Family feedback in Child Welfare Services: A systematic review of measures

Lara Ayala-Nunes[a,b*], Lucía Jiménez[b], Mª Victoria Hidalgo[b], & Saul Jesus[a]

[a]Research Centre for Spatial and Organizational Dynamics, University of Algarve. Campus de Gambelas, s/n., 8005-139 Faro (Portugal)

[b]Department of Developmental and Educational Psychology, University of Seville. C/ Camilo José Cela, s/n, 41018 Seville (Spain).

Authors' e-mail address:

Lara Ayala-Nunes: layala@us.es

Lucía Jiménez: luciajimenez@us.es

Victoria Hidalgo: victoria@us.es

Saul Jesus: snjesus@ualg.pt

[*] Corresponding author. Postal address: Departamento de Psicología Evolutiva y de la Educación, Universidad de Sevilla. C/ Camilo José Cela, s/n, 41018 Sevilla (Spain). Telephone number: +34 692 922 070

Family feedback in Child Welfare Services: A systematic review of measures

Abstract

Background: Assessing family feedback in Child Welfare Services is embedded in family-centered practice, and the availability of validated, reliable instruments to perform this evaluation is essential for front-line practitioners, managers, and policymakers. Nonetheless, to date, no study has systematically identified and analyzed the measures in this field.

Objectives: A systematic review of the literature following the PRISMA guidelines was conducted to identify, describe and conceptually and psychometrically assess all of the published measures of family feedback in Child Welfare Services.

Method: A search of the measures published in peer-reviewed scientific journals in English, Spanish, Portuguese, Italian, and French from 1980 to October 2013 was performed. Ten electronic databases and reference lists of relevant studies were consulted. In total, 13 studies including eight instruments were identified and analyzed.

Results: Most studies omitted information regarding the descriptive characteristics of the instruments and made no reference to a conceptual model. In most cases, the development and validation processes of the instruments and their psychometric characteristics were insufficiently reported. Additionally, some relevant elements of family-centered practice were frequently omitted in the dimensions of the questionnaires.

Conclusions: The scarcity of validated measures to evaluate family feedback in Child Welfare Services demands further research to develop new instruments that overcome these limitations. Recommendations for designing and validating future instruments are provided.

Keywords: Family feedback, user satisfaction, child welfare, family preservation, systematic review, instruments.

1. Introduction

In the last three decades, child welfare systems in most Western countries have evolved from a traditionally investigative and deficit-focused approach towards a more family-centered and strengths-based perspective (Connolly, 2007). This shift in social work practice has translated into the recognition of the family as a critical context in a child's life as well as the need to provide families with adequate support. Such practice is embodied in family preservation services, which aim to prevent children's out-of-home placement and to ensure children's well being within their families of origin. From a family preservation approach, the

notions of user consultation, user involvement, user input and user-driven services have a central role (Kelly & Blythe, 2000). In this line, studies about service user feedback may serve as a means to empower parents and to give them a chance to have a voice with regard to their experiences with services. The process of being asked about their own opinions may actually change their perceptions about the staff and the services; it may also help to reduce the pronounced power asymmetry that exists between practitioners and child welfare users, and to provide client-centered and family-focused casework practice (Alpert, 2005; Baker, 2007; Tilbury, Osmond, & Crawford, 2010). Additionally, systematically assessing families' opinions about services is consistent with the principle of turning users into active agents of the intervention and promoting their autonomy (Rodrigo, Maiquez, Martín, & Byrne, 2008). All of these aspects lead us to consider that family feedback encompasses more than just user satisfaction; it is a broader concept that includes all the perceptions and opinions that users hold about services (their characteristics and effectiveness), practitioners, and outcomes.

The importance of assessing family feedback with Child Welfare Services (CWS) has been widely recognized (American Humane Association, 1998; Baker, 2007; Cortis, 2007; Kapp & Vela, 1999; Lietz, 2009; Tilbury et al., 2010). Parents involved with child welfare have unique perspectives and interpretations about the processes, events, and decisions that occur in those services. Such input is helpful in assessing program efficacy, designing service delivery improvement strategies and enhancing professional practice, which in turn may increase parental engagement with interventions (Tilbury et al., 2010). Within this field, client or user satisfaction has been the most studied dimension. The American Humane Association (1998) considers family satisfaction to be one of the outcomes that should be assessed in CWS, as it improves accountability to families. Furthermore, user satisfaction has been associated to positive outcomes in child welfare, such as practitioner estimates of client progress, fewer further notifications and children remaining at home (Trotter, 2008) as well as a greater likelihood of service completion (Damashek, Doughty, Ware, & Silovsky, 2011). Additionally, the parent-worker relationship, which is an essential element of user satisfaction with CWS, has been found to be a consistent predictor of intermediate outcomes (Marsh, Angell, Andrews, & Curry, 2012). Specifically, the strength of the parent-worker relationship predicted service completion (Girvin, DePanfilis, & Daining, 2007); staff perception of family involvement (Korfmacher, Green, Spellmann, & Thornburg, 2007), child and family well being (Johnson & Ketring, 2006; Johnson, Wright, & Ketring, 2002; Southerland, Mustillo, Farmer, Stambaugh, & Murray, 2009), and improvements in child safety (Johnson & Ketring, 2006; Lee & Ayón, 2004). The perception of services may also affect child and family outcomes indirectly. For instance, satisfaction has been associated with program completion, collaboration with practitioners,

engagement and compliance in interventions. Compliance, in turn, predicts reductions in the likelihood of subsequent reports of child maltreatment and out-of-home placements (Littell, 2001), and engagement is positively associated with parents' perceptions that their children were safer because of their involvement with services and that their parenting had improved (Gladstone et al., 2012). It is likely that parents who feel better about the services they receive and the relationships established within those services will be more receptive to suggestions, referrals, and assistance offered by child welfare agencies (Alpert, 2005; Chapman, Gibbons, Barth, McCrae & the NSCAW Research Group, 2003).

