

# Satisfaction with social welfare services

– A review

*You may quote text from the National Board of Health and Welfare in, for example, educational material provided at cost price, providing you state the source. You may not however use the text in any commercial context. The National Board of Health and Welfare has the sole right to decide how this work shall be used, in accordance with the Act (1960:729) on copyright in literary and artistic works (the Copyright Act). Pictures, photographs and illustrations are also protected by the Copyright Act and their use requires the permission of the author.*

ISBN 978-91-7555-107-4

Article no 2013-10-23

---

Published [www.socialstyrelsen.se](http://www.socialstyrelsen.se), oktober 2013

# Förord

---

Idag är det en självklarhet att undersöka vård- och omsorgstagares tillfredsställelse med de tjänster de erbjuds. Utan den kunskapen minskar möjligheten att ge bästa möjliga hjälp. I denna litteraturöversikt sammanfattas kunskapen om metoder för att mäta vård- och omsorgstagares tillfredsställelse inom socialtjänstområdet.

Litteraturöversikten har på Socialstyrelsen uppdrag gjorts av professor Mark W. Fraser och doktorand Shiyu Wu, School of Social Work, University of North Carolina Chapel Hill, USA. Haluk Soydan från Socialstyrelsen har fungerat som kontaktperson till Mark W. Fraser.

Stockholm i oktober 2013

Knut Sundell  
Socialstyrelsen

# Content

---

<i>Förord</i>	3
<i>Executive Summary</i>	5
<i>Sammanfattning</i>	8
<i>Satisfaction with Social Welfare Services: A systematic Review</i>	11
<i>Part 1. Historical and Theoretical Foundations of Consumer Satisfaction</i>	13
<i>Part 2. Current Conceptualizations of Consumer Satisfaction</i>	23
<i>Part 3. The Psychometric Characteristics of Consumer Satisfaction</i>	32
<i>Part 4. Consumer Satisfaction Scales</i>	37
<i>Part 5. Summary and Recommendations</i>	48
<i>References</i>	56
<i>Appendix A: Summary of Consumer Satisfaction Instruments, 2003 – 2013</i>	68

# Executive Summary

---

In program evaluation and quality assurance, the use of *satisfaction with services* as an outcome measure is consistent with broad movements, such as New Public Management and Evidence-based Practice. In addition, the idea of consumer satisfaction recognizes the importance of client-centered practice. Moreover, through the empowerment and popular consumer movements, the involvement of service recipients in the design and evaluation of social welfare programs is considered as a wise and ethical professional practice.

In this context, the purpose of this report is to assess recent research on satisfaction with services in social welfare, as well as to review satisfaction scales that have been used in evaluations of social welfare programs. Drawing on the literatures in social welfare, public administration, and business, the report is organized into five parts: (a) the historical and theoretical foundations of consumer satisfaction in social welfare and business; (b) current conceptualizations of satisfaction in social welfare; (c) synthesis of research on the reliability and validity of satisfaction measures; (d) systemic review of satisfaction scales in social welfare published between 2003 and May 2013 ; and (e) summary and recommendations.

The review included nine databases and identified 58 satisfaction-related instruments developed or refined in the last 10 years. The number and variety of instruments suggest that no single scale currently captures all the aspects of satisfaction. On average, scales were brief, with a 10-item format being most frequently used. The length of a scale is an important consideration because the burden of completing a survey may affect the quality of the data. That is, longer surveys with a greater number of items can reduce response rates.

On balance, reliability was acceptable. The average reliability across instruments was .85, which falls in the acceptable range.

No common dimensionality emerged across the 58 satisfaction measures. Some scales have one dimension, whereas others have two or more dimensions. Some scales explicitly use the term *satisfaction* (e.g., *How satisfied were you with <x>?*), while others do not. Some scales include net promotion items (e.g., *Net of everything in your experience at <x>, would you recommend this service?*) or word-of-mouth recommendation (e.g., *If a friend were in need of similar help, would you recommend our program to him or her?*). After more than 35 years of development in social welfare, the dimensionality of satisfaction with services remains unclear.

The research on satisfaction with services has serious limitations. Many satisfaction scales appear insufficient in representing alternative aspects of satisfaction. Most reports (64%) contained no validity analyses. The methods used to score instruments rarely accounted for different elements of service and the relative importance of service elements. On balance, the measurement of satisfaction with services is less sophisticated and nuanced than

the measurement of other constructs – such as social problems and mental health symptoms – in social welfare.

In addition, a variety of problems affect the interpretation of the consumers' satisfaction with services. To have validity in making an inference about a service, for example, satisfaction information is needed from all participants who begin services or, alternatively, attrition must be shown to be missing at random. A frequent result is that those that prematurely terminate a service are on the average less satisfied.

In addition, at least three other confounding factors may affect satisfaction ratings: the image of the service provider (e.g., the reputation and community standing), the affective or utilitarian aspects of service (e.g., the courtesy of staff, the availability of parking, the attractiveness of the facility), and the consumer's sense of equity relative to the services received by others. These are considered confounding influences because they complicate using satisfaction to indicate the quality of services. Confounding factors may influence satisfaction ratings independent of the actual effect of a service.

Further complicating the use of satisfaction as an evaluation measure is the fact that satisfaction ratings are known to be high for participation in nearly all social welfare services. Participants involved in reading groups, discussions, and social support groups, or who receive a placebo intervention often report as high satisfaction scores as those that receive an active intervention. These high satisfaction ratings are sometimes attributed to the *Thank-You* effect. The thank-you effect derives from genuine appreciation that stems from participation in any service – including study groups, seminars, and the like. These thank-you effects are nontrivial, and in assessing the impact of a service, satisfaction ratings must be controlled with an attention-only or support services control condition.

An emerging feature of some satisfaction measures is the use of subjective causal appraisals. On the order of inviting reports of perceived change, some instruments invite appraisals of the impact of social welfare services. Compared to global satisfaction scores, these kinds of satisfaction measures may have greater predictive validity for long-term outcomes.

In sum, many scales and measures of satisfaction are available. They range from simple single-item measures to multi-dimensional scales of the acceptability and the perceived effects of service participation. Reliability is generally in the acceptable range. However, the degree to which satisfaction predicts behavioral and other theoretically important outcomes remains uncertain.

## **Recommendations**

From our review, satisfaction with services provides useful information. Satisfaction is a function of service engagement. More engaged program participants are likely to report higher satisfaction and to observe greater benefit from their receipt of services. In this sense, satisfaction is an outcome that sometimes can predict behavioral and other outcomes. However, satisfaction with services should not be used alone in program evaluation and quality assurance. It should complement the use of other relevant out-

comes. If satisfaction with service is to be used in evaluation or quality assurance, several factors should be considered.

- Satisfaction measures should include at least three kinds of questions: (a) satisfaction with service elements (e.g., *To what degree are you satisfied with [service element x, y, or z]?*), (b) promotion given an entire service experience (e.g., *Would you recommend this program a friend?*), and (c) perceived change (e.g., *To what degree did participation in the program resolve your problems?*). Because of the potential for these three elements of satisfaction to be differentially related to other outcomes, they should be considered separately as well as combined in data analysis.
- Satisfaction items that are related directly to the key elements of services and that invite a subjective causal appraisal of perceived changes may be preferable to global satisfaction ratings.
- Strategies must be developed to reduce attrition and to secure satisfaction ratings from program dropouts. These include conducting exit interviews with dropouts and collecting satisfaction ratings incrementally throughout service periods.
- If consumer satisfaction data are to be used to compare services across providers, information on potential confounding variables must be collected. In particular, the image or reputation of service providers must be controlled in analyses.

# Sammanfattning

---

Att undersöka vård- och omsorgstagares tillfredsställelse med tjänster ligger i linje med nya paradigmer som New public management och Evidensbaserad praktik. Det är också i enlighet med klientcentrerad praktik. Vård- och omsorgstagares delaktighet i utformning och utvärdering av sociala välfärdstjänster är också etiskt önskvärt utifrån synen på egenmakt (*eng. Empowerment*).

Syftet med denna rapport är att granska den senaste forskningen om vård- och omsorgstagares tillfredsställelse med välfärdstjänster samt att granska bedömningsmetoder som har använts för att utvärdera tillfredsställelse med välfärdstjänster. Rapporten är indelad i fem avsnitt: (1) den historiska och teoretiska grunden för vård- och omsorgstagares tillfredsställelse inom socialtjänst och näringsliv, (2) begreppsbyggnad om tillfredsställelse inom socialtjänst; (3) syntes av forskningen av reliabilitet och validitet hos bedömningsmetoder som mäter tillfredsställelse inom socialtjänst, (4) systematisk översikt av de bedömningsmetoder som mäter tillfredsställelse inom socialtjänst och som publicerats mellan 2003 och maj 2013, samt (5) sammanfattning och rekommendationer.

Genom sökning i nio databaser identifierades 58 bedömningsmetoder. Antalet och variationen i bedömningsmetoder tyder på att det saknas en enskild bedömningsmetod som fångar alla aspekter av. I allmänhet är bedömningsmetoderna korta, ofta omfattande ungefär tio frågor. Antalet frågor spelar roll eftersom många delfrågor kan minska svarsfrekvensen och därmed påverka kvaliteten på resultaten.

Som helhet var reliabiliteten i de 58 bedömningsmetoderna acceptabelt, i genomsnitt 0,85.

Det saknas gemensamma dimensioner i de 58 bedömningsmetoderna. Några har en dimension medan andra har två eller flera. Vissa bedömningsmetoder använder begreppet tillfredsställelse (t.ex. "hur nöjd var du med X?") medan andra inte gör det. Vissa inkluderar frågor om den sammanvägda upplevelsen (t.ex. "med hänsyn taget till hela din upplevelse av X, skulle du rekommendera den?") eller rekommendation (t.ex. "om en vän var i behov av liknande hjälp skulle du rekommendera X till honom eller henne?"). Efter mer än 35-års utveckling förblir dimensionerna i tillfredsställelse inom social välfärd oklara.

Forskningen om personers tillfredsställelse med insatser har tydliga begränsningar. Många bedömningsmetoder beskriver inte tillräckligt väl alternativa dimensioner i tillfredsställelse. De flesta (64 procent) saknar information om validitet. De metoder som används för att poängsätta bedömningsmetoderna redovisar sällan den relativa betydelsen av olika komponenter i välfärdstjänster. Som jämförelse är bedömningsmetoder som mäter tillfredsställelse i allmänhet mindre sofistikerade och nyanserade än bedömningsmetoder som mäter andra företeelser inom socialtjänsten, exempelvis förekomsten av sociala problem och psykisk hälsa.



