International Journal of Methods in Psychiatric Research
Int. J. Methods Psychiatr. Res. 25(1): 33–43 (2016)
Published online 15 July 2015 in Wiley Online Library
(wileyonlinelibrary.com) **DOI:** 10.1002/mpr.1477

Changes in mental health outcomes with the intensive in-home child and adolescent psychiatric service: a multi-informant, latent consensus approach

BAPTISTE BARBOT,[1,2] JOHANNA BICK,[3] MARY JANE BENTLEY,[2] KATHLEEN M.B. BALESTRACCI,[2] JOSEPH L. WOOLSTON,[2] JEAN A. ADNOPOZ[2] & ELENA L. GRIGORENKO[2,4]

1 Department of Psychology, Pace University, New York, USA
2 Child Study Center, Yale University, New Haven, CT, USA
3 Boston Children's Hospital/Harvard Medical School, Boston, MA, USA
4 Moscow State University of Psychology and Education, Moscow, Russia

Key words
severe emotional disturbance, IICAPS, in-home intervention, children, adolescents, Ohio Scales, psychiatric hospitalizations

Correspondence
Elena L. Grigorenko, PhD, Child Study Center, 230 South Frontage Road, P.O. Box 207900, New Haven, CT 06520-7900, USA.
Telephone (+1) 203-737-2316
Email: Elena.Grigorenko@yale.edu

Received 30 December 2014; accepted 18 March 2015

Abstract

This study investigates the Intensive In-home Child and Adolescent Psychiatric Service (IICAPS), a large-scale home-based intervention that collaboratively engages the family, school, and various other service providers (e.g. health practitioners or judicial systems) to prevent the hospitalization, institutionalization or out-of-home placement of children and adolescents with serious emotional disturbance. Multi-informant data (youth, parents and clinician) on the level of youth problem severity and functioning was gathered from 7169 youth and their families served by the IICAPS network, pre- and post-intervention. A newly developed "Multi-informant Latent Consensus" (MILC) approach was employed to measure mental health "baseline levels" and change, within a Structural Equation Modeling framework. The MILC approach demonstrated promise integrating information from multiple informants involved in the therapeutic process to yield a more accurate and systemic view of a child's level of functioning and problem severity than each report taken individually. Results indicated that the IICAPS family and community based intervention model led to a reduction of problem severity and improved functioning in children and adolescents with severe emotional disturbance.

Introduction

Children and adolescents with severe emotional disturbance (SED) are at high-risk for psychiatric hospitalizations, residential treatment or placement in out-of-home care. The adverse impact of out-of-home placement for these youth has been evidenced by many studies (e.g. Gowers *et al.*, 2000; Harpin *et al.*, 2013). For example, psychiatric hospitalization may cause strong affective reactions in adolescents and affect psychosocial dimensions of youth (Haynes *et al.*, 2011), and may even lead to significantly worse mental health outcomes than in-home treatment (Gowers *et al.*, 2000). Further, residential treatment programs are costly, and may not address the complex and multi-systemic needs of children and their families. In this study, we present evidence that an ecologically oriented, family-focused, and home-based intervention, the Intensive In-home Child and Adolescent Psychiatric Service (IICAPS), led to improvements in psychosocial functioning of children with SED, which should significantly decrease their risk of psychiatric hospitalizations, residential treatment and placement in out-of-home care.

IICAPS is a widely implemented structured mental health treatment supported by the Connecticut Department of Children and Families (DCF), authorized for payment by Medicaid, and delivered at 20 provider sites within the state of Connecticut. Grounded in developmental psychopathology, the IICAPS intervention posits that a child's developmental trajectory is determined by the complex, continuous interactions between the youth's inherited characteristics and his/her environment. Many children and adolescents enrolled in IICAPS face exposure to adverse conditions, which often include being reared in families with multiple physical and mental health problems and socio-economic risk factors. Youth served by IICAPS are often exposed to increased rates of violence within the family and surrounding community and are likely to have significant trauma histories. To address the multiple and complex factors that contribute to youth risk, IICAPS strives to enhance the "quality of fit" between the child and the all-encompassing family, school, and community systems in which he or she is embedded (Adnopoz *et al.*, 2012). This is accomplished by working collaboratively with the child and family in the delivery of all phases of the treatment, by incorporating additional stakeholders (i.e. from judicial and/or school settings) into the treatment process, and by eventually facilitating connections with additional community resources (i.e. in mental health, vocational, or recreational arenas) to support sustained improvements in youth well-being.

The IICAPS program includes home-visits with the child and family by a clinical team four to five hours per week for an average treatment length of six months. The first step of the treatment program is to identify each child's "Main Problem," defined as the behavior most likely to lead to the child's psychiatric hospitalization. The Main Problem is co-constructed by the youth and family with the clinical team immediately following the start of treatment. The IICAPS intervention targets the Main Problem as it is manifested in four domains influencing the child's functioning: child, family, school and community. The IICAPS intervention proceeds through three treatment phases: Engagement and Assessment (creation of an initial treatment plan focused upon the child's Main Problem), Work and Action (focused upon meeting the Goals and Action Steps that constitute the treatment plan), and Ending and Wrap-up (focused upon strengthening the linkages between the child and family and the services needed to sustain the gains they have made during treatment). A complete description of the intervention can be found elsewhere (Adnopoz *et al.*, 2012; Woolston *et al.*, 2007).

Multi-informant approaches to the assessment of mental health outcomes: in search of consensus

Over the past two decades, IICAPS has been shown to effectively decrease utilization of psychiatric hospitals and emergency rooms and improve child and family functioning (Adnopoz *et al.*, 2012). Among other evidence of treatment efficacy, these outcomes were drawn from analyses of multi-informant reports of child functioning, which were analyzed separately for each informant, which could include the child, parent, and clinician. Methodological approaches that capture a multi-dimensional view of the child (i.e. integrate multi-informant reports into a "consensual" compound) may provide more realistic and reliable estimates of child functioning and may more fully capture the degree of improvement in children's functioning across the IICAPS treatment program.