In spite of the importance of measuring such aspects, there is still a lack of information about family feedback on child welfare systems, and family participation in performance measurement is rare (Baker, 2007; Cortis, 2007; Tilbury et al., 2010). Additionally, most studies have focused solely on satisfaction with services, ignoring other relevant aspects of the experience with CWS. Kapp and Vela (1999) attribute the underdevelopment of this area to several reasons: public social service agencies generally are not economically supported by users, and therefore, some of them may lack the incentive to measure user feedback; most of the families involved with CWS are involuntary users, and funding or performance measurement is typically based on outcomes (e.g., number of children in out-of-home placement) as measures of service effectiveness, which are partly independent of user perceptions. Lastly, an undervaluation of user opinions, considering them unreliable or biased may also be responsible for the lack of emphasis on family feedback in child welfare (Russell, 1990).

An important segment of the research in this field has been criticized for lacking methodological rigor (Heneghan, Horwitz, & Leventhal, 1996). The methodological weaknesses of the research on child welfare program evaluation may be partly attributable to the absence of well-established, sound, reliable and valid measures to assess family feedback (Baker, 2007; Berrick, Frasch, & Fox, 2000; Harris & Poertner, 1998). Most of the instruments lack psychometric analysis or have not been adequately tested (Harris & Poertner, 1998), which may mean that studies may not be measuring user satisfaction well or not even be measuring the factors they believe they are measuring (Young, Nicholson, & Davis, 1995). On one hand, the majority of the studies tend to use ad hoc questionnaires without a conceptual framework that can only be used for a specific service or program or offer insufficient information about instrument development, validation or reliability (Kapp & Vela, 1999). On the other hand, many studies have assessed user feedback through qualitative methods, such as open interviews or focus groups (e.g., Buckley, Carr, & Whelan, 2011; Cortis, 2007). These methods may offer a rich and valuable vision of participants' perceptions, but a comparison of the

results is not feasible. This hampers the collection of user feedback data as well as the possibility of comparing the findings over time and across different programs and services; it also diminishes the impact that such data could have in improving organizations or changing social policies (Baker, 2007). In addition, there is a challenge in identifying core constructs that should be included in family feedback measures, and the construct itself has yet to be consensually defined. The lack of a conceptual model that frames the empirical findings is common in user satisfaction studies (Pascoe, 1983). On a research level, more user feedback studies are warranted to conceptualize and identify its main components and contribute to theory development in this area (Baker, 2007).

Taking into account the aforementioned reasons, it seems that having validated, reliable instruments available to assess the perceptions of users of CWS is key for front-line practitioners, managers and policy makers. Nonetheless, to the best of our knowledge, no attempt has been made so far to systematically identify the existing instruments in this field or to assess what is being measured and how. We should mention two previous efforts as exceptions: the reviews by Harris and Poertner (1998) and Kapp and Vela (1999). In the first review, the authors examined the measurement of client satisfaction in several human service agencies and examined instruments in terms of their ability to reflect clients' experiences with services, the dimensions of satisfaction they measured and the quality of the data. They managed to identify four client satisfaction instruments in child welfare/protection services. Kapp and Vela (1999) reviewed the instruments used to measure consumer satisfaction in family preservation services and other related areas and found nine instruments. In spite of the usefulness of these reviews, neither of them followed an explicitly systematic review procedure, and more than a decade has passed since. To overcome these gaps, a systematic review of the extant literature was performed, following the PRISMA guidelines for conducting and reporting systematic reviews and meta-analyses (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009). The aims of this systematic review are:

1) To identify and to characterize all of the published measures of family feedback with CWS that are suitable for research and service evaluation purposes;

2) To assess the conceptual framework and psychometric features of these measures;

3) To offer guidelines for the construction of new instruments that might overcome the current limitations.

2. Methods

2.1. Search and eligibility criteria

A search for studies that included or referenced instruments assessing family feedback with CWS was completed. The following inclusion criteria were considered: (a) The target population included the measures designed for caregivers of families at psychosocial risk whose children had not been placed in out-of-home care. The instruments assessing satisfaction with other services (e.g., mental health services) or developed with other populations (e.g., foster parents; parents with children in out-of-home care) were excluded; (b) The dates ranged from 1980 to October 2013. Articles published prior to 1980 were excluded because of the changes in policy, organizations and structure that CWS have undergone in the last three decades; (c) The types of sources and languages included peer-reviewed articles published in scientific journals in English, Spanish, Portuguese, Italian and French.

Two search methods were followed to identify the studies: An Internet-based search of the literature and a scan of the reference lists of articles that were found in the previous search and deemed relevant. The following electronic databases were searched: (i) PsycInfo, (ii) MedLine, (iii) PsycArticles, (iv) ProQuest Psychology Journals, (v) Social Services Abstracts, (vi) FRANCIS, (vii) ERIC, (viii) Web of Science, (ix) OVID and (x) Psychology and Behavioral Sciences Collection. To identify subject-related research terms, the authors performed a preliminary literature search and consulted child welfare experts. Based on these findings, an iterative search on each electronic database was carried out by matching two sets of terms, including both truncated and thesaurus terms. The truncated search strategy is displayed in Figure 1.

> su(social services OR child social services OR child welfare OR child welfare services OR child protect* services OR child protective agenc* OR family preservation services OR family preservation) AND su(satisfaction OR client satisfaction OR parent* satisfaction OR mother* satisfaction)

Fig. 1. Truncated search strategy

An example of the thesaurus search for PsycInfo is presented in Figure 2.

