)
School	k = 0.45	k = 0.31	k = 0.28	k = 0.33
	n = 7 (39%)	n = 20 (44%)	n = 13 (72%)	n = 31 (69%)
Non Professional	k = 0.65	k = 0.56	k = 0.77	k = 0.47
	n = 8 (44%)	n= 20 (44%)	n = 8 (44%)	n = 28 (62%)
Juvenile Justice	k =0	k = 0.91	k = 0	k = 1.0
	n = 1 (6%)	n = 7 (16%)	n = 1 (6%)	n =9 (20%)
Number of Service Settings	ICC = 0.62	ICC = 0.79	ICC = 0.73	ICC = 0.71

[a]n = number of children who reported having used the service (at T_1 and/or T_2).
% = "n"expressed as a percentage of the group.

details was calculated from a select subset of children who reported reliably on the broader measure of service use.

Reliability of "Benefits"

Did children reliably report subjective benefits of services? In the CASA children were asked "Do you think that what they did was helpful to you or your family?" Further probes resulted in a final coding of "definitely positive," "neutral or no effect," or "definitely negative." If there were multiple contacts with a given service type, perceived benefit was recorded for the most recent episode. As would be expected for this type of question, responses for many services did not display distributions that warranted analysis (e.g., nearly everyone said "positive" or nearly everyone said "neutral"). Table 5 displays kappas on perceived benefit for those services that had enough children reporting use of the service and for which "benefit" was variable rather than constant. Using very lax criteria (valid data from at least 5% of the sample), 10 of the 30 services were included in this table. Because of the specificity of this information, results are displayed for specific service settings rather than for service categories throughout this section.

Perhaps the most interesting finding in Table 5 is the striking change in kappas depending on how the "neutral" response is coded. In the first column of Table 5, positive or neutral was coded "1" while negative was coded "0." In the second column, neutral responses were coded the opposite way: negative or neutral = "1" while positive = "0." For 7 of the services, this latter coding produced higher reliability coefficients. This

Table 5. Reliability of Perceived Benefit of Services

Treatment Setting	Positive/Neutral v. Negative	Negative/Neutral v. Positive
Psychiatric Hospital	k = 0.30	k = 0.22
General Hospital	k = -0.21	k = 0.56
Mental Health Center	k = -0.01	k = 0.66
Emergency Room	k = 0	k = 0.57
Social Services	*	k = 0.18
Private Professional	k = -0.01	k = 0.32
School Guidance/Psychologist	k = 0.34	k = 0.33
Family MD/Other MD	k = 0	k = 0.20
Non Professional Adult	k = 0	k = 0
Probation Officer	k = 0	k = 0.50

*Kappa could not be calculated.

seems to indicate that when children said that a service "made no differ-
ence" or "had no effect," they were viewing this as a negative evaluation
of benefit.

Reliability of Dates

Can children reliably report on when they were first in contact with
a particular type of service? As with the reliability of benefits, reliability
of dates could be calculated only for those children who reported reliably
on whether they had experienced a particular service. Of those children
who reported use of a specified service at both interviews, the majority (>
80% for most services) could provide information on the date of service
onset. Therefore, reporting dates of service onset was not beyond the ca-
pabilities of these children.

Table 6 explores the reliability of these dates provided by children.
The intraclass correlation coefficients in this table show quite good reli-
ability for most of the services. All ICCs (with the exception of school
guidance counselor/psychologist) were larger than 0.60. For 6 of the 9 re-
ported services, ICCs were larger than 0.75, with several coefficients
exceeding 0.90.

These coefficients suggest quite high reliability. However, they do
not provide an intuitive sense of what this level of reliability means. The

Table 6: Reliability of Date of Onset of Services

Treatment Setting	ICC for Dates at T_1 and T_2	% with Same Date at T_1 and T_2	% Within 6 Months at T_1 and T_2	% Within 1 Year at T_1 and T_2
Psychiatric Hospital	0.61	25%	75%	83%
General Hospital	0.93	45%	90%	95%
Group Home line	0.91	0%	67%	67%
Mental Health Center	0.92	21%	57%	71%
Emergency Room	0.99	43%	100%	100%
Private Professional	0.69	25%	45%	60%
School Guidance	0.46	22%	33%	61%
Peers line	0.76	6%	50%	56%
Probation Officer	0.87	25%	75%	88%

final three columns of Table 6 examine, more concretely, the reliability of children's reports. The first of these columns shows the percentage of children (who provided dates at both T_1 and T_2 for a service) who gave the same date at both times. The second and third columns relax the criterion to indicate the percentage of children whose two dates were within six months of each other and the percentage whose dates were within a one year span.

These figures show that exact reliability (i.e., giving the same date at both interviews) was not common. With a few exceptions in each direction, approximately one-quarter of children who provided dates at both interviews provided the same date. Reliability increased substantially if the level of precision was relaxed to include a six month span. Using this level of precision, at least half of the children reported dates reliably for most services. Relaxing precision to include a one year period increased the percentage of children who reported dates reliably slightly further.