Flera problem påverkar tolkningen av vård- och omsorgstagares tillfredsställelse. För att det ska gå att dra slutsatser om en tjänsts värde behövs till exempel information från alla deltagare som börjat använda tjänsten, alternativt att bortfallet är slumpmässigt. Ett vanligt resultat är att de som avbryter en behandling i förtid i genomsnitt är mindre tillfreds.

Dessutom kan minst tre andra faktorer påverka tillfredsställelsen: tjänsteleverantörers anseende (t.ex., rykte), bemötande och praktiska aspekterna av tjänsten (t.ex., hjälpsam personal, parkeringsmöjligheter, trevliga lokaler) samt deltagarnas upplevelse av likvärdighet i de tjänster som andra i samma situation erhåller. Dessa faktorer påverkar också tillfredsställelsen utöver kvaliteten på tjänsten och kan påverka tillfredsställelsen oberoende av om insatsen är effektiv eller inte.

En ytterligare komplikation med att använda tillfredsställelse som ett utvärderingsmått är att vård- och omsorgstagare nästan alltid är tillfreds med välfärdstjänster. Individer som deltar i läsgrupper, diskussioner och sociala stödgrupper, eller som får en placebo är ofta lika tillfreds som de som får en mer aktiv behandling. Denna "tacksamhetseffekt" (*eng. thank-you effect*) är en äkta uppskattning av att ha fått delta i en tjänst. Tacksamhetseffekten är inte trivial och vid utvärdering av effekterna av en specifik tjänst behöver den kontrolleras för genom att undersöka tillfredsställelse hos individer som fått enbart uppmärksamhet (*eng. attention-control*) eller någon annan form av stöd.

Ett ökat inslag i bedömningsmetoder för tillfredsställelse är användningen av subjektiva kausala bedömningar, dvs. att individen får beskriva upplevelsen av insatsens effektivitet. Jämfört med globala mått på tillfredsställelse kan den här typen av mått ha bättre prediktiv validitet av långsiktiga utfall.

Sammanfattningsvis finns det många bedömningsmetoder för vård- och omsorgstagares tillfredsställelse. De inkluderar allt från enstaka frågor till flerdimensionella skalor som mäter deltagarnas acceptans och upplevelser av effekter av tjänsten. Tillförlitligheten i dessa är i allmänhet acceptabel. Däremot är det oklart hur väl personers tillfredsställelse kan förutse beteendemässiga och andra teoretiskt viktiga utfall.

## Rekommendationer

Vård- och omsorgstagares tillfredsställelse med tjänster ger viktig information. Tillfredsställelse är en funktion av engagemang i behandlingen. De som är mer engagerade är benägna att rapportera högre tillfredsställelse och dra större nytta av insatsen. I det avseendet kan tillfredsställelse även förutsäga beteendeförändring och andra utfall. Men bedömningsmetoder ska inte användas som det enda utfallsmåttet i utvärdering och kvalitetssäkring utan är ett viktigt komplement till andra relevanta utfallsmått.

Om tillfredsställelse med tjänster ska användas är flera faktorer viktiga att ta hänsyn till:

- Bedömningsmetoder om tillfredsställelse bör omfatta minst tre typer av frågor: (a) tillfredsställelse med specifika komponenter i behandlingen (t.ex. "I vilken utsträckning är du nöjd med [komponent x, y eller z]?"); (b) tillfredsställelse med hela tjänsten (t.ex. "Skulle du rekommendera den här insatsen till en vän?"); samt (c) upplevd för-

ändring (t.ex., ”I vilken grad har deltagandet i insatsen löst dina problem?”). Eftersom dessa tre typer av tillfredsställelse kan skilja sig i relation till andra utfallsmått bör de både analyseras separat och kombinerat.

- Frågor om centrala komponenters betydelse för upplevd förändring kan vara att föredra framför allmänna omdömen om tillfredsställelse.
- Det behövs strategier för att minska bortfall av svar och för att säkra svar från deltagare som inte fullföljer behandlingen. Det inkluderar att intervjua de som avbryter en behandling och att samla information om tillfredsställelse upprepade gånger under behandlingsperioden.
- Om vård- och omsorgstagares tillfredsställelse ska användas för att jämföra tjänster mellan leverantörer måste information om potentiella störande variabler samlas in, i synnerhet tjänsteleverantörers anseende.

# Satisfaction with Social Welfare Services: A systematic Review<sup>1</sup>

---

Satisfaction with services is selected often as an outcome in evaluations of social welfare programs (see, e.g., Royse, Thyer, & Padgett, 2010). The logic that has supported the use of satisfaction is compelling: If recipients of services are satisfied, services are considered, at least in part, successful. This perspective derives from the historical roots of customer or consumer satisfaction as an indicator of purchasing sentiment in business. For more than 50 years, customer satisfaction has been considered an indicator of the performance of both commercial products and services. Consumers who report satisfaction with a particular product – for example, an automobile – are thought more likely to purchase that same product or brand in the future. Brand loyalty is known to increase demand for products and, in the long run, to contribute to cash flows and gross margins (e.g., Homburg, Koschate, & Hoyer, 2005; van Doorn, Leeflang, & Tijs, 2013). However, is this logic sustained when satisfaction is used with social welfare services?

The use of satisfaction in social welfare is consistent with broad movements that have influenced public policies in Sweden and other countries.<sup>2</sup> Indeed, the idea of *consumer* satisfaction recognizes the personal agency of service recipients (i.e., the power and authority) and implies that services should be person-centered. The theme of agency is consistent with New Public Management (NPM), in which the administration of many programs has been decentralized, managed through municipal authorities, and privatized to increase efficiencies and consumer responsiveness (Ferlie, Ashburner, Fitzgerald, & Pettigrew, 1996; Hood, 1991). In addition, consumer agency is consistent with the themes of stakeholder participation in the emerging Public Value perspective (Moore, 1995; Stoker, 2006).<sup>3</sup> At the same time, evidence-based practice requires greater client involvement in both the selection and evaluation of practice strategies. Moreover, through the empowerment and popular consumer movements, the involvement of service recipients in the design and evaluation of social welfare programs is considered as a wise and ethical professional practice.

In contrast to its role in social welfare, consumer satisfaction has a more utilitarian role in business. Consumer satisfaction in business is based on the

---

<sup>1</sup> The report was supported by a Grant 13-0789 from the National Board of Health and Social Welfare, Sweden. We thank Haluk Soydan, PhD, Knut Sundell, PhD, and Diane Wyant, MDiv, who provided helpful comments on drafts of the manuscript. Rea Gibson, BFA, assisted in developing figures and artwork.

<sup>2</sup> Throughout this report, the term “social welfare” is used to refer to a range of services in aging, child welfare, corrections, developmental disabilities, housing, income assistance, job promotion and employment, juvenile justice, mental health, and substance abuse.

<sup>3</sup> The New Public Management and the Public Value perspectives represent alternative approaches to the design and implementation of social welfare programs. The description of these two perspectives is beyond the scope of this report.

ideas of product choice in a competitive marketplace and the exchange of money for services or goods. From this perspective, satisfaction is viewed largely as a cognitive appraisal of product performance. If a chosen product or service performs as expected, consumers will give it higher satisfaction scores. In turn, higher satisfaction scores are known to increase demand and profits through brand loyalty and word-of-mouth recommendations.

In social welfare, the relationship of consumer or client satisfaction to service provision is different from the business model. Indeed, rather than increasing product demand, the goal of social welfare services is to decrease demand by solving problems; in doing so, services achieve public value (Moore, 1995). In addition, as opposed to customers in business, consumers in social welfare often have more limited choices. Clients in social welfare agencies might not only have few services from which to choose, but in some fields – child welfare, juvenile justice, and substance abuse – freedom to exercise choice may be constrained. For example, in child welfare, an abusive or neglectful parent might have the choice of either participating in a family preservation program (involving caseworker supervision and parenting training) or having a child removed from the home and placed in protective care. This delimited choice is not the same agency afforded a consumer of a commercial product or service.

Moreover, it is not clear whether the principles underwriting consumer satisfaction in business operate when service participation is compelled by a court order or when services have an investigative element (e.g., Martin, Petr, & Kapp, 2003). In a study of pediatric hospitalizations in which 120 children were compared based on whether medical staff conducted a maltreatment screening, parents in the maltreatment-assessment group reported significantly lower satisfaction (e.g., *treatment was respectful*, *information provided was honest*) as compared with parents whose children were in the non-screened condition (Ince, Rubin, & Christian, 2010). Although screenings for maltreatment are necessary and have public value (i.e., the protection of children), they can be associated with lower consumer satisfaction. The principles and practices associated with customer satisfaction in business might not be easily compared with the principles and practices that guide the provision of many social welfare programs.

In this context, the purpose of this report is to assess the validity and reliability of satisfaction with services in social welfare, as well as to review satisfaction scales that have been used in evaluations of social welfare programs. Drawing on the literatures in social welfare, public administration, and business, the report is organized into five parts: (a) the historical and theoretical foundations of consumer satisfaction; (b) current conceptualizations of satisfaction in social welfare and business; (c) synthesis of research on the validity of satisfaction measures; (d) systemic review of satisfaction scales in social welfare; and (e) summary and recommendations.

# Part 1. Historical and Theoretical Foundations of Consumer Satisfaction

---

Developed in the 1960s, consumer satisfaction is often described as a pragmatic indicator of the success of social welfare programs (see, e.g., Copeland, Koeske, & Greeno, 2004; Fox & Storms, 1981; Locker & Dunt, 1978; Young et al., 1995). The term *consumer* is variably defined to include clients, patients, users, and others who participate in – or consume – a social or health service (Sharma, Whitney, Kazarian, & Manchanda, 2000). Reflecting a lack of unanimity on a proper term, some studies of child maltreatment, partner violence, health events such as heart attacks, and even weather events such as hurricanes use the term *survivor* satisfaction (e.g., Allen, Brymer, Steinberg, Vernberg, Jacobs, Speier, & Pynoos, 2010). On balance, no term seems fully appropriate for social welfare. In this report, we will use *consumer satisfaction* and *client satisfaction* interchangeably, and acknowledge that the field needs more inclusive and socially nuanced terminology.