> SU.EXACT("Dissatisfaction") OR SU.EXACT("Needs") OR SU.EXACT("Psychological Needs") OR SU.EXACT("Clients") OR SU.EXACT("Quality of Services") OR SU.EXACT("Consumer Satisfaction") OR SU.EXACT("Client Satisfaction") **AND** SU.EXACT("Family Preservation") OR SU.EXACT("Protective Services") OR SU.EXACT("Child Abuse") OR SU.EXACT("Social Services") OR SU.EXACT("Child Welfare") OR SU.EXACT("Child Neglect")

Fig. 2. Example of thesaurus search terms for PsycInfo

2.2. Selection procedure and study characteristics

Records were sorted by relevance, and duplicates were removed. In the cases in which the search retrieved more than 100 records, a filter was applied to exclude studies with unrelated subjects (such as job satisfaction, mental health or substance-related disorders). A preliminary study selection was performed by the first author. The relevance of the studies was determined through the screening of the titles and/or the abstracts (if the study was not in the relevant subject area, it was excluded on the basis of the title) of the retrieved records. After this selection, the studies that met the eligibility criteria were evaluated in further detail. Studies that made reference to or used at least one questionnaire on family feedback on child welfare or family preservation services or programs were included in the review. For data collection, a data extraction sheet was developed and adjusted after testing it with the first selected study (obtainable from the corresponding author). The first author extracted data from all of the selected studies, and the second author checked and confirmed the accuracy of the extracted data. In the case of disagreement between the first and second authors, the opinion of the third and fourth authors was solicited.

The search of the aforementioned databases provided a total of 17,954 records[1] (see Figure 3). After adjusting for unrelated subjects, dates, types of publications and publication status, 3,839 records remained. Of these, 3,791 were discarded because after reviewing their abstracts, it was concluded that those studies did not meet the criteria for inclusion. The full text versions of the remaining 48 studies were examined, and 37 studies did not meet the inclusion criteria as described. Ten studies met the inclusion criteria and were included in the review, and three relevant studies were identified by checking the reference lists of the studies that met the inclusion criteria. Finally, a total of 13 studies including eight instruments assessing family feedback on CWS and family preservation programs were identified and analyzed for this review.

---

[1] This number was obtained by summing all of the records retrieved for each database. Because the research was performed in 10 electronic databases separately, it is not possible to calculate the exact number of total records. In some cases, the same records were retrieved in more than one database.
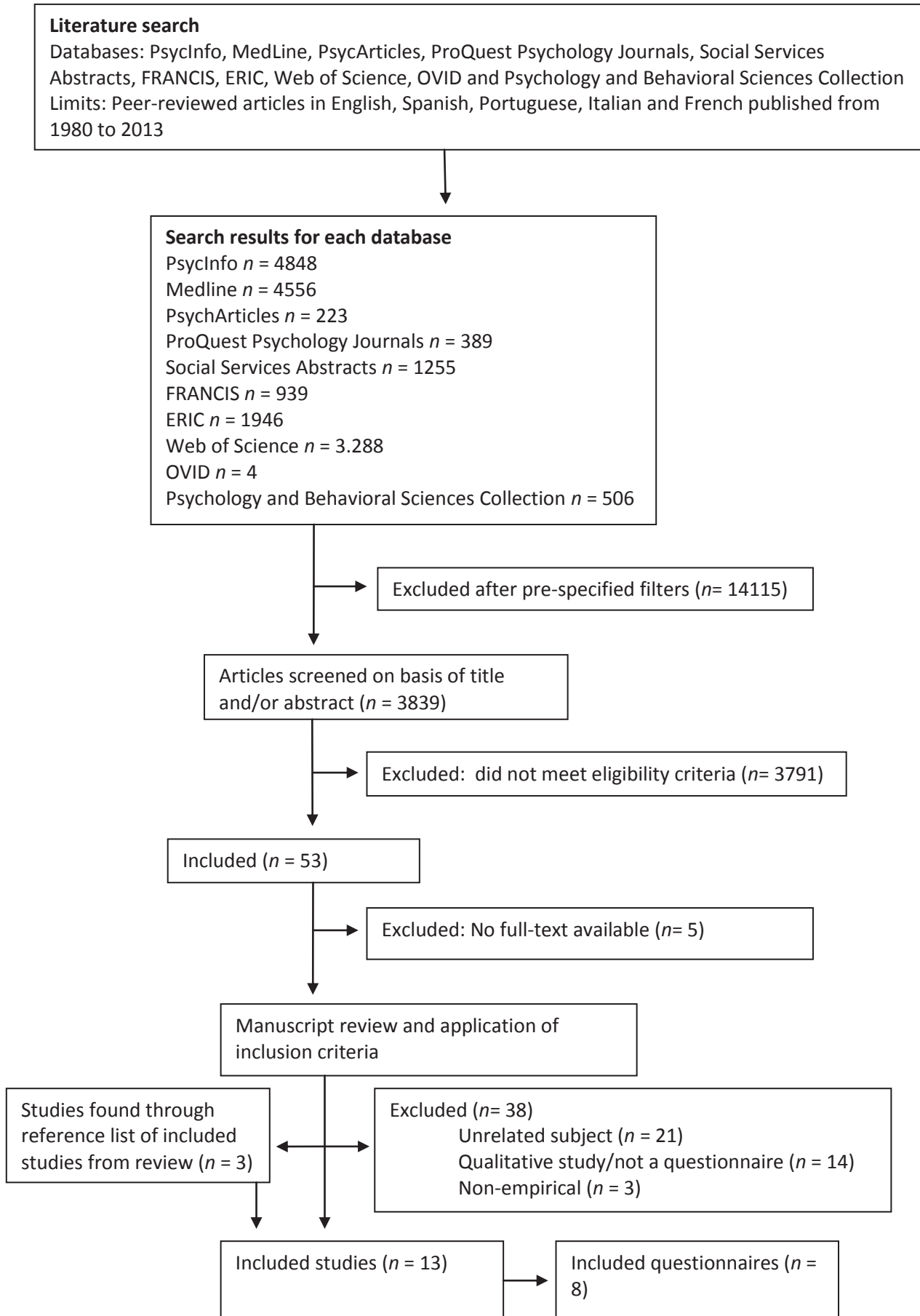
Fig. 3. Systematic review procedure.