Multi-informant approaches have traditionally been used to provide an "objective" assessment of child and adolescent behavioral and emotional symptoms (Carlston and Ogles, 2009). Classical approaches have combined information from multiple informants in a number of ways (e.g. when symptom or disorder is endorsed as present by at least one informant, or when the sum of each informant report is used). However, these methods may not be more incrementally reliable than approaches that consider informants' ratings independent from one another (e.g. De Los Reyes and Kazdin, 2005; Offord *et al.*, 1996) and they lead to different conclusions regarding the prevalence, comorbidity and correlates or risk factors of disorders (e.g. De Los Reyes and Kazdin, 2005; Youngstrom *et al.*, 2000).

In order to integrate information from multiple informants in a meaningful way, a body of work has focused on discrepancies in reports of child functioning across multiple informants. Research has shown that parent and child reports of child behavior and emotional functioning often diverge with regard to symptom presence and severity, yielding a significant amount of contradictory diagnostic information (e.g. Achenbach *et al.*, 1987; Carlston and Ogles, 2006; De Los Reyes and Kazdin, 2005). For example, research with clinical populations suggests that parents often report greater levels of child symptoms than their children (e.g. Ivens and Rehm, 1988). Informants may differ in their motivations for providing ratings of children, and in their perceptions of what constitutes abnormal behavior in their child (Richters, 1992). Additional sources of informant discrepancies may include variations in demographic and cultural background, representation of "typical child behavior" (e.g. Carlston and Ogles, 2006; Richters, 1992), informant's own symptoms, relationship to the child (De Los Reyes and Kazdin, 2005; Treutler and Epkins, 2003), and environmental features (Kanne *et al.*, 2009). Potential methodological issues such as variations in item content for each informant may contribute to these discrepancies as well (Epkins, 1996). Indeed, some multi-informant assessments of mental health rely on different sets of items for each informant, in order to measure the same underlying construct (e.g. depression). Previous research has shown that such differences in item content results in a lack of measurement invariance that has a dramatic impact on the estimation of mean differences (e.g. Barbot *et al.*, 2014b; De Beuckelaer and Swinnen, 2010), hence potentially obscuring the "true" concordance versus discrepancies in reporting between informants.

While the degree of discrepancy across informant reports undoubtedly represents a clinically important area of study (Achenbach, 2011; De Los Reyes, 2011), the "concordance" portion between informants may be equally meaningful in assessment of child functioning. Yet, estimates of concordance in multiple informants' reports have received surprisingly limited attention (e.g. Dunlop *et al.*, 2011; Grigorenko *et al.*, 2010). This may be because the "concordance" portion between informants is often underestimated (inter-correlation coefficients between reports are likely to be attenuated due to measurement error).[1] However, even attenuated, those correlations are not trivial, usually in the 0.40–0.50 range; (for meta-analyses of inter-informant concordance in the evaluation of behavioral and emotional problems among both clinical and non-clinical samples, see Achenbach *et al.*, 1987, 2005).

Indeed, the "concordance" often represents a greater portion of the total variance than the "discrepancies" in reports across informants. For example, based on a large scale study of 2850 youth using the multi-informant Ohio scales (Ogles *et al.*, 2001) as part of the Ohio Department of Mental Health Consumer Outcomes system (Ohio Department of Mental Health, 2003), we ran a secondary analysis suggesting that the amount of variance explained by inter-informant concordance reports on child functioning is substantial. Specifically, the first component of a principal component analysis (PCA) explained 69% of the total variance in children, parent, and clinician reports of a child's Problem Severity (resulting from an average inter-correlation between informants of $r = 0.53$), and 63.6% of the total variances in a child's Functioning scores across informants (average inter-correlation $r = 0.45$). In this example, only 31% to 37% of the total variance was due to discrepancies related to informants' "uniqueness" (specificity/subjectivity), as well as measurement error.

The current study

In the current study, we apply a newly-developed "Multi-informant, Latent Consensus" (MILC) approach to integrate multiple informant reports of Ohio scales collected as part of the IICAPS program. The aim of this approach was to derive a reliable estimate of the child's mental health that was free of each informant's subjectivity and measurement error. This approach partitions out the "concordance portion" from the "discrepancy portion" of multiple informant reports modelled within the Structural Equation Modeling (SEM) framework.

In this study, MILC was used to derive latent consensus estimates of a child's functional assessment of mental health (functioning and problem severity) across child, parent, and clinician reports on the Ohio scales administered at both intake and discharge of the IICAPS intervention. Specifically, we first derived latent consensus (MILC) estimates of child functioning and problem severity at intake and discharge. Next, we examined the degree of change in functioning and problem severity across measurement occasions. Finally, we compared the MILC-derived latent consensus estimates of child functioning and problem severity to "community benchmarks" estimated from normative data previously published in the literature. We hypothesized that the IICAPS intervention would yield significant improvement in mental health, as

[1] Traditional "discrepancy scores" (i.e. generally arithmetic difference scores between two informants' reports) that are the focus of a large body of research, may be particularly unreliable due to cumulated measurement error (accumulation of measurement error of each informant report).

operationalized by MILC-derived latent consensus estimates of child functioning and problem severity, incorporating parent, youth and clinician reports. We also hypothesized that the latent consensus estimates would be more similar to "community benchmark" estimates of child functioning at discharge when compared with intake assessments.

Method

Participants

Data for this study were collected across 14 IICAPS agencies in Connecticut serving children and families from July 2006 through June 2012. The total sample included 7169 children and adolescents (37.8% girls, 62.2% boys), their parents or other guardians and clinicians. Youth age ranged from 3 to 19, with a mean of 12.2 years (standard deviation [SD] = 3.5). Their racial/ethnic backgrounds included: 43.0% Caucasian, non-Hispanic, 14.3% African American, non-Hispanic, 35.2% Hispanic, and 7.5% multi-racial or another race. Youth referred for IICAPS services were discharged from psychiatric hospitals, were identified as being at risk for psychiatric hospital-based treatment, or were found to be unresponsive to less intensive outpatient clinic-based services (Adnopoz *et al.*, 2012). Referral sources included outpatient mental health providers, inpatient psychiatric hospitals, schools, DCF, and the court or juvenile justice systems. Presenting problems of the participants involved in IICAPS included aggression, acting out behaviors, and affective, anxiety, attachment, obsessive-compulsive, psychotic and stress-related psychiatric disorders. Among enrolled participants, 65.0% were identified as having completed treatment; 18.2% were discharged before treatment was completed due to a family-related interruption (e.g. family moved or was not involved in treatment), 10.6% were discharged to out-of-home treatment or care (e.g. psychiatric hospitalization, incarceration, residential treatment), and 5.6% were discharged by the agency or for "other" reasons (e.g. child needs other outpatient services, family not complying with agency policy).