The term *satisfaction* is usually interpreted as the appeal, acceptability, and approval of a service experience. Sometimes satisfaction includes liking or feeling personally involved in elements of service and contentment with outcomes (Nelson & Steele, 2006). The core argument for using satisfaction as an outcome in social welfare is simple: If clients feel satisfied with a program, they are more likely to have been engaged in program activities, to have adhered to program recommendations, and to have experienced program-derived benefits. From this perspective, consumer satisfaction is conceptualized as a predictor of concrete outcomes such as adaptive functioning and distal outcomes that have public value. These distal outcomes might include academic success, civic participation, sustained employment, and positive health behavior.

Notwithstanding this logic, research on satisfaction has yielded mixed findings. After assessing the use of consumer satisfaction in allied health, Koch and Rumrill (2008, p. 358, 362) commented, “Satisfaction...is often unrelated to the actual quality of technical services.... [It] has proven to be a very difficult variable to measure and interpret.” In a systematic review of 195 studies that assessed satisfaction in various health care settings (i.e., hospital inpatient, hospital outpatient, mental health service, and other care settings such as dental, maternity, and palliative care), Sitzia (1999, p. 327) found “only 6% ... used instruments which demonstrated the overall minimum level of evidence for reliability and validity.” He concluded (p. 327) that findings based on consumer satisfaction “lack credibility.”

Despite such criticism, credible studies increasingly use satisfaction as an outcome. For example, in an experimental test of a parent, child, and teacher

training in a sample of 159 children diagnosed with oppositional defiant disorder, Webster-Stratton, Reid, and Hammond (2004) used a multi-item satisfaction measure consisting of *treated problems are improved*, *feeling optimistic about child's problems*, *expecting good results*, *willing to recommend program to others*, and *confidence in managing child's problems*. Positive reports of satisfaction covaried with positive parent- and teacher-reports of child behavior. That is, satisfaction had concurrent validity in that it was related consistently to other measures (in this case, reports of child behavior) that had high theoretical relevance.

The mixed findings and, indeed, views on consumer satisfaction in the literature raise questions about the psychometric properties of satisfaction-related service appraisals and the relationship of satisfaction-based measures to both proximal and distal services outcomes.<sup>4</sup> Are the Webster-Stratton et al. (2004) findings anomalous? Or is consumer satisfaction a valid and reliable outcome that varies logically with other important outcomes? The purpose of this report is to describe consumer satisfaction, its properties, and its potential as an outcome in social welfare services.

### **Emergence of Consumer Satisfaction in Business**

In the 1960s and 1970s, consumer satisfaction emerged in business as researchers sought to explain the preferences and purchasing habits of potential buyers of goods and services. Consumer satisfaction was conceptualized as a posthoc evaluation of a buying experience. Further, researchers sought to explain the extent to which consumer evaluations of purchasing experiences might be tempered by expectations, which were perceived as representing the sum of prior experiences with similar products plus the influence of product-related advertising and packaging.

To conceptualize consumer satisfaction, three theoretical models emerged: (1) the contrast model, (2) an assimilation model, and (3) an integrated contrast-assimilation model (Day, 1977; Hunt, 1977a). The contrast model focused on the discrepancy between expectations and product performance. According to Pascoe (1983, p. 187), "Performance that is somewhat higher than expectations will be evaluated as satisfactory, whereas performance slightly less than expected will be judged as unsatisfactory." The assimilation model predicted, "...performance that is moderately lower than expectations will not cause dissatisfaction because perceptions of performance will be assimilated to match higher expectations" (Pascoe, p. 187). The assimilation-contrast model integrated the contrast and assimilation perspectives:

Expectations serve as a standard for judging a product or service, but there is latitude of acceptance surrounding this standard. Discrepancies that are within this latitude will be assimilated as follows: (a) expectations that are lower than outcomes will lead to de-

---

<sup>4</sup> Proximal outcomes are related strongly to the short-term goals of services. They may include, for example, changes in skills or increases in knowledge. They are often measured at the conclusion of services. Distal outcomes are related to the long-term goals of social programs. These may include reductions in hospitalizations or arrests over time, a sustained pattern of employment over time, or housing stability.

creased performance evaluations, lessening satisfaction more than if performance matches expectations; (b) expectations that are higher than outcomes will increase performance judgments, causing greater satisfaction than if performance and expectations matched. Contrast effects occur when discrepancies between performance and expectations are relatively large. In such cases, the latitude of acceptance is exceeded and the predictions of the contrast model are considered to apply (Pascoe, 1983, p. 187-188).

The contrast model of satisfaction was the first of the three conceptual models to be experimentally tested. In 1964, Helson found that consumer satisfaction judgments were determined by previous experience within a general product category and the discrepancy between previous experience and new experiences. Today, the contrast model continues to be implicit in common conceptualizations of consumer behavior (Cardozo, 1965; Engel, Kollat, & Blackwell, 1973; Howard & Sheth, 1969). That is, satisfaction is predicted by the intersection of prior and contemporary experience with a product or service.

In 1972, Olshavsky and Miller (1972) challenged the contrast model and re-conceptualized consumer satisfaction by proposing an assimilation model. In their view, “if the discrepancy is slight (within the latitude of acceptance), one would tend ... toward ... expectation” (p. 20). Olshavsky and Miller conducted an experiment to investigate the effects of both overstatement and understatement of product quality on product ratings (i.e., consumer satisfaction). In what would today be considered a dubious sample, the research used 100 male volunteers from undergraduate marketing classes. Based on a 2x2 factorial design with high and low levels for both expectation and performance, study participants were randomly assigned to the test conditions in blocks of four. The results showed overstatement produced more favorable ratings whereas understatement produced less favorable ratings. Applied (cautiously) to social welfare, the findings suggest that satisfaction ratings may be influenced by recruitment and retention strategies that promote the potential positive effects of program participation. That is, *recruitment and retention strategies that advertise the potential benefits of a social welfare program may contribute to satisfaction ratings beyond the actual impact of services*.

At nearly the same time as Olshavsky and Miller’s proposal of the assimilation model, Anderson (1973) proposed a model that integrated the contrast model and an assimilation model. Based on his findings from an experiment involving 144 students enrolled in an undergraduate marketing course, Anderson argued that, “...product perceptions will vary directly with expectations over a range around actual performance, but above and below this threshold, product perceptions will vary inversely with the level of expectation” (p. 41). Modern conceptualizations of consumer satisfaction are founded, in part, on the contrast-assimilation model (see also, Hunt, 1977b). Hunt (1977a), for example, concluded that consumer satisfaction is rooted in cognitive appraisal rather than emotional reaction. That is, he argued that satisfaction derives from an interaction of the service (or product), the situation, and expectations.

The field, however, has rarely achieved consensus. Shortly after the introduction of the third theoretical model, Linder-Pelz (1982b) defined satisfaction as an affective response to a consumer experience. Observing that satisfaction might be both an affective response and an outcome of experience, Yi (1990) later argued that two distinct types of satisfaction must be conceptualized: (a) satisfaction as a measure of outcome derived from a consumption experience; and (b) satisfaction as a measure of process during consumption experiences.

### **Consumer Satisfaction in Social Welfare**

By the late 1970s, scales for measuring consumer satisfaction began to penetrate the social welfare and health fields. The accountability movement, in which policymakers and scholars sought to make services more responsive to public need, accelerated the use of consumer satisfaction as a program outcome. At the same time, watchdogs and critics who had voiced concerns about the provision of ineffective and possibly iatrogenic services argued for consumer involvement in program evaluation. Further, the use of consumer satisfaction ratings was supported by advocates of quality assurance, who sought to prevent the delivery of poorly implemented or bogus interventions. Indeed, from a variety of sources, greater consumer involvement in both the design and evaluation of social and health services was demanded.

Within this context of increased emphasis on consumer involvement, researchers asked an interesting question: Independent of the effect of a service, what might influence satisfaction ratings? Their question implied that satisfaction might be influenced by an array of factors, including the extent to which a service is perceived as having public acceptance; the degree to which participation is voluntary (versus court ordered); the extent to which potential participants hold positive attitudes toward a service; the extent to which positive attitudes are held by people (staff and other service recipients) who consumers come to know through service involvement; and, more methodologically, the length and complexity of satisfaction measures (Kiesler, 1983). From this perspective, satisfaction began to be understood as a multiply determined construct that was influenced by the convenience, availability, efficacy, cost, and pleasantness of services. Moreover, satisfaction was understood as comparative. Satisfaction ratings involved assessing a service occurrence against expectations, weighing the service against values (i.e., what is good, what is bad), comparing a service experience to the experiences of others, and testing a service experience against normative beliefs (i.e., did the service fulfill commonly recognized expectations; Linder-Pelz, 1982).

Satisfaction in social welfare started to be conceptualized as having both cognitive and affective elements. From this perspective, it was thought to involve cognitive processes focused on the confirmation of expectations and affective processes focused on the more hedonic aspects of service or product experiences (see, e.g., World Health Organization [WHO], 2000).



## **Determinants of Satisfaction: The Expectancy-Performance Disconfirmation Model**

Although consumer satisfaction was used in social welfare and medicine, its uptake was eclipsed by its uptake in business, where it became common to find phrases such as “Satisfaction Guaranteed!” and where periodic consumer satisfaction surveys were (and continue to be) reported in national media. With the development of the European Customer Satisfaction Index and the American Customer Satisfaction Index, both of which were adapted from the Swedish Customer Satisfaction Barometer (Fornell, 1992), consumer satisfaction has secured a role as a major indicator of economic activity at the national and international levels. In business – more than in social welfare and medicine – dozens of studies have been undertaken to identify the determinants and consequences of customer satisfaction.

Across these studies, patterns have emerged that helped to illuminate the relationships among expectations, service or product experience, satisfaction, brand loyalty, and distal outcomes such as the financial health of companies that enjoy high customer satisfaction versus those that contend with low customer satisfaction. Of course, findings from research in business cannot be extrapolated to social welfare, where societal outcomes in addition to cost must be considered and where choice is often constrained. Nonetheless, findings from business research are helpful in identifying the range of variables that must be considered when consumer satisfaction is used as an outcome measure in social welfare. Overall, five constructs have been identified as predicting consumer satisfaction: expectations, service/product performance, disconfirmation, affect, and equity.