2.3. Data analysis

Information was extracted from each selected instrument on: 1) questionnaire subscales, 2) the number of items, 3) number of response options, 4) administration length, and 5) reading level. The following information about the sample to which the instrument was administered is provided: 6) size, 7) country, 8) program or service in which participants were engaged. Measures were also characterized regarding their theoretical and their psychometric strengths and weaknesses, according to Carretero-Dios and Pérez's (2005) guidelines, in the following aspects: 1) a basis on a theoretical model, 2) the inclusion of the definition of the construct, 3) a content validity analysis, 4) a statistical analysis of the items, 5) a dimensionality analysis, 6) a reliability estimation, and 7) evidence of external validity (criterion, convergent and discriminant).

3. Results

3.1. Descriptive characteristics

The characteristics of the questionnaires included in the review are displayed in Table 1. Two studies did not report instrument subscales, and one did not report the number of items. The number of subscales ranged from two to five, and they made reference to dimensions such as the quality of the program, the satisfaction with the service and satisfaction with staff. Only three of them included outcomes as a dimension, and only the Strengths-Based Practices Inventory included specific family preservation-related dimensions (e. g., empowerment). The instruments had an average of 15 items (range = 6-27).

In all cases, the range of response options was reported, and in five cases, examples of responses were included. On average, the instruments had five response options (range = 4-7). The administration length was only specified in two studies (maximum = 10 minutes), and the required reading level to complete the questionnaire was only indicated by three studies (range = 4.9 – 6th grade). Sample sizes ranged from 24 to 4,168. The vast majority of the instruments (88%) were developed with populations from the USA. All participants were engaged with child welfare/protection agencies, family preservation services and/or structured parenting programs (e.g., Early Head Start).

Table 1.

Characteristics of the questionnaires and the samples.

	Client Satisfaction Survey (ClSSu)	Strengths-Based Practices Inventory (SBPI)	Client Satisfaction Inventory (CSI)	Customer Satisfaction Survey (CuSS)	Current Client Satisfaction with Agency Staff (CCSAS)	Parent Satisfaction Questionnaire (PSQ)	Client Satisfaction Scale (CISS)	CPS Father Survey (CPSFS)
Authors and date	Chaffin, Bard, Bigfoot, & Maher (2012)	Green, McAllister, & Tarte (2004)	McMurtry & Hudson (2000)	Huebner, Jones, Miller, Custer, & Critchfield (2006)	Winefield & Barlow (1995)	Reid, Webster-Stratton, & Beauchaine (2001)	Damashek, Bard, & Hecht (2012)	Huebner, Werner, Hartwig, White, & Shewa (2008)
Subscales	- Quality of the program - Satisfaction with services - Realized benefits from the services	- Empowerment approach - Cultural competence - Staff sensitivity-knowledge - Relationship-supportive	Unidimensional	Unidimensional	NR	- Overall satisfaction - Program usefulness - Leader satisfaction - Techniques' ease -Techniques' usefulness	- Service satisfaction - Personal improvement	NR
Number of items	27	16	9	11	6	NR	18	19
Number of response options	4 (0 = not at all; 3 = a lot) and 5-point scale (0 =	7 (1 = very strongly disagree; 7 = very strongly	7 (1 = none of the time; 7 = all of the time) and	5 (1 = strongly disagree; 5 = strongly agree)	4 (1 = strongly; 4 = strongly agree)	7	4	5

	did not receive; 4 = mostly helpful)	agree)	X (does not apply)					
Administration length	NR	NR	5 minutes or less	Less than 10 min	NR	NR	NR	NR
Reading level	NR	NR	5th grade	At or below the 6th grade	NR	NR	NR	4.9 grade
Sample size	354	Study 1: 275 Study 2: 68	329	4168	24	634	1.305	339 fathers, 1.203 social service workers
Country	USA	USA	USA	USA	Australia	USA	USA	USA
Program or service	SafeCare home-based model	Early Head Start program	Human service agencies	Comprehensive family services	Child protection agency	Incredible Years Program	SafeCare+	Child protective services

Note. NR = Not reported.

3.2. Theoretical and psychometric characteristics

Concerning the theoretical and psychometric quality of the analyzed instruments (Table 2), the vast majority of the instruments (88%) were atheoretical (i.e., not explicitly based on a theoretical model). Only the SBPI was based on a conceptual model, and only the CSI attempted to define the construct of interest. The conceptual model for the SBPI was based on two sources: a review of the literature and focus groups with parents involved with a strengths-based family support program. Through the review of the literature, the authors identified ten practices that a strengths-based approach to family services would include: an empowering orientation, cultural competence, a relationship-based approach, family strengthening, active partnering between family members and program staff, a community orientation, knowledge of community-based providers, a family-centered approach, a goal-oriented approach, and the individualization of services to address specific family needs. In the focus groups, parents were asked what made the strengths-based approach work. In the resulting model, three pathways of influence were underlined: boosting parents' motivation to engage and participate in program services, improving parents' sense of competence and empowerment and investing in relationships with parents, and building community relations. Regarding the CSI, although the authors did not make reference to any conceptual model, they did attempt to define satisfaction, recognizing that there is still no consensus regarding its definition. They adopted the position of Young, Nicholson and Davis (1995), considering a generalist and contextualist perspective of satisfaction. Unlike most studies, McMurtry and Hudson (2000) devoted a few lines to reflecting about the obstacles to theoretical work in the satisfaction field, such as ad hoc questionnaires and a body of research based on instruments with unknown psychometric qualities.