Procedure

All IICAPS treatment and data collections are administered by a team of two mental health professionals, of whom at least one is a master's level. Each team is required to participate in 15 hours of training before being fully engaged in the work. As part of their IICAPS training, all team members learn to administer the IICAPS tools and measures, which provide data on family history, functioning, socio-economic status, strengths, vulnerabilities and the factors which have promoted or precipitated the child's problem behaviors. This information is combined with the clinical data presented during weekly case supervision with experienced supervisors and at regularly scheduled rounds co-led by a child psychiatrist. It is used to establish treatment plans and identify the goals and actions steps which the family will pursue. Additional data is collected at specified time-points throughout the intervention and serves multiple purposes: it provides a rating of progress towards treatment goals, allows for monitoring of adherence to the model and serves as a quality assurance mechanism, and enhances ongoing engagement by family members. IICAPS has utilized the Ohio scales as one measure of program effectiveness in two important areas, functioning and problem severity as well as parent hopefulness and satisfaction. The Ohio scales are administered to youth and their parents or guardians by their assigned IICAPS team during the Engagement and Assessment phase, which occurs in the initial treatment period. The scales are administered again during the Ending and Wrap-up phase of treatment which occurs prior to discharge.

Measures

Child functioning was assessed using the Ohio scales short-form (Ogles *et al.*, 2001). Parents or guardians, children, and clinicians completed the assessment at both intake and discharge. The "youth report form" was developed for children 12 and older; therefore, youth self-report data was not available for children ages 11 years and younger (48.6% of the sample). Two Ohio scales were used for this study. The first was the "Functioning" scale, which assesses behaviors and competencies, such as the ability to get along and develop relationships with others, practice proper hygiene, control emotions, accept responsibility, and tackle tasks of daily life. For example, informants rate how well children are "Doing things without supervision" or "Getting along with friends". The second was the "Problem Severity" scale, which assesses the degree to which the designated youth's problems affect his or her current ability to engage in everyday activities. For example, informants rate the degree to which children are "Getting into fights" or "Feeling sad or depressed". Both of these scales contain 20 items (identical across informants); the Problem Severity items are rated on a six-point scale, and the Functioning items are rated on a five-point scale. The Ohio scales have demonstrated good reliability and evidence of criterion validity with DSM-IV corresponding diagnosis across informants (Turchik *et al.*, 2007).

Data analyses

A set of preliminary analyses was conducted to examine the internal consistency and distributional features of the

Ohio scales. The inter-scales correlation matrix was inspected to ensure the suitability of the data for use in the planned analyses. Confirmatory factor analysis (CFA) as implemented in AMOS 18 (Arbuckle, 2009) was conducted to test the MILC model of Functioning and Problem Severity at both intake and discharge. The MILC model was defined by two latent variables (LVs) referring to Functioning and Problem Severity. Each LV was loaded into the three corresponding observed variables (one for each informant report). Correlations between both LVs and residual terms of each common-informant's scales were freely estimated to account for potential systematic error variance (informant "bias"). The resulting model is depicted in the SEM diagram in Figure 1. To identify the model, we used an arbitrary marker variable (the clinician report variables) by fixing its factor-loading at one and its intercept to zero, providing a metric for the underlying LV.

In a second step, both intake and discharge MILC models were integrated into a single longitudinal multivariate factor model (e.g. Barbot *et al.*, 2014a; McArdle, 2009; McArdle and Nesselroade, 1994) to test for measurement invariance over time (e.g. Meredith and Horn, 2001). This procedure was used to ensure that the latent consensus captured at intake and discharge had the same substantive meaning (e.g. an identical "weight" of each informant in the consensual score) so that latent means could be interpreted meaningfully over time. Accordingly, a Congeneric model (no restrictions on factor loadings and covariance structure across measurement occasions) was compared to models with increasing invariance stringency. Stringency steps included Weak Invariance (which imposed equal factor loadings over time), Strong Invariance (which added the constraint of equal indicators intercepts over time to the previous model), Strict Invariance (which added the constraint of equal residual variances over time to the previous model), and Structural Means (which assumed that latent means of Functioning and Problem Severity MILC were equal at intake and discharge). Finally, we implemented a multiple-common-factor Latent Change Score (LCS) model (e.g. Barbot *et al.*, 2013) which allowed for the estimation of mean-level change in the MILC Problem Severity and Functioning scores over time, as well as the correlations between these LCSs.

All SEM parameters were estimated using a Full Information Maximum Likelihood algorithm (Schafer and Graham, 2002). We emphasized practical fit indices for assessing model fit (Kline, 2010): normed fit index (NFI), comparative fit index (CFI), and root mean square error of approximation (RMSEA) with a 90% confidence interval (CI). Invariance decision was based on the CFI between the Congeneric and most restricted model, with values lower than 0.01 denoting strict invariance (Byrne, 2010; Cheung and Rensvold, 2002).