These results suggest that children can report dates of service onset with reasonable reliability within a six-month span of precision. The pattern of higher reliability for more restrictive services that was noted earlier can also be seen in these data (although the pattern was not as clear as for the dichotomous measure of service use). For highly restrictive services (particularly hospitalization and probation), 75% of children were reliable to within six months. For the least restrictive services (peers and school guidance), 33-55% of children were reliable at the same level of precision. Reporting of the mid-level set of services (represented here by mental health center, emergency room, and private professional) was mixed. Emergency room had perfect reliability at the six-month level of precision. The other two mid-level services suggest that reliability for these types of services may again fall between the rather high reliability of restrictive services and the low reliability of least restrictive services.

Detailed Child Services Form

A Detailed Child Services Form (DCSF) was completed for each service that a child reported having used in the preceding three months. As part of this form, the interviewer gathered specific information about the type of provider the child saw (e.g., psychiatrist, psychologist, social worker) and the length of stay or number of visits. Such information is essential for calculating estimates of cost and to assess intensity of service use. Information on sources of payment is gathered only in the parent interview.

An initial set of interviews with children suggested that children were unable to provide information about the financial aspects of their care. Therefore, these questions were deleted from the child interview.

Children appear to be able to quite reliably report the length of stay of inpatient episodes. A sufficient number of children reported inpatient experiences in psychiatric hospitals and in psychiatric units of general hospitals to allow analysis. Intra-class correlation coefficients for the length of stay across the T_1 and T_2 interviews were 0.99 for psychiatric hospitals and 0.88 for general hospitals. As noted previously, children reported very reliably whether they had ever been in an inpatient setting. From the DCSF, we see that they also displayed a high level of reliability when reporting on the duration of such stays (at least for inpatient episodes that occurred during the previous three months). Taking these two findings together, it appears that children's self-reports were a reliable source of information about their inpatient experiences.

For outpatient services, it was possible to examine the reliability for two details—the number of visits and the specific provider. A sufficient number of children reported reliably for two specific outpatient settings—community mental health centers and private professionals—to allow analysis.

Reliability of reported number of visits was rather different for these two outpatient settings: ICC for mental health centers = 0.96, ICC for private professionals = 0.58. On average, children reported more visits to mental health centers during the preceding three months than to private professionals (means of 11.4 and 7.0 visits, respectively).

The second analysis of these details focuses on the professional discipline of the provider that the child saw. The small number of children who reliably reported mental health center visits and visits to a private professional made it difficult to quantify reliability of these reports. Various ways of examining the data, however, suggested that children did not report providers' disciplines very reliably. Of children who reported services in a mental health center, 29% (n = 2) reliably reported the provider's discipline. Similarly, 33% (n = 2) of children who reported visiting a private professional reliably reported the provider's discipline. In both settings, children who reported reliably identified the provider as a psychiatrist. "Psychiatrist" was also the most frequently named provider. Because we have not yet evaluated the validity of these responses, it is unclear whether children actually reported visits to psychiatrists more reliably than to other professionals or if children were likely to respond that they had seen a psychiatrist regardless of the professional's discipline.

DISCUSSION

What does this analysis say about the reliability of children's self-reports of service use? The overall message is generally encouraging. Intraclass correlation coefficients for the total number of services reported at the two interviews were fairly high. The designated timeframe of recall did not appear to affect the reliability of children's responses—children reported as reliably on lifetime service use as they did on services used in the three months preceding the interview. Looking at reliability for specific types of services, analysis showed that children reported some services more reliably than others. Level of reliability appeared to correspond to the restrictiveness of the service. Services at the most restrictive end of the continuum (out of home placements, inpatient services, and contacts with the juvenile justice system) were reported most reliably. Services provided "in vivo" (school services and non-professional help) were reported least reliably. A set of services that fall in the middle of this continuum of restrictiveness (e.g., outpatient and crisis services and services from non-mental health professionals) displayed moderate reliability.

Demographic and psychiatric characteristics of the child were associated with different patterns of service use and reliability. These results, though, must be interpreted carefully. Younger children initially appeared to be less reliable reporters than were older children. On closer inspection, though, it appeared that this apparent difference rested on the types of services the two age groups reported having used, not on an inability of younger children to report services reliably. For the most part, younger children reported use of the less restrictive types of services. Such services were not reported very reliably by any children, regardless of age. Younger children who did report use of the most restrictive services appeared to report these as reliably as did older children.

Girls reported a different distribution of services than boys. However, there was no evidence of systematic differences in the reliability of their reports.

The type of psychiatric symptoms a child reported may be associated with both the types of services reported and the reliability of these reports. Children with both emotional and behavioral problems displayed ICCs for lifetime service use that were somewhat higher than those displayed by children with emotional problems only. Children with symptoms in both areas also reported more encounters with the most restrictive services and higher reliability of reports of mid-level services (e.g., outpatient and crisis services). Type of psychiatric problem, then, was associated with the types of services received and may have affected the reliability with which at least some of these services were reported.

From the analysis of responses to more detailed information about services, it appears that children did not report subjective benefit very reliably. This analysis suggested that children may have used "neutral" responses to indicate that a service was "not helpful." When reporting dates of service onset, at least half of the children were reliable within a six-month window for most services. Analysis of information from the Detailed Child Services Form indicated that children very reliably reported lengths of recent inpatient stays and number of visits to mental health centers. They did not report the discipline of the provider very reliably. Recall that analyses of such details were run using the subset of children who reported reliably on the more global measure of service use.