Table 2.

The theoretical and psychometric quality of the questionnaires.

	ClSSu	SBPI	CSI	CuSS	CCSAS	PSQ	ClSS	CPSFS
Basis on a theoretical model	NR	Yes	NR	NR	NR	NR	NR	NR
Inclusion of construct definition	No	No	Yes	No	No	No	No	No
Content validity analysis	NR	Yes	NR	Yes (NEIP)	NR	NR	NR	Yes (NEIP)
Statistical analysis of items	NR	Yes	Yes	NR	NR	NR	NR	Yes (NEIP)
Dimensionality analysis	Yes (EFA and CFA)	Yes (CFA)	NR	Yes (EFA) (NEIP)	NR	NR	Yes (EFA and CFA)	NR
Reliability estimation	α = .94	EA α = .92 CC α = .72 SSK α = .81 RS α = .82	α = .89. SEM = 4.11	α = .90 - .94 (NS)	α = .69 for the first worker α =.87 for the second worker	OS α = .57 PU α =.95 LS α =.80 TE: α =.92 TU α = .93	α = .95	α = .95
EVIDENCE OF EXTERNAL VALIDITY — **Criterion**	NR	Yes	NR	NR	NR	NR	NR	NR
EVIDENCE OF EXTERNAL VALIDITY — **Convergent**	NR	Yes	NR	NR	NR	NR	NR	NR
EVIDENCE OF EXTERNAL VALIDITY — **Discriminant**	NR	NR	NR	NR	NR	NR	NR	NR

Note. NR = Not reported NEIP = Not enough information provided EFA = Exploratory factor analysis CFA = Confirmatory factor analysis NS = Not specified EA = Empowerment approach CC = Cultural competency SSK = Staff sensitivity-knowledge RS = Relationship-supportive OS = Overall satisfaction PU = Program usefulness LS = Leader satisfaction TE = Techniques' ease TU = Techniques' usefulness SEM = Standard error of measurement

In most cases (63%), little or no information about the development and validation of the instruments was offered. It is important to note that only three of the included studies (27%) were aimed at instrument development or validation, while the rest of the studies used ad hoc measures aimed at assessing specific interventions, services or programs.

Regarding content validity, the SBPI, the CuSS and CPSFS provided information about the procedures followed to ensure item content validity. However, the information in the latter two was rather vague and unspecific for item construction and selection (e.g., "Survey items were generated from existing literature and through focus groups with child protection workers, policy specialists, field supervisors, university faculty, and several fathers"). For the SBPI, a list of items that reflected strengths-based practice principles was developed and consequently refined by a team of researchers and program staff. Five items were generated for each of the 10 principles (previously identified through a literature search).

Statistical analysis of items was reported in three cases (38%) - the CPSFS, the SBPI and the CSI. The first one briefly mentioned the survey's initial form piloting, whereas in the SBPI, authors referred to dropping items that were reported by respondents as being unclear. The subscales of this instrument were computed on the basis of the originally hypothesized factor structure (10 principles of strengths-based practice) and reduced to seven subscales. Reliability analyses were conducted, and items that did not fit well were dropped or combined with other factors. However, the observed asymmetry and Kurtosis superior to two in some subscales were not discussed by the authors. The authors of the CSI conducted corrected item-total correlations and observed moderate to large scores. No information pertaining to the variance of neither items' scores nor alpha if an item was deleted was provided. None of the analyzed studies reported the criteria followed to delete problematic items after the initial administration of the instrument (pilot). Those that made reference to the questionnaire piloting (e.g., SBPI, CPSFS) only mentioned revisions through iterations and field-testing but did not provide any criteria for dropping items.

Concerning dimensionality analysis, four of the eight instruments (50%) were tested for internal structure. In one case (CuSS), the authors performed an exploratory factor analysis but did not provide any information regarding the model's robustness. For the rest of the instruments, a confirmatory approach was followed (with or without a previous exploratory analysis), showing psychometrical robustness in terms of item retention and significant correlations between the resulting factors. However, the goodness of fit of the final model was only reported in two cases (the SBPI and the ClSSu).

The reliability estimation of the instruments was reported using Cronbach's alpha in all cases with acceptable to excellent results (George & Mallery, 2003), with the exception of the

PSQ (lowest α = .57). For the CSI, the standard error of measurement was used as an additional alternative method to estimate reliability.

All studies failed to report evidence of any facet of external validity, except for the SBPI and the CSI study. Information about criterion and convergent validity for the SBPI was offered. Regarding its criterion validity, regression analyses were performed to assess whether the extent of the strengths-based practices of the institutions predicted empowerment, social support or parental outcomes ten months later. None of the analyses indicated that the SBPI score predicted an improvement in such outcomes. Hence, its criterion validity may be compromised. In contrast, the SBPI's convergent validity seemed to be acceptable: it positively correlated with families' level of engagement in services, the frequency of services, support satisfaction, empowerment, efficacy-related variables, parental competency and quality of the home. As for the CSI, the authors only focused on discriminant validity. The mean correlation between its items and other scales that measured unrelated constructs (Generalized Contentment Scale, Index of Self-Esteem and Index of Peer Relations, as well as clients' background variables) was negligible, as hypothesized by the authors.