Results

Preliminary analyses

The Ohio scales for each domain showed high internal consistency coefficients at intake (Cronbach's alpha ranged from 0.86 to 0.91, median = 0.88) and at discharge

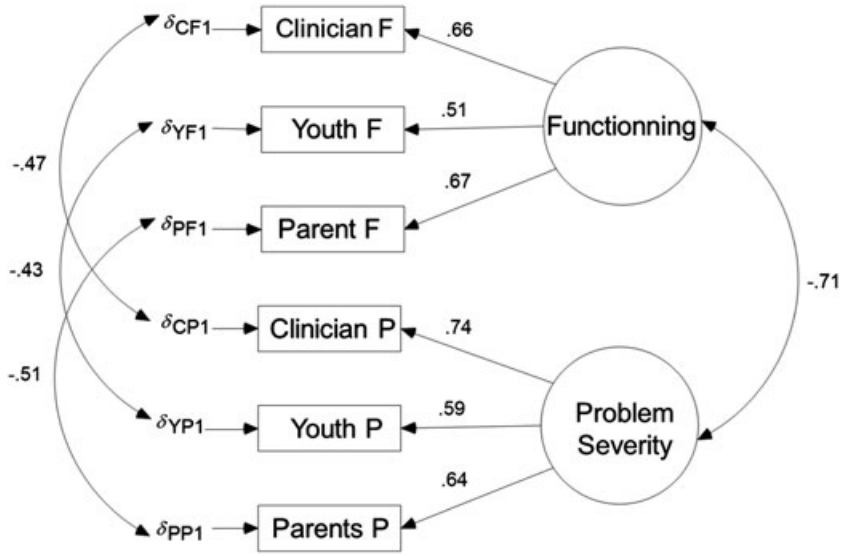

Figure 1. Multi-informant Latent Consensus (MILC) model and its estimated standardized parameters at intake. (F = Child Functioning; P = Problem severity).

(Cronbach's alpha ranged from 0.86 to 0.94, median = 0.91), with acceptable distributional features for planned analyses (Table 1). The correlation matrix used as the basis of all models in this study (Table 1) was not an identity matrix (Bartlett's test of sphericity = 13886.7, $dl = 66$, $p < 0.001$), and showed an acceptable global measure of sampling adequacy (KMO = 0.75; MSAs [0.73–0.76]). Consistent with prior research (Achenbach *et al.*, 1987; Ohio Department of Mental Health, 2003), inter-informant consensus ranged from mean $r = 0.37$ (Functioning at intake) to 0.50 (Problem Severity at discharge) with a mean of $r = 0.44$ across Ohio scales and measurement occasions (all p values < 0.001). Correspondingly, the first component of the PCAs on the multi-informant data explained between 57.4% and 66.1% (mean = 62%) of the total variance in Functioning and Problem Severity scores across informants and measurement occasions. Hence, a minor portion of the total variance (i.e. 38% on average) could be attributed to informants' discrepancies as well as measurement unreliability, justifying the development of the MILC models to capture the consensual portion of the multi-informant report.

Child functioning and problem severity MILC cross-sectional and longitudinal models

As reported in Table 2, both intake and discharge MILC models returned an adequate fit to the observed data with highly similar estimates of model parameters at both time points. Figure 1 illustrates the SEM diagram of the Functioning and Problem Severity MILC model, and estimated model parameters obtained with the intake data. All factor loadings and covariance estimates were significantly different from zero ($p < 0.001$), with standardized loadings ranging from 0.51 to 0.74 at intake (mean = 0.64), and from 0.55 to 0.80 at discharge (mean = 0.72). Estimated correlations between Functioning and Problem Severity MILCs scores were high, with $r = -0.71$ at intake and $r = -0.84$ at discharge.

Both intake and discharge MILC models were integrated into a longitudinal multivariate factor model to test for the measurement invariance of the underlying constructs over time. As indicated in Table 2, the Configural, Weak Invariance, and Strong Invariance models yielded a similar and adequate fit to the data, with the most constrained model (Strong Invariance) associated with a marginal degradation in model fit (CFI = 0.003) when compared to the Configural model. However, the most stringently invariant model (Strict Invariance) was associated with a significant degradation in model fit (CFI = 0.019). Modification indices suggested that

Table 1. Descriptive statistics and inter-correlations of input data

		Descriptive					Correlations											
	Scale	N	Mean	SD	Skew (SE)	Kurt.(SE)	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	11.	12.
1.	Intake Parent Functioning	6723	39.04	14.80	0.10 (0.03)	−0.31 (0.06)	—											
2.	Intake Parent Problem Severity	6726	34.27	16.82	0.31 (0.03)	−0.25 (0.06)	−0.60	—										
3.	Intake Child Functioning	3685	52.86	13.82	−0.48 (0.04)	0.29 (0.08)	0.33	−0.27	—									
4.	Intake Child Problem Severity	3690	25.37	16.17	0.76 (0.04)	0.14 (.08)	−0.26	0.40	−0.51	—								
5.	Intake Clinician Functioning	6633	37.36	11.81	−0.12 (.03)	−0.01 (0.06)	0.45	−0.32	0.31	−0.22	—							
6.	Intake Clinician Problem Severity	6633	32.34	13.71	0.43 (0.03)	0.25 (0.06)	−0.36	0.48	−0.30	0.45	−0.59	—						
7.	Discharge Parent Functioning	5578	47.18	15.52	−0.32 (0.03)	−0.28 (0.07)	0.45	−0.28	0.18	−0.16	0.29	−0.21	—					
8.	Discharge Parent Problem Severity	5585	22.98	15.25	0.95 (0.03)	0.73 (0.07)	−0.32	0.46	−0.13	0.23	−0.21	0.29	−0.68	—				
9.	Discharge Child Functioning	2903	58.52	12.88	−0.74 (0.05)	0.99 (0.09)	0.20	−0.16	0.39	−0.28	0.19	−0.16	0.40	−0.31	—			
10.	Discharge Child Problem Severity	2905	17.20	13.01	1.32 (0.05)	2.31 (0.09)	−0.18	0.26	−0.27	0.49	−0.15	0.25	−0.35	0.45	−0.54	—		
11.	Discharge Clinician Functioning	6652	44.29	14.19	−0.31 (0.03)	−0.21 (0.06)	0.24	−0.14	0.14	−0.10	0.42	−0.26	0.57	−0.48	0.38	−0.34	—	
12.	Discharge Clinician Problem Severity	6653	23.64	14.03	0.87 (0.03)	0.54 (0.06)	−0.19	0.22	−0.13	0.21	−0.31	0.41	−0.49	0.59	−0.35	0.47	−0.74	—

Note: Skew. = Skewness; Kurt. = Kurtosis; SD = standard deviation; SE = standard error; all correlation coefficients are significant at $p < 0.001$.