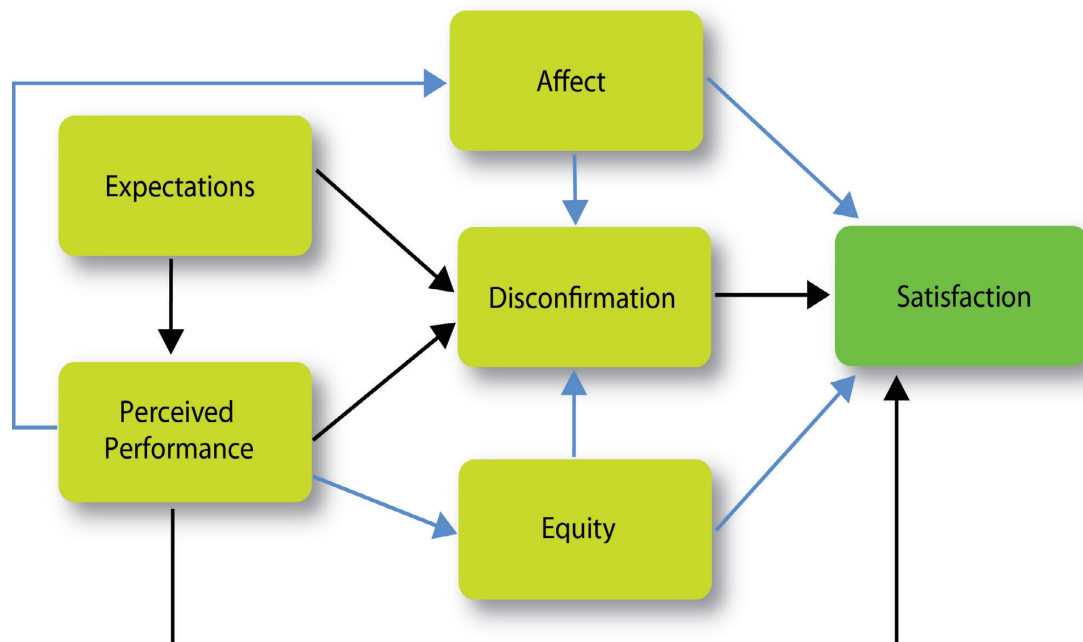
**Expectations.** Expectations are beliefs about how well a product will perform and attitudes about the likely outcomes of making a purchase or receiving a service. Performance expectations are shaped by prior experiences with products, services, advertising, and comparative referents. The key idea is that perceptions of the quality of a product or service can be shaped prior to product purchase and with little, if any, direct experience with a product or service (Fornell, Johnson, Anderson, Cha, & Bryant, 1996). From this perspective, expectations are influenced by product knowledge and all prior interactions – or referents – that might bear on the product.

**Performance.** Assessment of performance is a subjective appraisal of the quality of a service or product. One approach to this assessment has been driven by the perspective that consumers evaluate all services and products using a common set of performance indicators. Working from this perspective, researchers have attempted to develop latent performance criteria across which quality might be scored in many different service sectors. However, the findings have varied and seem dependent on the setting. A second approach, which has been less driven by research, has tended to conceptualize performance in terms of physical characteristics and interactional quality. Tied to specific transactions, interactional quality is defined as staff courtesy, knowledge, speed or promptness, and engagement. Physical performance indicators include appearance and functionality (e.g., ease of use, reliability). Appraisals of performance appear to be conditioned also on cost

heuristics. That is, a consumer might report “high satisfaction” with a restaurant that falls in the medium price range, but that consumer’s satisfaction ratings would decline if the same medium-price restaurant were to be compared with restaurants in a higher price category. The calculus of performance appears to be contextualized by price and schema.<sup>5</sup>

**Disconfirmation.** Disconfirmation was one of the earliest theoretical perspectives regarding consumer satisfaction, and the term refers to the congruence of performance with expectations. Shown in Figure 1, the Expectancy-Performance Disconfirmation Model has become a staple in the field. When product performance exceeds expectations, a positive disconfirmation (of expectations) is thought to produce high satisfaction ratings. When expectations exceed product performance, a negative disconfirmation (of expectations) is thought to produce low satisfaction ratings. Performance that meets expectations will produce weak satisfaction.

**Figure 1.** *The Expectancy-Performance Disconfirmation Model with Affect and Equity*



This suggests that social welfare services that conform to the hopes and perceptions of clients will not disconfirm expectations and will, therefore, receive high satisfaction. In contrast, services of lower quality than expected will disconfirm expectations and receive poor satisfaction scores.

**Affect.** Although early versions of the Expectancy-Performance Model did not include affect, we include it in Figure 1 because research has shown that arousal and the sensory experience related to receiving a service or purchasing a product contribute to satisfaction ratings (Mano & Oliver, 1993). In

<sup>5</sup> Schema are cognitive frames of reference for categorizing information and interpreting life experiences.

fact, as shown by the blue arrows in the figure, affect (e.g., ranging from delight to anger in appraising a service experience) can have both a direct effect on satisfaction and an indirect effect that operates through disconfirmation (Szymanski & Henard, 2001).

The way in which affect might influence satisfaction is social welfare is scarcely understood. However, affect would clearly be related to the holistic experience of being a client in an agency. One aspect of affect may be, for example, the extent to which a client feels involved with and understood by workers. If affect has an effect on satisfaction that is independent of service content, then worker competence and skill might matter beyond the active ingredients of service. For example, in a service designed to provide housing assistance, satisfaction is likely to be related not just to a housing-related outcome (e.g., whether a dispute with a landlord is resolved) but it will also be related to the courtesy of workers, the speed of processing, the nature of the waiting experience, and, more generally, with the “pleasant-ness” of the service experience. In situations where consumers are highly involved with service providers, affect may be as important as disconfirmation (Krampf, Ueltschy, & d’Amico, 2002). This could explain why attention-only conditions in social welfare studies sometimes earn comparatively high satisfaction ratings (for a discussion, see Ingram & Chung, 1997).

In sum, independent of the impact of a particular service strategy, satisfaction might be elevated because clients find workers to be friendly, supportive, and caring. We often regard these as common or, perhaps, foundational elements of professional practice and, indeed, they are. At the same time, however, as the evidence base for services grows, specific intervention strategies are used. Satisfaction ratings of these more focused intervention strategies will be confounded with foundational aspects of practice. *Because of this, satisfaction measures should focus on elements of services and produce ratings that are tied to distinct practice strategies as well as the overall service experience. Both must be measured.*

**Equity.** Similar to affect, the early versions of the Expectancy-Performance Model also ignored equity. Research has shown that satisfaction is influenced by perceptions of norms related to a service, including cost. That is, equity influences satisfaction by introducing fairness judgments (also indicated by blue arrows in Figure 1). These judgments are based on perceptions of what others have received in terms of both the physical quality of a service (or a product) and the interactional aspects of a service experience. In a meta-analysis of 50 published and grey literature studies focused on customer satisfaction in business, Szymanski and Henard (2001) found satisfaction had significant correlations with expectations ( $r = .27$ ), performance ( $r = .34$ ), disconfirmation ( $r = .46$ ), affect, ( $r = .27$ ), and equity ( $r = .50$ ). The findings showed equity had the highest association with customer satisfaction and, in regression analyses, equity’s influence was surpassed only by disconfirmation.

As consumer satisfaction came to be viewed as a complex construct determined by multiple factors, the construct also came to be considered as a legitimate, independent service outcome in social welfare. The consumer movement in mental health pressed for expanding the role of clients in

planning and evaluating services. The emergence of evidence-based practice in medicine encouraged physicians to involve patients in making treatment decisions and appraising treatment outcomes (Sackett, 1997). Related to a growing emphasis on quality assurance and evaluation, consumer satisfaction began to be perceived as an appropriate and clinically justifiable outcome for social welfare programs.

### **Growing Concern: Does High Satisfaction Mean Social Welfare Services Were Effective?**

However, amidst growing appreciation for the complexity of consumer satisfaction and its adoption in mental health, substance abuse, and other sectors of social welfare, critics urged caution (e.g., Lebow, 1982, 1983). In 1984, British social work scholar Ian Shaw warned:

First, satisfaction with service may often co-exist with criticism of that service. Second, satisfaction verdicts are influenced by the frame of reference within which they are reached. Third, judgments about satisfaction need to be clearly distinguished from judgements about the success or otherwise of welfare services. (pp. 279-280)

Shaw argued that personal satisfaction with service could exist concurrently with a critical appraisal. Precisely because satisfaction is multi-dimensional, global measures of satisfaction could mask both positive and negative appraisals. Shaw maintained that consumer judgments are contextually conditioned, emotionally influenced, and unreliable. Using data from a 2-week follow-up of satisfaction ratings in brief family therapy, Shaw found only 66% test-retest agreement on a satisfaction item ("satisfied/very satisfied with service") and 55% agreement on the effect of therapy ("attribute change at least in part to treatment"). Shaw (p. 283) called the use of consumer satisfaction ratings "problematic." He pressed for consideration of the psychometric properties of satisfaction scales and developing a better understanding of the relationship between satisfaction and more objective service outcomes.

Others also voiced concern about using consumer satisfaction as an outcome in social welfare (e.g., Lebow, 1982, 1983). Sol Garfield (1983, p. 238), a U.S. psychologist, cautioned the field about overemphasis on consumer satisfaction: "[there could be] a possible problem if we get overly concerned with matters of consumer satisfaction at the expense of other measures of outcome, particularly, behavioral change." At the U.S. National Institute of Mental Health, Morris Parloff (1983, p. 245) expressed doubt that satisfaction could become a useful outcome because "the majority of consumers appear to be satisfied with any and all services provided." Concurring with Parloff, Garfield (1983, p. 241) argued that, "Gullible people may be satisfied with the services of charlatans, and the history of the placebo response is well known." Concluding glumly, Parloff (p. 246) called the growing use of consumer satisfaction an "unfortunate ... stratagem."

Still others charged that satisfaction ratings were unrelated to changes in the symptoms of clients. From this perspective, satisfaction ratings lack suf-

ficient precision and specificity to be used as an indicator of the impact of a service (see, e.g., McNeill, Nicholas, Szechy, & Lach, 1998). In a review, Williams (1994, p. 515) concluded, "The original motivation behind satisfaction surveys was to introduce some element of consumerism and accountability to health care; however, through high levels of relatively meaningless expressions of satisfaction an illusion of consumerism is created which seldom does anything but endorse the status quo." Specifically, after studying 176 outpatient clients at three urban mental health centers in the U. S. Midwest region, Pekarik and Wolff (1996) observed no significant correlations between satisfaction and symptom reduction. They concluded (p. 206), "...satisfaction is not meaningfully related to traditional client measures of outcome."