Conclusions and Recommendations for Use of the CASA

Existing data sources (e.g., insurance company and service facility records) do not provide adequate information on the services that children use to address mental health issues. The CASA is proposed as a new means of gathering these crucial, yet difficult to collect, data. This analysis of the reliability of the instrument suggests that children can be useful and informative reporters on the types of services they have experienced. Children at least as young as eight-years-old can complete the CASA with reasonable reliability. Children were particularly reliable in their reports of very restrictive services. Unfortunately they were much less reliable reporters of services that were provided in their natural environment (i.e., from school personnel and informal supports such as peers). These latter services (particularly informal sources) are also the types of services that are least likely to be captured by other data sources (e.g., official records or reimbursement information).

There may by several factors contributing to the low reliability of reports on non-professional and school services. For the most part, non-professional services refer to help from peers. Given the informality of such services, it is not surprising that they may be reported differently at separate interviews. It may be that school services also display low levels of reliability for a similar reason (i.e., they were provided within the natural environment, and therefore, did not stand out in the child's view). Another explanation for the low reliability of school services may include the CASA's attempt to link services to behavioral or psychiatric problems that were revealed earlier in the interview. Following this explicit linking, special school services provided to address learning disabilities, for example, are not included here because such problems are not assessed in the diagnostic interview. Since children may not understand why they received a particular school service, there may be confusion and inconsistency in the reporting

of such services. Whatever the cause, school services were the least reliably reported type of service in the CASA.

From these analyses it does not appear that reliability is dependent upon characteristics of the child. Age and sex were associated with the types of services received, but they did not appear to influence the reliability of children's reports of services that they received. Children with isolated emotional symptoms may be slightly less reliable reporters than children who also displayed conduct problems, but these differences were not large. Reliability seemed to be associated with the type of service more significantly than with characteristics of the child.

Two further analyses of the CASA's reliability and validity are currently being conducted. In the first, we are comparing reports provided by children and parents. A preliminary analysis of the first 45 parent-child pairs suggests that parents' reliability and reports of service use are quite similar to those provided by their children. We have not yet analyzed the parent data in detail. One goal of this analysis will be to determine if interviewing both parent and child provides redundant information or if dual respondents produce a stronger and more complete record than that produced by either respondent alone. The second analysis assesses the validity of the CASA by comparing self- and parent-reports to formal records (e.g., clinic charts, police records, insurance information). The present analysis suggests that children could report reliably on many of the types of services covered by such official records. As noted in the introduction, such organizational records may contain significant errors and omissions. However, they remain the best standard against which to validate self-report data.

From the completed analysis, it appears that the CASA can be a useful instrument for researchers interested in children's mental health issues. Children appear to report very reliably on the most restrictive types of services. They also reported mid-level services (e.g., out patient) with reasonable reliability. In addition, children can report not only that they have used services, but they can also reliably report on some details of their service use (e.g., date of onset, length of stay, number of visits). Using the current version of the CASA, however, reports of informal and school-based services should not be regarded as reliable. Future instrument development needs to focus on ways to more reliably gather data on services provided by these informal and school-based sources.

ACKNOWLEDGMENTS

REFERENCES

Angold, A., Cox, A., Pendergast, M., Rutter, M., & Simonoff, E. (1991). Child and Adolescent Psychiatric Assessment (CAPA). Durham, NC: Developmental Epidemiology Program, Duke University Medical Center.

Burns, B. J., Angold, A., Magruder-Habib, K., Costello, E. J., & Patrick, M. K. S. (1991). Child and Adolescent Services Assessment (CASA). Durham, NC: Developmental Epidemiology Program, Duke University Medical Center.

Burns, B. J., Angold, A., & Costello, E. J. (1992). Measuring child, adolescent, and family service use. In L. Bickman & D. J. Rog (Eds.), Evaluating mental health services for children (pp. 17-30). San Francisco: Jossey-Bass.

Cannell, C. F., Marquis, K. H., & Laurent, A (1977). A summary of studies of interviewing methodology. Vital and Health Statistics, 69: 1-78.

Edelbrock, C., Costello, A. J., Dulcan, M. K., Kalas, R., & Conover, N.C. (1985). Age differences in reliability of the psychiatric interview of the child. Child Development 56: 265-275.

Golding, J. M., Gongla, P., & Brownell, A. (1988). Feasibility of validating survey self-reports of mental health service use. American Journal of Community Psychology, 16, 39-50.

Jobe, J. B., White, A. A., Kelley, C. L., Mingay, D. J., Sanchez, M. J., & Loftus, E. J., (1990). Recall strategies and memory for health-care visits. The Millbank Quarterly, 68, 171-189.

Tuabe, C. A., Schlenger, W. E., Rupp, A., & Whitmore, R. W. (1986). Validity of medicaid household respondent reporting of ambulatory visits for mental disorders. Journal of Economic and Social Measurement, 14, 243-256.