Int. J. Methods Psychiatr. Res. 25(1): 33–43 (2016). DOI: 10.1002/mpr

Table 2. Goodness-of-fit indices and invariance testing of the parenting models

Model	χ^2	df	χ^2/df	*p*	$\Delta\chi^2$	NFI	CFI	Δ CFI	RMSEA (CI)
Cross-sectional models									
Intake MILC	52.31	5	10.46	0.001	—	0.995	0.996	—	0.036 (0.028–0.046)
Discharge MILC	24.34	5	4.87	0.001	—	0.998	0.999	—	0.023 (0.015–0.033)
Longitudinal models									
Configural	146.76	30	4.89	0.001	—	0.995	0.996	—	0.023 (0.02–0.027)
Weak Invariance	192.29	34	5.66	0.001	0.001	0.994	0.995	0.001	0.025 (0.022–0.029)
Strong Invariance	236.59	38	6.23	0.001	0.001	0.992	0.993	0.003	0.027 (0.024–0.030)
Strict Invariance	751.94	44	17.09	0.001	0.001	0.975	0.977	0.019	0.047 (0.044–0.050)
Partial Invariance	362.33	42	8.63	0.001	0.001	0.988	0.989	0.007	0.033 (0.030–0.036)
Structural Means	2898.35	44	65.87	0.001	0.001	0.904	0.905	0.101	0.095 (0.092–0.098)

Note: χ^2 = chi-square; df = degrees of freedom; *p* = *p* value of the chi-square test; $\Delta\chi^2$ = *p* value of the chi-square difference test; CFI = comparative fit index; ΔCFI = difference in the CFI value (assuming baseline model to be correct); RMSEA = root mean square error of approximation; CI = 90% confidence interval of RMSEA value.

constraints on the equality of unique factors variance in both the child and parent reports on Problem Severity were the main source of misfit. In other words, these informants' unique perceptions of child's Problem Severity changed substantially over time.

Hence, a Partial Invariant version of this model relaxing the constraint on the aforementioned uniqueness variance was developed. This model included sufficient invariance constraints, associated with a marginal degradation in model fit (CFI = 0.07), to validly estimate differences in latent means overtime. Parameter estimates of this model suggested a moderate rank-order stability of child Functioning ($r = 0.50$, $p < 0.001$) and Problem Severity ($r = 0.52$, $p < 0.001$) between intake and discharge. Latent means of the Functioning MILC were 37.48 at intake and 44.23 at discharge, while means of the Problem Severity MILC were 32.51 at intake and 23.52 at discharge (all *p* values < 0.001).

Mean change under the influence of IICAPS

Building upon the Partial Invariant model, the longitudinal Structural Means model of child Functioning and Problem Severity MILC revealed a poor fit of the model to the observed data, with dramatic decrease in fit when compared to the Congeneric model (CFI = 0.101; Table 2). In other words, means of the Functioning and Problem Severity MILC were substantially different at intake and discharge, suggesting a significant change over time, assumedly driven by the IICAPS treatment. Hence, the Partial Invariant model was extended into a multivariate Latent Difference Score (LDS) model (e.g. McArdle, 2009) to capture the latent change between MILC scores at intake and discharge. This model returned an adequate fit to the observed data ($x^2 = 362.23$, df = 42, $p < .001$, x^2/df = 8.63, CFI = 0.989, NFI = 988, RMSEA [CI 90%] = 0.033 [0.030–0.036]).

Figure 2 presents the SEM diagram of this model along with the standardized estimates of the main parameters. The average latent difference scores were statistically different from zero ($p < 0.001$), denoting a statistically significant change over time, with an average increase of 6.91 units for the Functioning MILC score and an average decrease of 8.68 units for the Problem Severity MILC score between intake and discharge. In addition, the variances of the latent change scores were also statistically different from zero ($p < 0.001$), suggesting substantial inter-individual differences in the rate of change under the influence of IICAPS. As illustrated in Figure 2, intercorrelations between change in Functioning and Problem Severity were strong and negative ($r = -0.85$, $p < 0.001$) suggesting that, expectedly, a decrease in Problem Severity between intake and discharge was strongly associated with an increase in Functioning.

Community estimates of the functioning and problem severity MILC levels

The structural means of a sample of youth from the community were modeled in a multi-sample analysis that was used to develop "community benchmarks" for the Functioning and Problem Severity MILC constructs (based on descriptive statistics of Functioning and Problem Severity reports of 166 children, 329 parents and 40 clinicians available in Ogles *et al.* [2004]). Our goal was to estimate "normative" levels of the MILC constructs (since they have not been estimated to

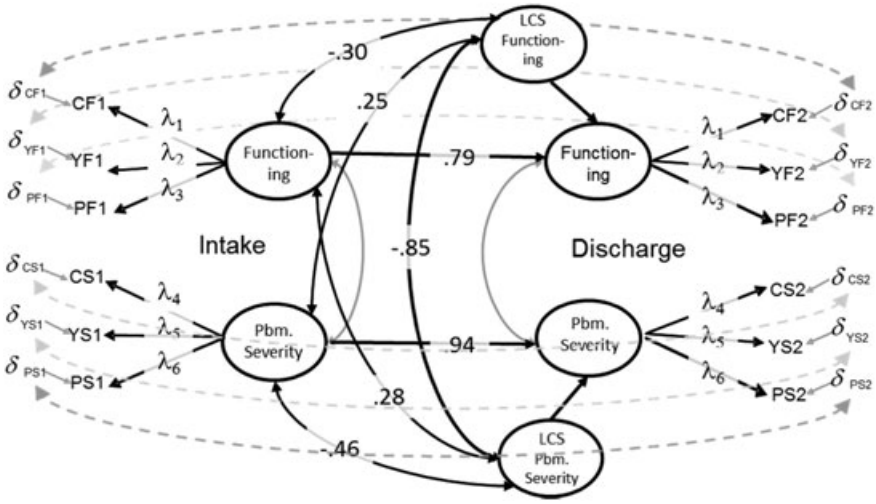

Figure 2. Multivariate Latent Change Score (LCS) model of the Ohio scales Functioning and Problem Severity Multi-informant Latent Consensus (MILC) constructs.

date), providing norm-referenced MILC scores of the IICAPS sample at both intake and discharge, and a complementary appraisal of the effectiveness of the IICAPS intervention.