In recent years, the criticism of satisfaction has continued with little interruption. For example, Schraufnagel and Li (2010) focused on two methods of establishing child support orders intended to ensure that dependent children receive both financial and medical assistance. They tested traditional court processing versus mediation procedures. Their comparison of people who went through routine court processing with those who went through mediation showed satisfaction was significantly higher for the mediation condition. However, the groups did not differ on compliance with the child support orders. Schraufnagel and Li found that satisfaction was not correlated with the protection of children through child support agreements, supporting Garfield's implicit suggestion that client satisfaction ratings have low concurrent validity with concrete outcomes.

Overall, the findings regarding the validity of consumer satisfaction ratings are mixed. Conclusions seem to depend on what outcome is used as an indicator of effectiveness. For example, Lunnen, Ogles, and Pappas (2008) compared satisfaction, symptom change, perceived change, and end-point functioning in a study of 66 clients in two community-based mental health centers in the U.S. Midwest. Clients were referred for mood or anxiety disorder (61%), adjustment or stress disorder (16%), substance abuse (7%), personality disorder (7%), and other diagnoses (10%). Participants rated symptomatic change, perceived change (amount of change), end-point functioning (e.g., *On the whole, how well do you feel you are getting along now?*), and satisfaction. The study findings revealed that although satisfaction was not significantly related to symptomatic change, satisfaction was significantly related to perceived change ( $r = .826$ ) and end-point functioning ( $r = .334$ ). The authors speculated the significant relationship found for end-point functioning was due to a method variance. That is, satisfaction, perceived change, and end-point functioning involved a cumulative retrospective assessment that lacked the specificity of a symptom checklist. Lunnen et al. concluded:

On the surface, it seems reasonable to assume that a person reporting significant symptomatic reductions would concurrently report higher levels of satisfaction with services when compared with a person who reported little or no symptomatic improvement.... The present results demonstrate that this assumption may not be empirically supportable. (p. 148)

Finally, from a policy orientation, scholars adopting a Public Value perspective have challenged as limited the frameworks related to consumer satisfaction that were invoked under New Public Management (Moore, 1995). In essence, they view the business conceptualization of *consumer* as insufficiently descriptive of the roles clients and, indeed, citizens should fulfill in the design and evaluation of social welfare services. Stoker (2006, p. 56) argued, for example, that the NPM perspective reduces citizens “to mere consumers, not allowed to question the objectives of service delivery but only encouraged to comment on their quality.”

Even as scholars raised concern about the consumer perspective and the validity (e.g., the relation of satisfaction scores to symptom change) and reliability (e.g., stability over time) of consumer satisfaction measures, other scholars set about developing satisfaction measures for social welfare. Among the first to be developed, the Client Satisfaction Questionnaire (CSQ-8; Nguyen, Attkisson, & Stegner, 1983) and the Reid-Gundlach Social Service Satisfaction Scale (R-GSSSS; Reid & Gundlach, 1983) were published in 1983. These scales were tailored for general use, and premised on the position that satisfaction is one of many potentially important program outcomes.

## Part 2. Current Conceptualizations of Consumer Satisfaction

---

Since 2000, the development of consumer satisfaction has followed three lines of inquiry. First, consumer satisfaction has been used in social welfare and business as a transaction-specific measure, which when summed indicates satisfaction with the quality of a particular service or product. This work has focused on designing satisfaction scales, particularly multi-item scales, and on examining the determinants of satisfaction scores.

Second, starting in the 1990s, interest grew in creating a cumulative measure of consumer satisfaction across industries in the private and public sectors. Almost entirely outside social welfare, this work did not focus on specific transactions because consumers could have multiple transactions with a service or product in a given period. Satisfaction ratings were based on perceptions across all transactions within service or produce categories. With the creation of the Customer Satisfaction Barometer in 1989 (Fornell, 1992), Sweden became the first country to measure satisfaction at the national level. Assessing cumulative satisfaction across 130 companies in 32 sectors of the economy, Sweden influenced the development of the American Customer Satisfaction Index (ACSI) and the European Customer Satisfaction Index (ESCI).

Finally, although work on the Expectancy-Performance Disconfirmation Model came largely from business and economics, applications of the model to a variety of government and government-regulated services (e.g., public utilities) have begun to appear in the literature. To our knowledge, no transaction-level applications of the disconfirmation perspective have been reported in social welfare. However, researchers have used ACSI data in applications that model local, state, and federal public services. Together, these three streams of investigation have important implications for conceptualizing client satisfaction in social welfare.

### **Multi-Item Measures of Consumer Satisfaction in Social Welfare**

Criticism of global scales, which focus on service as a one-dimensional construct, prompted efforts to design measures with greater complexity that could be used to rate satisfaction within elements of services. The business field supplemented single-item satisfaction scales (i.e., *In general, how satisfied are you with Service X?*) with multi-item satisfaction scales that assessed various aspects and degrees of satisfaction (i.e., *How satisfied are you with the performance of Service X regarding x, y, and z?*). In addition to scoring services or products per se, multi-item scales divided services into alternative facets of the purchasing or service event. These facets included the availability and clarity of information, the helpfulness of personnel or staff, and the courtesy of customer services in resolving complaints. Finally, entirely new measures, such as the Net Promoter® Score (NPS; *How likely*

*are you to recommend Service X to a friend, relative, or colleague?*), were developed.

The marketing literature is rich with efforts to test these new, more complex measures. For example, a Dutch research team used an Internet survey of 11,967 customers in the banking, insurance, utilities, and telecom industries to compare satisfaction ratings gathered via single-item, multi-item, and NPS scores (van Doorn, Leeflang, & Tijs, 2013). Van Doorn et al. estimated the effect of each type of satisfaction rating on current and future sales revenues, growth margins, and cash flow in 46 companies. Although they had hoped to discover differences, all three metrics performed equally well in predicting current gross margins and cash flow. However, none of the metrics was predictive of future (12-month) performance. In this test, each perspective on consumer satisfaction had concurrent but not predictive validity. Thus, from the van Doorn et al. study and given the added burden of multi-item and NPS ratings, it appears that new metrics confer no advantage over traditional metrics.

In social welfare, work has also focused on developing multi-item scales, but these efforts have been less tied to outcomes. Nonetheless, services have been broken down into elements that have, in turn, been rated. As in business, rather than averaging across all items to produce a single satisfaction score, scores have been summed or averaged within aspect areas. Extending work in business psychology in the Norway, Sweden, and the United States (e.g., Johnson, Gustafsson, Andreassen, Lervik, & Cha, 2001), researchers have attempted to weight service aspects to pinpoint relatively more influential facets of service that, if scores were low, might warrant redesign.

One example of the efforts to develop scales that hold the potential to tease apart aspects of satisfaction is the Client Satisfaction: Case Management (CSAT-CM) scale that was developed by Chang-Ming Hsieh (2006) at the University of Illinois in the United States. The CSAT-CM permits ranking and then assessing various elements of services. Hsieh designed the scale using the concept of overall satisfaction as a function of satisfaction scores within various aspects of service, which were ranked for importance. Hsieh designated five elements of service, two of which assessed the service performance (assessment of client needs, development of the plan of care) and three that assessed the case manager's performance (case manager's knowledge of available services, case manager's ability to get services for clients, and the availability of the case manager). The CSAT-CM was tested using a sample of 112 older adults ( $M = 76.4$  years of age,  $SD = 7.3$ ), and satisfaction and importance ratings were obtained for each of the five elements. Using a score weighted by importance, test-retest reliability and concurrent validity were estimated. The CSAT-CM had acceptable test-retest reliability ( $r = .81$ ) and concurrent validity with the CSQ-8 ( $r = .70$ ). The CSAT-CM approach demonstrated that multi-item weighted scales have adequate psychometric properties. If using scores within aspects of service were to provide a focus for refining service strategies, the multi-item, multi-dimensional approach would have potential impact greater than one-dimensional satisfaction scales (see also, Hsieh, 2012).



Taken together, the Hsieh (2006, 2012) and van Doorn et al. (2013) studies represent contemporary attempts to improve transaction-specific satisfaction measures. Largely developed for use in practice and practice research, these scales test new ways to measure consumer satisfaction. This work is informed by the Expectancy-Performance Disconfirmation model, in which the determinants of satisfaction are conceptualized as a partial function of an appraisal of the performance of a service or product. However, in social welfare, the work is psychometrically oriented and has rarely moved from testing measurement models to testing theoretical models.

### National Models of Customer Satisfaction: The Swedish Barometer and Beyond

During the same period in which one stream of work focused on transaction-specific satisfaction measures, a second stream of work focused on cumulative ratings of satisfaction. This perspective sought to score satisfaction across a consumer's aggregate experiences with making a purchase or receiving a service. These efforts superseded the transaction-specific orientation and focused on estimating an overarching satisfaction score. In breaking away from a specific event as the anchor in measurement, researchers were forced to ask consumers to rate average performance (e.g., service experiences) against a perception of ideal performance.

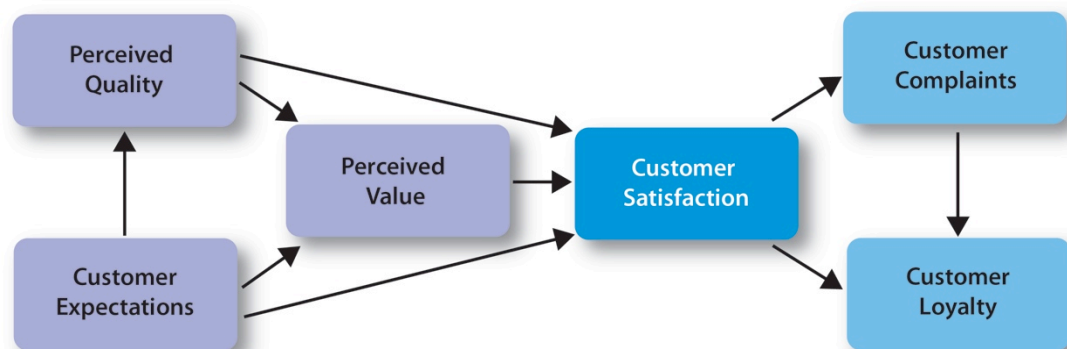
The first of these models was the Swedish Customer Satisfaction Barometer (SCSB; Fornell, 1992). Shown in Figure 2, the initial Swedish model introduced *perceived value* as a performance measure. That is, performance was gauged as “value for the money” based on recent experiences with a product or service. This model defined expectations as beliefs about how well a product or service *will* perform, and thereby avoided normative expectations (e.g., how well it *should* perform). *Customer loyalty* was defined, in part, as the willingness to purchase the same product or service again (i.e., repurchase), and *customer complaints* referred to the courtesy, speed, and consumer orientation of personnel who manage item returns and other product or service problems.