4. Discussion and conclusions

4.1. Summary of evidence

This review has identified eight user feedback questionnaires in CWS that were appropriate for research and service evaluation purposes. This scarcity of measures is somewhat at odds with the shift in policies towards family-inclusive practices in child welfare that has occurred over the last decades. Most instruments omitted information regarding their characteristics. This finding replicates those of Kapp and Vela's (1999) and Harris and Poertner's (1998) reviews. For instance, few instruments reported administration length or reading levels. An underrepresentation of the instruments developed outside the USA was also found, and therefore, the use of these questionnaires may not be entirely appropriate in other countries. It is important to note that most studies from which the instruments were drawn were not aimed at developing or validating such instruments. Instead, these studies utilized user feedback questionnaires to evaluate service or program effectiveness, so it is understandable that much of the information pertaining to the psychometric characteristics of the instruments was omitted or unexplored. Nonetheless, this fact mirrors the extended practice of creating and using ad hoc, non-validated questionnaires, which cannot be generalized to other contexts.

Regarding the dimensions included in the questionnaires, most instruments reflected a somewhat narrow view of feedback and did not incorporate elements that are central to

14

family-centered practice, such as a strengths- (e.g., did the intervention focus on the family's strengths?) and empowerment-based approach (e.g., has the intervention improved the family's autonomy and their ability to solve their problems without professional help?). It is also noticeable that none of them have included participants' previous expectations about services in spite of the importance of this construct (Pascoe, 1983). Assessing prior expectations about the effectiveness and helpfulness of the services is essential to understanding the causes of their (dis)satisfaction with the actual services received. This is particularly the case for non-voluntary users in disadvantaged situations, such as child welfare users. For instance, a person's degree of satisfaction with a service or a program may not be the result of its quality, but of the person's low expectations. Harris and Poertner (1998) also found that this dimension was infrequently included in child welfare client satisfaction questionnaires. Likewise, most instruments had little if any theoretical basis: the definition of the construct that they were attempting to measure was only provided in one study, and we did not observe any discussions about the potential factors or the interaction of elements that shape participants' perceptions of CWS with the exception of the SBPI. Therefore, this field still has significant conceptual gaps that need to be addressed to develop solid assessment tools.

Concerning the methodological characteristics, a significant lack of information regarding instruments' content validity was found, precluding us to determine whether the efforts to guarantee content validity were not pursued or simply not reported. This is applicable for the rest of psychometric properties, with the exception of internal consistency. In either case, the vast majority of studies did not report how the items were generated or which procedures were followed to select them (e. g., experts' panel, focus groups, or the Delphi technique). Information regarding the piloting of the instrument and the criteria to delete problematic items was also insufficient in most cases. Additionally, only one study included the descriptive statistics of the items. A dimensionality analysis was performed for half of the analyzed instruments, with only two of them reporting information about the goodness of fit of the model. A confirmatory approach was followed in three instruments, with psychometrical robustness in item retention and significant correlations between the resulting factors. In contrast, internal reliability was tested and reported for all instruments, using Cronbach's alphas with generally acceptable values. The same was not observed for external validity: only two studies reported at least one aspect of the external validity of the instruments, and none of them evaluated all three facets.

In sum, if we consider the available information about the instruments analyzed in this review, we may conclude that the vast majority have considerable weaknesses or at least that there is a great deal of uncertainty about their conceptual and psychometric features.

4.2. Limitations

This systematic review has several limitations. First, full texts were not available from the consulted databases for some of the retrieved studies (see Fig. 3), and some books found in reference lists were unavailable. Second, although attempts were made to achieve an exhaustive search, it is possible that a relevant search term was not used and therefore that relevant studies were not retrieved. Additionally, in spite of authors' thorough scans of the retrieved results, it is possible that some studies were overlooked. Additionally, no attempt to assess publication bias was made. These limitations, which are common to the majority of systematic reviews, explain the main weakness of this study: the reduced number of instruments found. Nonetheless, to the best of our knowledge, this review is the first to systematically identify and assess family feedback questionnaires on CWS.

4.3. Recommendations for the development of new instruments and conclusions

We analyzed the instruments that could be used to assess family feedback on CWS and found several instruments that may be useful for this purpose. Nevertheless, we did not find any instrument that had satisfactory conceptual and psychometric characteristics or at least any instrument that reported sufficiently these features. Despite the current efforts made in program and service evaluations to include the perspectives of users, there remains a lack of quantitative and validated instruments. Further research is needed to develop new instruments that overcome the aforementioned limitations while using the strengths of the current instruments. Based on our findings and on Carretero-Dios and Pérez's (2005) guidelines for developing and validating scales, we suggest the following recommendations for future questionnaires:

1) When developing the questionnaire, a theoretical model that attempts to explain the determinants of service evaluations should be considered and a definition the construct of interest must be provided or at least discussed;

2) Evidence-based, relevant dimensions should be included, such as participants' previous expectations of services; the quality of the service or program; practitioners'' competencies and/or the quality of the user-provider relationship; service effectiveness in terms of empowerment and improving family and child well being; satisfaction with the characteristics of the service or program (e.g., facilities and accessibility). The rationale for choosing subscales should be given;

3) The instrument should allow a quick administration (30 items maximum), and items should be written in simple, colloquial language that can be understood by participants

with low levels of educational. A balance between positive and negative sentences is advisable to neutralize de acquiescence effect. It should be self-administered. Responses should be anchored on a four- or six-point scale to avoid the tendency to choose neutral answers.

4) To ensure content validity, the pool of items created for each dimension should be submitted to an external evaluation that could provide evidence about an item's relevance to the construct and an adequate representation of each one of the selected dimensions. It is also advisable to evaluate the clarity of the items. For this purpose, a panel of experts selected on basis of their knowledge or their similarities with the target population could be consulted. Experts' evaluations could be performed with a simple numerical scale (e.g., five to seven point scale) or with the Delphi method. Should authors decide to modify or add new items after an expert evaluation, the process must be repeated. Authors should indicate which items have been eliminated and why, while specifying the final pool of items.