To do so, we developed a multi-sample, Strict Measurement Invariance model (equal factor loadings, intercept, uniqueness variance and covariance across samples), where variances of the LVs were fixed at one in both samples, and the latent means were fixed at zero in the community sample, making it the reference group (latent means of the IICAPS sample were freely estimated). This model returned an acceptable fit to the data at both intake ($x^2 = 965.86$, df = 30, $p < 0.001$, CFI = 0.933, RMSEA [CI 90%] = 0.067 [0.063–0.070]) and discharge ($x^2 = 918.18$, df = 30, $p < 0.001$, CFI = 0.935, RMSEA [CI 90%] = 0.065 [0.062–0.069]), which suggests that both samples could be compared adequately.

As represented in Figure 3, "norm-referenced" estimates of latent means of the IICAPS sample were as follows: Functioning MILC = −3.04 (standard error [SE] = 0.082, $p < 0.001$) at intake, and −1.82 (SE = 0.075, $p < 0.001$) at discharge; Problem Severity MILC = 1.58 (SE = 0.069 $p < 0.001$) at intake, and 0.637 (SE = 0.067, $p < 0.001$) at discharge. Hence, although the IICAPS sample shows MILC Functioning level considerably below and MILC Problem Severity level noticeably above the community sample at both intake and discharge, these norm-referenced estimates suggest that IICAPS leads to significant mental health improvement (by about 1 SD in both MILC constructs), with mean MILC levels substantially closer to the community sample benchmark at discharge.

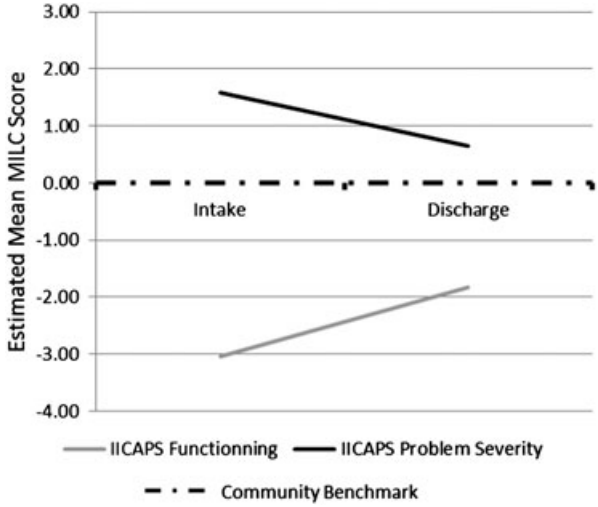

Figure 3. Estimated community benchmark and mean Multi-informant Latent Consensus (MILC) score of Functioning and Problem Severity of the Intensive In-home Child and Adolescent Psychiatric Service (IICAPS) sample at intake and discharge.

Discussion

Results from this study suggest that youth serviced by IICAPS showed levels of Functioning dramatically below (over 3 SDs), and levels of Problem Severity dramatically above (over 1.5 SDs) a community benchmark estimated from data published in the literature. After discharge from the IICAPS program, mean estimates of the youth Functioning and Problem Severity were much closer to the community benchmarks with an average improvement

Int. J. Methods Psychiatr. Res. 25(1): 33–43 (2016). DOI: 10.1002/mpr

of about 1 SD in Functioning, and a reduction of Problem Severity of nearly the same size. Further, results showed a robust dynamic of change under the effect of IICAPS. Overall, Problem Severity and Functioning as estimated by the MILC constructs were only moderately stable over time, with decrease in Problem Severity associated with strong increases in Functioning. However, differences in baseline level were associated with differential amount of change attributed to IICAPS: lower Functioning and higher Problem Severity at intake were related to less reduction of problems at discharge, suggesting a greater stability of mental health issues for the most severe cases.

The evidence of improvement among youth involved in IICAPS is noteworthy for several reasons. First, it suggests that a family and home-based treatment program can improve mental health functioning among youth with SED, which may prevent entry into out-of-home-placements and its associated adverse consequences. Second, it supports the efficacy of treatment models that engage youth, family, and community sectors, tailoring services to the multi-systemic needs of the target child and family. Third, it demonstrates the effectiveness the and generalizability of this program to diverse sectors of society and provides evidence of youth improvement, as demonstrated at the level of a large-scale, state-wide delivered program. Fourth, data provide preliminary support for the effectiveness of community-based approaches, demonstrated here with engagement of school and judicial stakeholders in addition to the partnerships between the child, family, and clinician, to collaboratively promote youth well-being.

In addition to showing evidence of the effectiveness of this large-scale treatment program, the present study established a new approach to the functional assessment of mental health baseline levels (Functioning and Problem Severity) and change, which capitalizes on the information gathered from multiple informant reports. In contrast to most multi-informant studies that have emphasized inter-informant discrepancies (De Los Reyes, 2011), investigated factors associated with each party's "unique perspective" (Carlston and Ogles, 2006; De Los Reyes and Kazdin, 2005; Kanne *et al.*, 2009), or explored which informants are more "valid" in the assessment of mental health outcomes (Smith, 2007), the study presented here has demonstrated that (1) the concordance portion of multiple informants' reports is not trivial and contributes a fair share of each informant perspective, (2) this concordance portion can be captured using the proposed MILC modelling approach, and (3) the resulting MILC constructs are reproducible over time, outlining the robustness of the measured constructs and their suitability for the study of reliable change.