**Figure 2.** The Swedish Customer Satisfaction Barometer



The SCSB was the basis for the American Consumer Satisfaction Index (ACSI). Indeed, Claes Fornell, now a professor at the University of Michigan in the United States, founded both indices (Fornell et al., 1996). Shown in Figure 3, the ACSI differed from the SCSB in that perceived (performance) value was predicted by perceived quality and expectations. Perceived quality involved assessing the physical characteristics (including reliability) and interactional aspects of product or service experiences. In both the Swedish and American models, customer loyalty is measured by willingness to repurchase and by the amount of change in price that would be required to alter a repurchase decision.

**Figure 3.** *The American Customer Satisfaction Index (Fornell et al., 1996).*

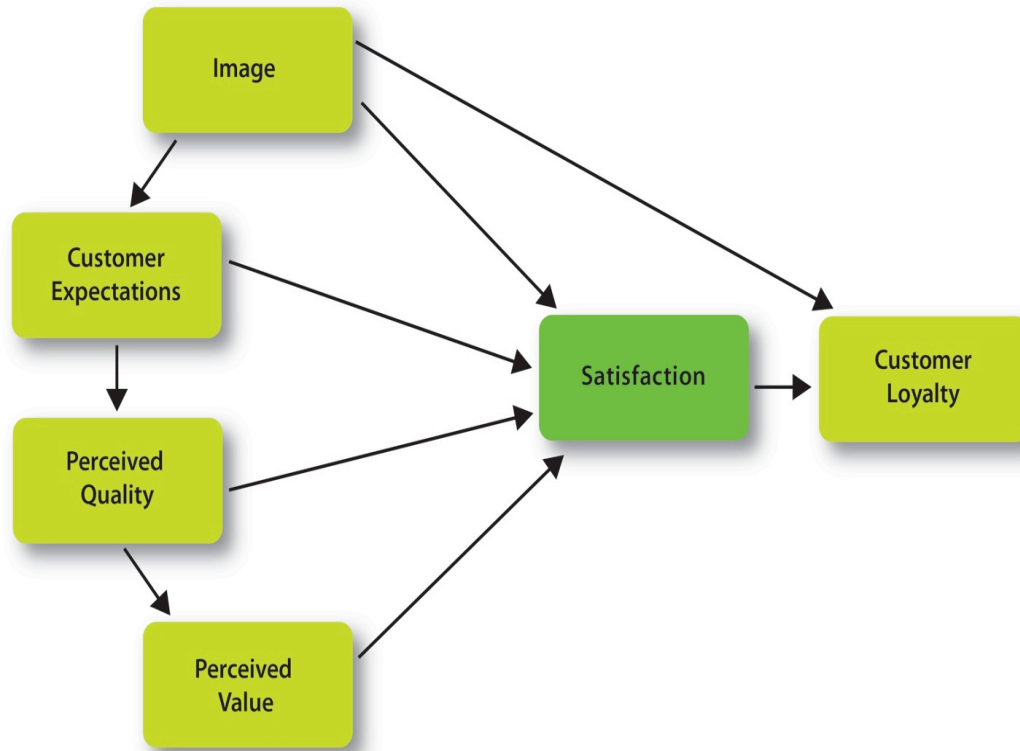


The European Union's ECSI was based, in part, on the ACSI. However, the ECSI was distinguished from the ACSI by the deletion of customer complaints and the addition of an indicator for corporate image (see Figure 4). This additional indicator was adapted from the Norwegian Customer Satisfaction Barometer (Andreassen & Lindestad, 1998), which conceptualized corporate image as having direct effects on consumer expectations, satisfaction, and loyalty.

The ECSI argues implicitly that the image of a service provider affects client satisfaction ratings (see, e.g., Martensen, Gronholt, & Kristensen, 2000). Research on the relationship of image to satisfaction has provided mixed results, and it is rare in social welfare, although studies of public utilities and services may have relevance. One such study focused on satisfaction with postal services on the Isle of Man in the United Kingdom. O'Loughlin and Coenders (2002) tested the ECSI model, including image (e.g., *Are postal services reliable, trustworthy, customer oriented, and do they provide a valuable service?*), with 280 residents selected in a simple random sample. Image was expected to predict perceived value, consumer satisfaction, and loyalty. However, the study findings showed image was related only to loyalty. Similarly, in a study of 551 banking consumers in Tehran, a research team found no significant relationship between image and satisfaction (Hamidizadeh, Jazani, Hajikarimi, & Ebrahimi, 2011). On the other hand, in a Thai study of 276 mobile phone users in Bangkok, research showed image was significantly related to satisfaction, when the

analysis controlled for perceived quality, emotional value, promotional advertising, and other factors (Leelakulthanit & Hongcharu, 2001).

**Figure 4.** *The European Consumer Satisfaction Index*



Perhaps the largest test of the ECSI (and the inclusion of image) was undertaken by the European Commission (2007, p. 19), which conducted a study of consumer satisfaction across 11 “services of general interest” (e.g., electricity supply, water distribution, transport [urban, extra-urban, air], postal services, and insurance) in 25 countries. More than 29,000 consumers were randomly sampled and interviewed face-to-face in their homes. Scored on a 10-point scale ranging from *low* (1) to *high* (10), satisfaction was defined as the extent to which a product or service “has met ... needs or expectations” (p. 8). Survey participants reported higher levels of satisfaction with air transport, mobile phones, insurance service, and retail banking; they reported less satisfaction with utilities, urban transport, and extra-urban transport.

Survey participants used the same 10-point response scale to rate quality, image, and pricing. Image was defined as a provider’s reputation for customer mindedness, technological innovation, and environmental sensitivity. Although pricing tended to be the main driver behind satisfaction scores, significant variation was observed across sectors. Price drove satisfaction scores for insurance, electrical utilities, retail banking, fixed telephone, mobile phones, and water distribution. In regression models, image drove satisfaction for postal services, urban transport, and extra-urban transport. In short, the Commission found that image was related to consumer satisfac-

tion in some sectors of the European economy but unrelated to consumer satisfaction in other sectors. The data suggest that *image may be particularly important for public services that require frequent interpersonal interactions with the service providers*.

This finding implies that the image of a provider could influence consumer satisfaction scores in social welfare. That is, the satisfaction ratings of a service provided by a well-known agency (e.g., Karolinska Institutet or, in the U.S., the Mayo Clinic) might garner higher satisfaction ratings given the influence of public trust and reputation. In the sense of the ECSI, the image of the source or auspices of services has the potential to influence satisfaction ratings beyond the actual experience or effect of the service.

**Relevance to social welfare of national customer satisfaction indices.** The efforts to develop the SCSB, ACSI, and ECSI are potentially useful in informing the conceptualization of consumer satisfaction in social welfare. In these models, customer satisfaction is a proximal outcome. On balance, satisfaction is predicted by consumer expectations, the quality of service experiences, and the perceived value of those experiences. In the ECSI, image is seen as differentially influencing expectations and directly contributing to satisfaction ratings.

If the ECSI assumptions about image are correct, then image is a potential confounding factor in using client satisfaction as an indicator of service effectiveness in social welfare. Assuming that two services are comparable in effectiveness, satisfaction scores would be expected to differ based on the influence of the image associated with the service provider (e.g., one service might be provided by a well-established, widely known agency while another is provided by a newly formed agency). If so, *client satisfaction in social welfare cannot be a valid indicator of the quality of services offered across providers whose images vary*.

### **Consumer Satisfaction as a Mediator in Social Welfare**

A third line of inquiry complements the work on multi-item scales and the development of national customer satisfaction indices through its focus on testing theoretical frameworks. By placing consumer satisfaction in causal models, this line of inquiry represents a conceptual advance that is just beginning to penetrate social welfare. To be sure, the focus in scales like the CSAT-CM on the active ingredients of services suggests that satisfaction may be conceptualized as a *mediator* of distal outcomes. Sometimes called an intervening variable, a mediator conveys the effect of a prior variable, such as expectations, on a more distal variable, such as loyalty. For example, as shown in Figure 4, image is hypothesized to operate on customer loyalty in at least three ways. Image has a direct effect on loyalty and a mediated effect on loyalty through satisfaction. It also has secondary effects that operate through expectations. With advances in structural equation modeling, the “fit” of these theoretical models can now be tested in research.

A promising line of inquiry involves work on conceptualizing client satisfaction as a mediator in social welfare. In a study of post-treatment abstinence among 208 clients of an outpatient alcohol treatment in Buffalo, NY (USA), researchers used a modified version of the CSQ-8 to estimate a

model that included service expectations, the client/therapist alliance, and session attendance (Dearing, Barrick, Dermen, & Walitzer, 2005). The effect of all three measures on post-treatment abstinence was mediated fully by CSQ-8 satisfaction. After testing a number of alternative models and observing good fit, Dearing et al. concluded, "Positive expectations about therapy, greater session attendance, and a positive perception of the working alliance appeared to predict greater client satisfaction and, in turn, more positive drinking-related outcomes" (p. 75). A mediation model in which client satisfaction predicted post-treatment abstinence was supported.

A common hypothesis related to client satisfaction is that program participants who have been more engaged in services will be more satisfied. This hypothesis was tested in a trial of Assertive Community Treatment (ACT) with 191 homeless adults who were diagnosed with serious mental illnesses (Fletcher, Cunningham, Calsyn, Morse, & Klinkenberg, 2008). After giving informed consent, participants were randomly assigned to integrated ACT (IACT), ACT only (ACTO), or Standard Care (SC), and each client was interviewed at 3, 15, and 30 months after treatment assignment. During the interviews, data were collected on satisfaction with services (using a multi-item scale), housing stability, substance abuse, psychiatric symptoms, and a variety of measures of participation in IACT, ACTO, and SC. Participants in IACT and ACTO reported significantly greater satisfaction and more stable housing arrangements. No significant differences were observed between IACT and ACTO. Although satisfaction received high ratings during the early stages of service participation, it declined over time for all three conditions. However, satisfaction was predicted by program contact (i.e., number of days per month in contact with ACT staff), assistance with daily activities, help with emotional problems, transportation support, and substance abuse referral. Although no effects were observed on the important outcomes of substance abuse or symptoms, the findings suggest that satisfaction with services covaried with housing stability; moreover, greater treatment engagement predicted higher service satisfaction. While rare in social welfare, mediation studies such as those conducted by Dearing et al. (2005) and Fletcher et al. (2008) are beginning to provide an empirical context for conceptualizing satisfaction with services as a function of treatment adherence (see also, Hawkins, Baer, & Kivlahan, 2000).