5) A previous analysis of the metric properties of the items should be performed, typically through a pilot administration of the items selected in the previous step to a similar sample. The aim is to select items that maximize the variance, so authors should select items with high discriminatory capacity, high standard deviations, and mean scores close to the medium point of the scale.

6) The internal structure of the scale must be explored to evaluate the extent to which the items and dimensions of the instrument are embedded in the construct. In sum, authors must determine whether their theoretical hypothesis about how the items were clustered is confirmed empirically. For this purpose, performing a confirmatory factor analysis is recommended, although a previous exploratory procedure might also be appropriate. In either case, authors should provide the number of resulting factors, the saturation of the items in such factors, the amount of variance explained by each factor and items together, the goodness of fit of the model, and the residuals.

7) The reliability of the instrument can be estimated using different methods. The one-test method is the most frequently used, specifically the calculation of internal consistency through Cronbach's alpha. Nonetheless, this index is highly influenced by the number of items, and very high internal consistency values (> .95) might indicate that items are redundant and therefore that the scale has an insufficient content validity rather than a high reliability. Hence, it is advisable to consider using alternative methods, such as the parallel form method or the test-retest method.

8) Lastly, it is advisable to determine if the predicted theoretical associations between test scores and other relevant external variables are confirmed. In brief, in this step, the psychological coherence of the construct is established. Authors should study the associations between the instrument scores and a criteria that are expected to be predicted by these scores (criteria validity), other scales that measure the same construct (convergent validity), and other variables that should be differentiated (discriminant validity). For instance, the instrument scores should predict outcomes such as program completion or engagement with the intervention, be highly associated with other instruments that evaluate the satisfaction with services, and be negatively associated with outcomes such as service drop-out.

All of the aforementioned procedures and results should be clearly reported.

Developing standardized and rigorous instruments and procedures for obtaining family feedback as well as incorporating such feedback into performance improvement strategies in CWS could have an impact that is not achievable through informal or ad hoc mechanisms (Tilbury et al., 2010). It could also overcome the inherent limitations of using data generated through focus groups, open interviews or suggestions. Consumers of public services are also citizens, and this fact implies rights to equity, representation, and participation (Pollitt, 1998), particularly in the context of overt user-provider asymmetries, such as child welfare. Above all, from a family preservation and strengths-focused perspective and to guarantee a democratic functioning of institutions, it seems necessary to give a voice to the protagonists of the intervention – the families.

References

Alpert, L. T. (2005). Research review: Parents' service experience – a missing element in research on foster care case outcomes. *Child & Family Social Work*, *10*(4), 361-366. doi: 10.1111/j.1365-2206.2005.00387.x

American Humane Association (1998). *Assessing outcomes in child welfare services: Principles, concepts, and a framework of core outcome indicators*. Englewood, CO: American Humane Association.

Baker, A. J. (2007). Client feedback in child welfare programs: Current trends and future directions. *Children and Youth Services Review, 29*, 1189-1200. doi: 10.1016/j.childyouth.2007.05.003

Berrick, J. D., Frasch, K., & Fox, A. (2000). Assessing children's experiences of out-of-home care: Methodological challenges and opportunities. *Social Work Research, 24*(2), 119–127. doi: 10.1093/swr/24.2.119

Buckley, H., Carr, N., & Whelan, S. (2011). 'Like walking on eggshells': Service user views and expectations of the child protection system. *Child & Family Social Work*, *16*(1), 101-110. doi: 10.1111/j.1365-2206.2010.00718.x

Carretero-Dios, H., & Pérez, C. (2005). Normas para el desarrollo y revisión de estudios instrumentales. *International Journal of Clinical and Health Psychology, 5*(3), 521-551.

* Chaffin, M., Bard, D., Bigfoot, D. S., & Maher, E. J. (2012). Is a structured, manualized, evidence-based treatment protocol culturally competent and equivalently effective among American Indian parents in child welfare? *Child Maltreatment*, *17*(3), 242-252. doi: 10.1177/1077559512457239

Chapman, M. V., Gibbons, C. B., Barth, R. P., & McCrae, J. S., & the NSCAW Research Group. (2003). Parental views of in-home services: What predicts satisfaction with child welfare workers? *Child Welfare, 82*, 571–596.

Connolly, M. (2007). Practice frameworks: Conceptual maps to guide interventions in child welfare. *British Journal of Social Work, 37*(5), 825–837. doi: 10.1093/bjsw/bcl049

Cortis, N. (2007). What do service users think of evaluation? Evidence from family support. *Child & Family Social Work*, *12*(4), 399-408. doi: 10.1111/j.1365-2206.2007.00495.x

Damashek, A., Doughty, D., Ware, L., & Silovsky, J. (2011). Predictors of client engagement and attrition in home-based child maltreatment prevention services. *Child Maltreatment*, *16*(1), 9-20. doi: 10.1177/1077559510388507

* Damashek, A., Bard, D., & Hecht, D. (2012). Provider cultural competency, client satisfaction, and engagement in home-based programs to treat child abuse and neglect. *Child Maltreatment*, *17*(1), 56-66. doi: 10.1177/1077559511423570

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th ed.). Boston: Allyn & Bacon.

Girvin, H., DePanfilis, D., & Daining, C. (2007). Predicting program completion among families enrolled in a child neglect preventive intervention. *Research on Social Work Practice*, *17*, 674-685. doi: 10.1177/1049731507300285

Gladstone, J., Dumbrill, G., Leslie, B., Koster, A., Young, M., & Ismaila, A. (2012). Looking at engagement and outcome from the perspectives of child protection workers and parents. *Children and Youth Services Review*, *34*(1), 112-118. doi: 10.1016/j.childyouth.2011.09.003

* Green, B. L., McAllister, C. L., & Tarte, J. M. (2004). The Strengths-Based Practices Inventory: A tool for measuring strengths-based service delivery in early childhood and family support programs. *Families in Society: The Journal of Contemporary Social Services*, *85*(3), 326-334. doi: 10.1606/1044-3894.1493

Harris, G., & Poertner, J. (1998). *Measurement of client satisfaction: The state of the art*. Urbana, Ill: Children and Family Research Center.