Although the understanding of multi-informant discrepancies constitutes an important area of study that increases our understanding of the causes, consequences and treatments of child and adolescent psychopathology (Lewis *et al.*, 2014; Smith, 2007), it provides few directions for integrating multiple informant reports to derive an accurate evaluation of child and adolescent psychopathology. Therefore, the MILC approach addresses an ongoing debate on the best way to incorporate information from multiple informants (e.g. Achenbach, 2011). Indeed, because there is no "gold standard" to determine the child's true level of dysfunction (due to each informant's own level of subjectivity), there is a need to incorporate information from multiple informants (e.g. Richters, 1992). Contrary to a sum or average score of each informant's report that incorporates both the concordance and the portion of the total score variance related to each informant's unique perception, the MILC approach partitions out the discrepancy and measurement error portion, to capitalize only on the concordance portion. Hence, the MILC approach attempts to derive a consensual, systemic view on youth psychopathology from the perspective of multiple informants, which is considered to be an excellent approximation of the youth's "true" level of psychopathology.

In the study presented here, MILC was applied to the study of change in mental health, in response to the IICAPS intervention. Using a large sample of youth and their families, results showed that the hypothesized MILC model (and underlying inter-informant latent consensus) fit the observed data properly, and could be extended longitudinally with a strong level of invariance over time. That is, the underlying latent consensus between youth, parents and clinicians was structurally comparable over time, underlying the robustness of the measured construct (i.e. extrinsic stability of the Functioning and Problem Severity constructs), while permitting the study of reliable change through LDS modeling. The MILC approach demonstrated promise to integrate information from multiple informants, to provide a more accurate and systemic view of a child's level of Functioning and Problem Severity, to be used as and/or as a basis of subsequent analyses such as the study of change under the effect of interventions such as IICAPS.

In sum, the newly developed MILC approach has confirmed that intensive in-home interventions such as IICAPS lead to a reduction of problem severity and an improvement in functioning in children and adolescents with severe emotional disturbance. The population-based nature of the study sample with its diverse ethnicity and socio-economic background supports the external validity

of the results reported. However, we must acknowledge several limitations of this study. First, the dynamic of change emphasized in our multivariate latent difference model provides an average representation of the effect of IICAPS, as captured by a multi-informant, functional assessment of mental health. Future studies should estimate whether the dynamic and amount of change are similar among multiple sub-groups (e.g. based on initial diagnosis or "main problem", or relevant background variables). Such studies could lead to a better refinement of the IICAPS program, tailored to specific sub-group characteristics. Although results suggest a substantial improvement of mental health at discharge, whether these differences in outcomes can be sustained over the long term should be ultimately be addressed through further evaluation of the IICAPS model's effectiveness across problems and populations. Finally, we should note that community benchmarks for the MILC constructs were based on limited data from the literature and the assumption of identical covariance structure. These limitations deserve additional research to develop more robust benchmarks against which clinical samples could be compared. Regardless, effectiveness analyses of the IICAPS should be (and are currently being) investigated in more controlled designs such as Randomized Controlled Trials design. Hence, despite limitations, this study showed great promise with regard to the IICAPS model as an effective intervention and the MILC approach as a way to integrate meaningfully multi-informant functional assessment of mental health.

Acknowledgments

Preparation of this manuscript was supported by Grant NIHT32 MH18268. The authors thank Ben Ogles for his help with the Ohio scales.

Declaration of interest statement

The authors have no competing interests.

References

Achenbach T.M. (2011) Commentary: definitely more than measurement error: but how should we understand and deal with informant discrepancies? *Journal of Clinical Child and Adolescent Psychology*, **40**(1), 80–86. DOI: 10.1080/15374416.2011.533416.

Achenbach T.M., Krukowski R.A., Dumenci L., Ivanova, M.Y. (2005) Assessment of adult psychopathology: meta-analyses and implications of cross-informant correlations. *Psychological Bulletin*, **131**(3), 361–382.

Achenbach T.M., McConaughy S.H., Howell C.T. (1987) Child/adolescent behavioral and emotional problems: implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, **101**(2), 213–232.

Adnopoz J., Woolston J., Balestracci K.B. (2012) IICAPS: a treatment model for delinquent youths with co-occurring mental health disorders. In Grigorenko E.L. (ed.) Handbook of Juvenile Forensic Psychology and Psychiatry, pp. 357–369, New York, Springer. DOI: 10.1007/978-1-4614-0905-2_23.

Arbuckle J. (2009) Amos 18 user's guide. SPSS Incorporated: Armonk, New York, USA.

Barbot B., Crossman E., Hunter S.R., Grigorenko E.L., Luthar S.S. (2014a) Reciprocal influences between maternal parenting and child adjustment in a high-risk population: a five-year cross-lagged analysis of bidirectional effects. *American Journal of Orthopsychiatry*, **84**(5), 567–580. DOI: 10.1037/ort0000012.

Barbot B., Hein S., Luthar S.S., Grigorenko E.L. (2014b) Capturing age-group differences and developmental change with the BASC parent rating scales. *Journal of Applied Developmental Psychology*, **35**(4), 294–303. DOI: 10.1016/j.appdev.2014.04.003.

Barbot B., Hunter S.R., Grigorenko E.L., Luthar S. S. (2013) Dynamic of change in pathological personality trait dimensions: a latent change analysis among at-risk women. *Journal of Psychopathology and Behavioral Assessment*, **35**(2), 173–185. DOI: 10.1007/s10862-012-9331-4.

Byrne B. (2010) Structural Equation Modeling using AMOS, second edition, New York, Routledge.

Carlston D., Ogles B. (2006) The impact of items and anchors on parent–child reports of problem behavior. *Child and Adolescent Social Work Journal*, **23**(1), 24–37. DOI: 10.1007/s10560-005-0028-3.

Carlston D.L., Ogles B.M. (2009) Age, gender, and ethnicity effects on parent–child discrepancy using identical item measures. *Journal of Child and Family Studies*, **18**(2), 125–135.

Cheung G.W., Rensvold R.B. (2002) Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, **9**(2), 233–255.

De Beuckelaer A., Swinnen G. (2010) Biased latent variable mean comparisons due to measurement non-invariance: a simulation study. In Davidov E., Schmidt P., Billiet J. (eds) *Methods for Cross-cultural Analysis: Basic Strategies and Applications*, European Association of Methodology Series Edition, pp. 119–149, London, Taylor & Francis Group.