Although more research is needed, satisfaction is probably best conceptualized as one of several proximal outcomes (e.g., including theory-relevant outcomes such as changes in knowledge and skill) that might, in turn, mediate distal outcomes. To be sure, mediation was implied in the initial logic of satisfaction with social welfare services. That is, greater service engagement was expected to predict satisfaction. Implicitly, satisfaction was thought to link to behavioral outcomes that would follow from greater treatment adherence. However, in the 1980s and 1990s, satisfaction began to be viewed as a sufficient outcome on its own. Critics, such as Parloff (1983), argued that the use of satisfaction was driven by its convenience relative to the challenges of developing behaviorally focused outcomes (e.g., the measure of housing stability used by Fletcher et al., 2008).

In much the same way that customer satisfaction in business has been linked to concrete outcomes such as gross margins, client satisfaction can be

useful in understanding the causal mechanisms that operate in social interventions to produce distal (longer term) outcomes. However, it is not clear that satisfaction with services is a *necessary* component of the causal logic that underlies intervention models. Indeed, some studies suggest that satisfaction is, at best, only weakly related to symptomatic reductions and other behavioral outcomes (Lunnen et al., 2008). Nonetheless, a few areas of public service have applied the Expectancy-Performance Disconfirmation model to evaluations of “citizen” satisfaction with governmental programs. The findings from these studies have suggested that satisfaction might emerge as a link in the causal chains that explain program effects.

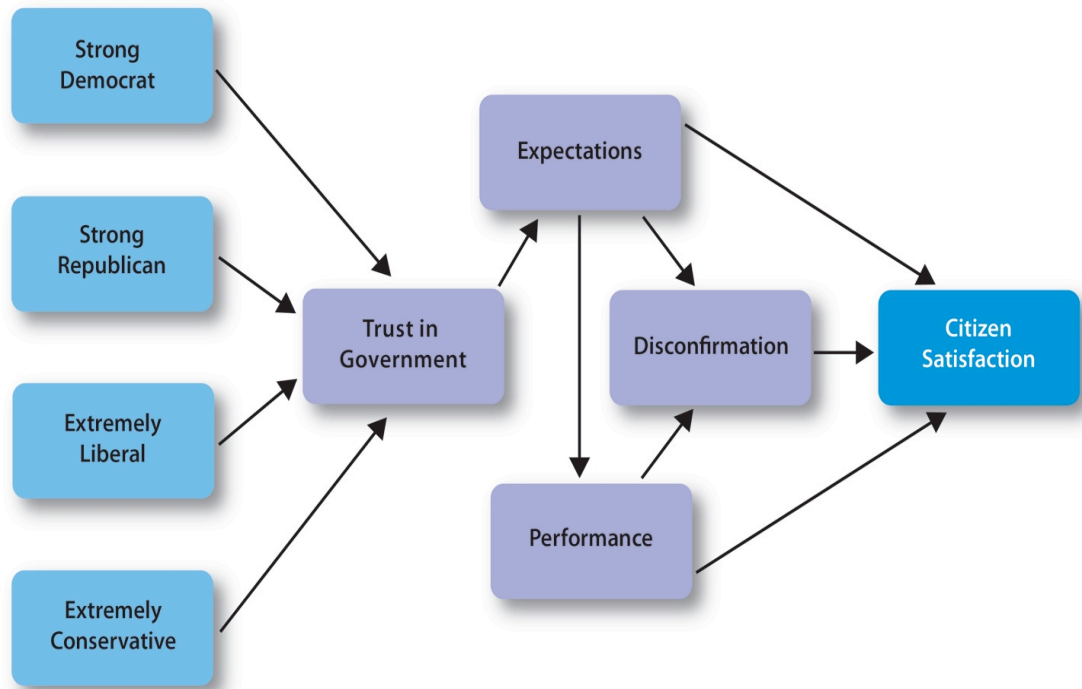
### **Applying the Disconfirmation and ECSI Perspectives to Satisfaction with Public Services**

Studies of citizen satisfaction with public services in areas such as policing and fire protection have included measures of expectations and performance (see, e.g., James, 2009; Poister & Thomas, 2011; Van Ryzin, 2006). For example, in 2013, Morgeson applied the Expectancy-Performance Disconfirmation model (EPDM) to ratings of satisfaction with United States federal government services. Citizen satisfaction was measured via responses to items such as, *Considering all your experiences to date, how satisfied are you with <x>’s services?* Using ACSI data from 1,480 consumers who reported service experiences with the government (i.e., any program or department, with the exception of postal services), Morgeson first tested the EPDM, and then expanded the EPDM to include measures of governmental trust.

The conceptualization of trust in satisfaction models has varied over time. In an earlier report, Morgeson and Petrescu (2011) used trust (i.e., *How much of the time do you think you can trust the government?*) as a function of satisfaction. That is, trust was seen as deriving from – as being an outcome of – satisfaction with public services. However, across six domains of service (e.g., health care, veteran’s affairs, and social security [pensions]) and with the exception of the Internal Revenue Service, Morgeson and Petrescu found satisfaction was not correlated with trust. In a later report, Morgeson (2013) re-conceptualized trust as influencing expectations. Thus, trust in public services appears to share some similarities with the image indicator in the ECSI in that both image and trust are thought to impact satisfaction by influencing expectations.

Morgeson’s findings (2013) were illuminating. At the zero-order level, satisfaction with public services, including health care and social security, correlated significantly with expectations of quality ( $r = .543$ ), performance ( $r = .863$ ), confirmation/disconfirmation ( $r = .832$ ), and trust ( $r = .360$ ). In structural equation models, the EPDM was confirmed. See Figure 5. Not shown in Figure 1 (because it is not consistently observed), a weak but significant path ( $\beta = .08$ ) was found from expectations to satisfaction. This path suggests that expectations have both direct effects and mediated effects, through performance and disconfirmation, on satisfaction with public services.

**Figure 5.** Test of the Disconfirmation Model with ACSI Data on Public Services in the United States. Note. Figure is adapted from Morgeson (2013).



In supplemental analyses that controlled for political party and ideology (shown on the left in Figure 5), trust was added to the model as a predictor of expectations. The findings showed trust was related significantly to expectations ( $\beta = .32$ ), but trust accounted for only 10% of the variance. Findings such as these begin to support the view, manifest in the ECSI, that image and trust in the agency providing services influence expectations, performance appraisal, disconfirmation, and satisfaction through a chaining of effects.

## Part 3. The Psychometric Characteristics of Consumer Satisfaction

---

Consumer satisfaction is often conceptualized as a proximal or near term program outcome (e.g., Royse, Thyer, & Padgett, 2010). That is, it is viewed as closely related to the immediate effects of program participation. This requires conceptualizing services in theoretical models. The specific aspects of a service (e.g., teaching a particular parenting strategy) may be described in a program theory (i.e., the theory of change for why a particular program is expected to work), and satisfaction may be one aspect of a chain of factors leading to longer term and, often policy-relevant, outcomes (Berghofer, Castille, & Link, 2011). In a causal model, then, consumer satisfaction is predicted *by* program content, and it is predictive *of* distal outcomes.

In 2000, the World Health Organization published a workbook on *Client Satisfaction Evaluation* in which the authors argued that using consumer satisfaction gives program participants a voice in program evaluation. Although including participant views is widely acknowledged as a feature of responsible professional practice and program evaluation, many evaluations gather participant feedback through self-reports of behavior. Compared to satisfaction reports, these provide more specific feedback on the social and health problems of program participants. What then is the proper place of client satisfaction?

Consistent with the WHO (2000), satisfaction may include assurance that services are provided in a consistent and dependable manner, that services are responsive to needs, and that providers are courteous. In this sense, the construct validity for using satisfaction rests on its theoretical relevance (i.e., specifying how program content is related to satisfaction and other outcomes) as well as the empirical findings showing that satisfaction is correlated with health and welfare outcomes, such as reductions in depression, declines in drug use, and stability in housing. As a function of service-related characteristics, *satisfaction should be helpful in understanding the success and failure of programs.*

In this section, we summarize research findings on satisfaction with services in social welfare. We focus on two types of validity: concurrent validity and predictive validity. When a measure is correlated contemporaneously with another measure and when the correlation makes sense in terms of theory, the two are said to have concurrent validity. If the original measure occurs prior in time and an argument can be made that it is a marker for the subsequent outcome, the former measure is said to have predictive validity. In this section, we explore the concurrent and predictive validity of client satisfaction.



## **Client Satisfaction as an Indicator of Program Processes in Social Welfare**

Although satisfaction with services has been discussed for many years in social welfare, the empirical literature is weak. Satisfaction is often used as an outcome in studies where there is no counterfactual. As a result, it is not possible to estimate satisfaction in a control group, which would provide a measure of satisfaction for participation in routine or other services. Moreover, because consumers in voluntary programs often appreciate any attention, satisfaction ratings tend to be high for all service participation (Ingram & Chung, 1997; McNeill et al., 1998). An attention-control condition is needed if the purpose of using satisfaction is to gauge reactions of program participants to a specific intervention as opposed to the receipt of attention only. Further, most measures of satisfaction are conducted at posttest only. Because consumers who drop out of programs are more likely to be dissatisfied, the use of satisfaction without conducting exit interviews with dropouts produces a potential attrition bias (Gottlieb & Wachala, 2007). *Using satisfaction as a program outcome has little validity if only participants who are so satisfied as to stay through program completion report satisfaction ratings.*

Most studies of consumer satisfaction are cross-sectional, and dropout plus length bias are threats to inferences. Length bias refers to the phenomenon that length of stay in a program is often correlated positively with satisfaction because consumers who are dissatisfied with a program tend to drop out. To control for length bias, a few studies have introduced program participation as a covariate in analyses. For example, Heinze, Jozefowicz, and Toro (2010) used consumer satisfaction as a dependent measure to assess the influence of a variety of program characteristics on 133 adolescents and young adults who were involved in services intended to reduce homelessness in a Midwestern U.S. city. Across six programs, a majority (78%) of the youth had attended a program for at least 1 week, and 47% had been involved for at least 6 months. Satisfaction was scored on an 11-item scale (e.g., *I am satisfied with the amount of help I receive at <x>*), and nine program processes were measured with multi-item scales. Program dropouts were not followed, and, suggesting an attrition effect, length of stay (i.e., months in program) was positively correlated with satisfaction ( $\beta = .21$ ). That is, the youths who were more satisfied with the program spent longer periods of time in the programs, whereas the youths who were less satisfied with the program spent less time in the programs.