Heneghan, A., Horwitz, S., & Leventhal, J. (1996). Evaluating intensive family preservation programs: A methodological review. *Pediatrics, 97*(4), 535–542.

* Huebner, R. A., Jones, B. L., Miller, V. P., Custer, M., & Critchfield, B. (2006). Comprehensive family services and customer satisfaction outcomes. *Child Welfare, 85*(4), 691-714.

* Huebner, R. A., Werner, M., Hartwig, S., White, S., & Shewa, D. (2008). Engaging fathers: Needs and satisfaction in child protective services. *Administration in Social Work*, *32*(2), 87-103. doi: 10.1300/J147v32n02_06

Johnson, L. N., & Ketring, S. A. (2006). The therapy alliance: A moderator in therapy outcome for families dealing with child abuse and neglect. *Journal of Marital and Family Therapy*, *32*, 345-354. doi: 10.1111/j.1752-0606.2006.tb01611.x

Johnson, L. N., Wright, D. W., & Ketring, S. A. (2002). The therapeutic alliance in home-based family therapy: Is it predictive of outcome? *Journal of Marital and Family Therapy*, *28*, 93-102. doi: 10.1111/j.1752-0606.2002.tb01177.x

Kapp, S. A., & Vela, R. H. (1999). Measuring consumer satisfaction in family preservation services: Identifying instrument domains. *Journal of Family Strengths*, *4*(2), 1-19.

Kelly, S., & Blythe, B. (2000). Family preservation: A potential not yet realized. *Child Welfare, 79*(1), 29–42. doi:

Korfmacher, J., Green, B., Spellmann, M., & Thornburg, K. R. (2007). The helping relationship and program participation in early childhood home visiting. *Infant Mental Health Journal*, *28*, 459-480. doi: 10.1002/imhj.20148

Lee, C. D., & Ayón, C. (2004). Is the client-worker relationship associated with better outcomes in mandated child abuse cases? *Research on Social Work Practice*, *14*, 351-357. doi: 10.1177/1049731504265833

Lietz, C. A. (2009). Examining families' perceptions of intensive in-home services: A mixed
methods study. *Children and Youth Services Review*, *31*(12), 1337-1345. doi:
10.1016/j.childyouth.2009.06.007

Littell, J. (2001). Client participation and outcomes of intensive family preservation services.
Social Work Research, 25(2), 103–113. doi: 10.1093/swr/25.2.103

Marsh, J. C., Angell, B., Andrews, C. M., & Curry, A. (2012). Client-provider relationship and
treatment outcome: A systematic review of substance abuse, child welfare, and mental
health services research. *Journal of the Society for Social Work and Research*, *3*(4), 233-
267. doi: 10.5243/jsswr.2012.15

* McMurtry, S. L., & Hudson, W.W. (2000). The Client Satisfaction Inventory: Results of an
initial validation study. *Research on Social Work Practice, 10*, 644–663.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group (2009). Preferred
reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS
Medicine, 6*, e1000097. doi: 10.7326/0003-4819-151-4-200908180-00135

Pascoe, G. C. (1983). Patient satisfaction in primary health care: A literature review and
analysis. *Evaluation and Program Planning*, *6*(3), 185-210. doi: 10.1016/0149-
7189(83)90002-2

Pollitt, C. (1998). Bringing consumers into performance measurement: Concepts,
consequences and constraints. *Policy and Politics, 16*, 77-87.

* Reid, M. J., Webster-Stratton, C., & Beauchaine, T. P. (2001). Parent training in Head Start: A
comparison of program response among African American, Asian American, Caucasian,
and Hispanic mothers. *Prevention Science*, *2*(4), 209-227. doi:
10.1023/A:1013618309070

Rodrigo, M. J., Maiquez, M. L., Martín, J. C., & Byrne, S. (2008) *Preservación familiar. Un
enfoque positivo para la intervención con familia*s. Madrid: Pirámide.

Russell, M. N. (1990). Consumer satisfaction: An investigation of contributing factors. *Journal
of Social Service Research, 75*(4), 43-56. doi: 10.1300/J079v13n04_03

Southerland, D. G., Mustillo, S. A., Farmer, E.M.Z., Stambaugh, L. F., & Murray, M. (2009).
What's the relationship got to do with it? Understanding the therapeutic relationship in
therapeutic foster care. *Child and Adolescent Social Work Journal*, *26,* 49-63.
doi:10.1007/s10560-008-0159-4

Tilbury, C; Osmond, J; Crawford, M. (2010). Measuring client satisfaction with child welfare
services. *Journal of Public Child Welfare, 4*(1), 77-90. doi: 10.1080/15548730903563160

Trotter, C. (2008). What does client satisfaction tells us about effectiveness? *Child Abuse
Review, 17*, 262-274. doi: 10.1002/car.1038

* Winefield, H. R., & Barlow, J. A. (1995). Client and worker satisfaction in a child protection

agency. *Child Abuse & Neglect*, *19*(8), 897-905. doi: 10.1016/0145-2134(95)00052-A

Young, S., Nicholson, J., & Davis, M. (1995). An overview of issues in research on consumer

satisfaction with child and adolescent mental health services. *Journal of Child and Family*

Studies, 4, 219-238. doi: 10.1007/BF02234097

- A systematic review following the PRISMA guidelines was conducted

- Eight measures assessing family feedback in Child Welfare Services were analyzed

- Measures' conceptual and psychometric characteristics were often unreported

- Further research is needed to develop and validate measures in this field