De Los Reyes A. (2011) Introduction to the special section: More than measurement error: Discovering meaning behind informant discrepancies in clinical assessments of children and adolescents. *Journal of Clinical Child & Adolescent Psychology*, **40**(1), 1–9. DOI: 10.1080/15374416.2011.533405.

De Los Reyes A., Kazdin A. (2005) Informant discrepancies in the assessment of childhood psychopathology: a critical review, theoretical framework, and recommendations for further study. *Psychological Bulletin*, **131**(4), 483–509.

Dunlop B.W., Li T., Kornstein S.G., Friedman E.S., Rothschild A.J., Pedersen R., Ninan P., Keller M., Trivedi M.H. (2011) Concordance between clinician and patient ratings as predictors of response, remission, and recurrence in major depressive disorder. *Journal of Psychiatric Research*, **45**(1), 96–103. DOI: 10.1016/j.jpsychires.2010.04.032.

Epkins C.C. (1996) Parent ratings of children's depression, anxiety, and aggression: a cross-sample analysis of agreement and differences with child and teacher ratings. *Journal of Clinical Psychology*, **52**(6), 599–608.

Gowers S., Weetman J., Shore A., Hossain F., Elvins R. (2000) Impact of hospitalisation on the outcome of adolescent anorexia nervosa. *The British Journal of Psychiatry*, **176**(2), 138–141.

Grigorenko E.L., Geiser C., Slobodskaya H.R., Francis D.J. (2010) Cross-informant symptoms from CBCL, TRF, and YSR: trait and method variance in a normative sample of Russian youths. *Psychological Assessment*, **22**, 893–911.

Harpin S., Kenyon D.B., Kools S., Bearinger L.H., Ireland M. (2013) Correlates of emotional distress in out-of-home youth. *Journal of Child and Adolescent Psychiatric Nursing*, **26**(2), 110–118. DOI: 10.1111/jcap.12030.

Haynes C., Eivors A., Crossley J. (2011) Living in an alternative reality: adolescents' experiences of psychiatric inpatient care. *Child and Adolescent Mental Health*, **16**(3), 150–157. DOI: 10.1111/j.1475-3588.2011.00598.x.

Ivens C., Rehm L.P. (1988) Assessment of childhood depression: correspondence between reports by child, mother, and father. *Journal of the American Academy of Child & Adolescent Psychiatry*, **27**(6), 738–741. DOI: 10.1097/00004583-198811000-00012.

Kanne S.M., Abbacchi A.M., Constantino J.N. (2009) Multi-informant ratings of psychiatric symptom severity in children with autism spectrum disorders: the importance of environmental context. *Journal of Autism and Developmental Disorders*, **39**(6), 856–864.

Kline R.B. (2010) Principles and Practice of Structural Equation Modeling, third edition, New York, The Guilford Press.

Lewis A.J., Bertino M.D., Bailey C.M., Skewes J., Lubman D.I., Toumbourou J.W. (2014) Depression and suicidal behavior in adolescents: a multi-informant and multi-methods approach to diagnostic classification. *Frontiers in Psychology*, **5**, 1–9.

McArdle J. (2009) Latent variable modeling of differences and changes with longitudinal data. *Annual Review of Psychology*, **60**(1), 577–605.

McArdle J., Nesselroade J.R. (1994) Using multivariate data to structure developmental change. In Cohen S., Reese H. (eds) Life-span Developmental Psychology: Methodological Contributions, pp. 223–267, Hillsdale, NJ: Lawrence Erlbaum Associates.

Meredith W., Horn J. (2001) The role of factorial invariance in modeling growth and change. In Linda M., Sayer A.G. (eds) New Methods for the Analysis of Change, pp. 203–240, Washington, DC: American Psychological Association. DOI: 10.1037/10409-007.

Offord D.R., Boyle M.H., Racine Y., Szatmari P., Fleming J.E., Sanford M., Lipman E.L. (1996) Integrating assessment data from multiple informants. *Journal of the American Academy of Child & Adolescent Psychiatry*, **35**(8), 1078–1085.

Ogles B.M., Dowell K., Hatfield D., Melendez G., Carlston D.L. (2004) The Ohio scales. In Maruish M. (ed.) The Use of Psychological Testing for Treatment Planning and Outcomes Assessment, pp. 275–304, Mahwah, NJ: Lawrence Erlbaum Associates.

Ogles B.M., Melendez G., Davis D.C., Lunnen K.M. (2001) The Ohio scales: practical outcome assessment. *Journal of Child and Family Studies*, **10**(2), 199–212.

Ohio Department of Mental Health. (2003) Ohio Mental Health Consumer Outcome System: Initial Statewide System, Columbus, OH: Department of Mental Health.

Richters J.E. (1992) Depressed mothers as informants about their children: a critical review of the evidence for distortion. *Psychological Bulletin*, **112**(3), 485–499.

Schafer J.L., Graham J.W. (2002) Missing data: our view of the state of the art. *Psychological Methods*, **7**(2), 147.

Smith S.R. (2007) Making sense of multiple informants in child and adolescent psychopathology: a guide for clinicians. *Journal of Psychoeducational Assessment*, **25**(2), 139–149. DOI: 10.1177/0734282906296233.

Treutler C.M., Epkins C.C. (2003) Are discrepancies among child, mother, and father reports on children's behavior related to parents' psychological symptoms and aspects of parent–child relationships? *Journal of Abnormal Child Psychology*, **31**(1), 13–27.

Turchik J.A., Karpenko V., Ogles B.M. (2007) Further evidence of the utility and validity of a measure of outcome for children and adolescents. *Journal of Emotional and Behavioral Disorders*, **15**(2), 119–128.

Woolston J.L., Adnopoz J.A., Berkowitz S.J. (2007) IICAPS: A Home-based Psychiatric Treatment for Children and Adolescents, New Haven, CT: Yale University Press.

Youngstrom E., Loeber R., Stouthamer-Loeber M. (2000) Patterns and correlates of agreement between parent, teacher, and male adolescent ratings of externalizing and internalizing problems. *Journal of Consulting and Clinical Psychology*, **68**(6), 1038–1050. DOI: 10.1037/0022-006X.68.6.1038.