This *length of stay* effect became nonsignificant as program features (i.e., clear expectations, supportive staff relationships, sense of belonging, and encouragement for skill development) were entered into regression models. Explaining 68% of the variance in satisfaction, program characteristics appeared to control for the greater satisfaction reported by longer-term program participants. Heinze et al. (p. 1371) concluded, "...findings support balancing caring, supportive staff-participant relationships with clear expectations and limit setting, while providing age-appropriate opportunities for strength identification, skill building and personal growth." Using satisfaction as a dependent measure appeared to produce important program in-

sights. Acknowledging the need for more distal outcomes, Heinze et al. cautioned:

Longitudinal research examining participants within agencies over time is needed to clarify temporal relationships among program dimensions and assess how dimensions and satisfaction ratings impact other outcomes of interest, such as length of stay, skill development, reintegration with families, school and occupational success, and a successful transition to independent living. (p. 1370)

Like others, Heinze et al. (2010) were beginning to use satisfaction with services in a causal modeling perspective. The study is promising, in part, because the researchers developed detailed measures of program processes. The findings showed a strong association between satisfaction and program processes. At the same time, the need for behavioral measures, such as successful transition to independent living, is clear and satisfaction alone is not regarded as an adequate outcome.

In other research, satisfaction with services also appears to be predicted by program processes (e.g., Dearing et al., 2005; Fletcher et al., 2008). For example, Kivlighan, London, and Miles (2012) examined the relationship between satisfaction and group leadership in a study of the *Choices Independent Living Program*, an intervention comprised of structured exercises, didactic materials, and discussion. Thirteen *Choices* groups were led by only one therapist, and 19 groups were co-led by a therapy team. Referred for a variety of conduct problems, 87 boys and 89 girls ages 13 to 15 years old were assigned to one group leadership condition or the other. At end of each of the eight weekly sessions, data on a variety of group-process measures were collected, and, at the conclusion of treatment, the youth completed the Youth Client Satisfaction Questionnaire (YCSQ; Shapiro, Welker, & Jacobson, 1997). Satisfaction was related significantly to leadership and other group process indicators (e.g., climate). Youths in co-led groups reported more *benefits from therapy*, a dimension of the YCSQ. Although youths were not assigned randomly to groups and the study contained no discussion of attrition, subjective causal appraisals of perceived program effects distinguished singly led from co-led groups.

Most studies of satisfaction use single-group or nonrandomized multi-group designs and, many report relationships between service engagement and client satisfaction (e.g., Trute & Hiebert-Murphy, 2007). In a study of 88 adult male sex offenders who attended outpatient cognitive-behavioral treatment groups in Connecticut, USA, satisfaction was significantly correlated with treatment engagement ( $r = .54$ ; Levenson, Prescott, & D'Amora, 2010). In a retrospective survey of 131 family members of nursing home residents who had recently died of dementia in Massachusetts and Rhode Island (USA), analyses showed satisfaction with care to be significantly correlated family communication ( $r = .68$ ) and a comfort-orientation ( $r = .65$ ; Liu, Guarino, & Lopez, 2012). Although design issues such as attrition confound inferences, the data often suggest that satisfaction is related to program processes.

## Consumer Satisfaction as an Indicator of Behavioral Outcomes in Social Welfare

Since Garfield (1983), Parloff (1983), Shaw (1984), Lebow (1982, 1983), and others challenged the use of consumer satisfaction, the most enduring question in the field has been whether satisfaction is associated with behavioral outcomes. To attempt to answer that question, it is necessary to compare the covariation of satisfaction with behavioral outcomes within studies. Webster-Stratton et al. (2004) found concurrent effects for satisfaction and reports of child behavior. Fletcher et al. (2008) found concurrent effects for satisfaction and housing stability; however, they found no concurrent effects for symptoms or drug use. Schraufnagel and Li (2010) found no relationship between satisfaction and compliance with child support orders. In a study of a shelter for homeless youth in Israel, satisfaction reported by 102 residents was related to adaptation to the program, but was unrelated to reasons for leaving the shelter (Spiro, Dekel, & Peled, 2009). In a small study ( $N = 19$ ) of patients in a pediatric unit in an Australian hospital, Walsh and Lord (2004) found no relationship between satisfaction (measured using the CSQ-8) and client reports of empowerment (measured as feeling empowered to change). Walsh and Lord concluded, "Satisfaction should not be used as an all-encompassing method of service evaluation or quality assurance" (p. 50). Skepticism has been widespread. So much so that in 2005, Weinbach warned, "...the major problem of using client satisfaction surveys as indicators of intervention effectiveness, or of quality of a service, is that satisfaction with services and successful intervention are not the same" (p. 38).

However, the findings may be conditioned on the specificity of satisfaction measures. In a study of 3,298 students in 23 Chinese secondary schools, Shek (2010) addressed the question directly. He compared self-reported responses to the Chinese Positive Youth Development Scale (CPYDS) with self-reported responses to the Chinese Subjective Outcome Scale (CSOS), which measured satisfaction with program attributes (e.g., *atmosphere of the class was good*), program implementation processes (e.g., *instructors could arouse my interest*), and perceived program effectiveness (e.g., *program can strengthen my ability to face challenges*). The total CPYDS score, which measured outcomes across 15 behavioral domains, was significantly correlated with all three satisfaction measures (respectively,  $r = .62$ ,  $r = .64$ ,  $r = .62$ ). In addition, controlling for Time 1 CPYDS scores, all three CSOS measures predicted Time 2 CPYDS scores. Shek (p. 299) concluded that there is "an intimate relationship between subjective measures of satisfaction and objective measures of behavior." *When constructed to measure specific program processes, satisfaction scores appear predictive of behavioral outcomes.*

In the same vein, Trotter (2008) tackled the issue of the concurrent and predictive validity of client satisfaction in child welfare. Through fraught with methodological problems, this work is important because many parents are involuntarily involved in child welfare services. It is not clear that satisfaction can be a relevant outcome when services are mandated and, though of high public value, when such services serve a social control rather than social care function (Martin, Petr, & Kapp, 2003). Trotter examined the relationship between caretaker reports of satisfaction, worker ratings of client progress, official reports of subsequent mal-

treatment events, and agency records of child placement in a sample of 205 families referred to child welfare authorities in Victoria, Australia. Although satisfaction ratings were not collected from all parents whose children were placed out of the home, satisfaction was weakly correlated with worker ratings of client progress, with subsequent maltreatment reports, and with out-of-home placement. The findings suggest that satisfaction may have concurrent and predictive validity (i.e., it varies consistently with other relevant outcomes) in fields in which choice is constrained and where agencies provide services that may be less subject to the inflated ratings observed when service participation is voluntary.

## Summary

As a program evaluation measure, satisfaction with services is evolving. Findings are mixed, and some studies where satisfaction is related to program features suggest that, as a measure, it has concurrent and predictive validity. Work on perceived change as an element of satisfaction with services is an important line of inquiry. These subjective causal appraisals (e.g., *my needs have been resolved as a result of the help I received.*) may be particularly useful when they are related to specific features or elements of services. At the cutting edge, researchers are using satisfaction as a mediator that conveys the effects of service properties to distal outcomes. On balance, however, *the literature is not sufficiently strong to draw conclusions about whether satisfaction is a necessary and sufficient component of mediation models.*

Although its value in designing interventions and its use in theories of change is not clear, satisfaction with services is consistent with prevailing views in public administration and professional practice that efforts are needed to give voice to consumers in the choice and evaluation of services. To that end, dozens of satisfaction scales have been developed. We review them in the next section.

## Part 4. Consumer Satisfaction Scales

---

Research suggests that consumer satisfaction is correlated with service engagement and, with substantially less certitude, with behavioral outcomes, such as subsequent offending and symptom changes. Although the data are not strong, studies also indicate that satisfaction is related to treatment adherence. However, it is not clear that satisfaction is needed in – indeed, whether it is a necessary component of – theoretical models of service outcomes. Models that use changes in skill or knowledge may be sufficient and may fit the data better than models that incorporate satisfaction with skills, knowledge, and other measures of the core features of interventions. Nonetheless, satisfaction has the potential to inform theories of change and provide an explanation as to how services engage and retain participants. Moreover, satisfaction might explain why participants adhere to protocols and observe particular outcomes. Much stronger research designs and more complicated data analyses are needed to sort out these complexities.

This part of the report presents a systematic review of consumer satisfaction scales. Undertaken in spring 2013, the aim of this review was to identify and describe all consumer satisfaction scales used in published evaluations of social welfare services, 2003-2013. Special attention was given to consumer or client satisfaction used in social and health behavior research that evaluated programs in aging, child welfare, criminal justice, developmental disabilities, education, housing, juvenile justice, mental health, and substance abuse.

### Methods

To identify studies meeting the aim of this review, we established four inclusion/exclusion criteria. First, studies were included if the research was related to consumer/client/patient/service user/psychiatric-survivor satisfaction and treatment/intervention/program satisfaction. This scope included studies in social work, sociology, psychiatry, psychology, and substantive areas such child welfare and mental health. Studies in business, economics, and medicine were excluded. Second, studies were included if satisfaction was reported as a process or outcome measure. Third, we included only studies published in English. The fourth criterion restricted inclusion to articles that were published in peer-reviewed journals between January 2003 and May 2013.

Restricting the sampling frame to this period was a pragmatic decision. Instruments for measuring satisfaction have been developed over the course of 30 years, and the social welfare literature on satisfaction has grown decade by decade. For example, when considering the nine databases used in this review and the “raw search” results (i.e., before duplicate reports were eliminated), expanding the 10-year inclusion period by even small increments had a dramatic effect on the sample size. Expanding the inclusion period by 5 years would have increased the number of included articles by

177, and another 5-year increase (i.e., literature published in the last 20 years) would have increased the included articles by 294, for a total of 675 articles. Given the practical limitations of conducting a thorough review, we chose to restrict the sample to the most recent decade of published literature. Further, most older instruments that are currently used to measure satisfaction have been refined or modified within the past 10 years. For instruments first developed before 2003 (e.g., CSQ-8, BIRS, and WAI-S,) the inclusion criterion limiting articles to those published in the past decade enabled us to focus on the more recent, more relevant iterations of older instruments.

**Search engines.** Shown in Figure 6, the following nine databases were used:

- Social Services Abstracts
- Social Work Abstracts
- Social Sciences Citation Index
- Sociological Abstracts
- PsycINFO
- ASSIA
- PubMed
- CINAHL
- Business Source Complete

**Search Terms.** The search code was: [("consumer satisfaction" OR "client satisfaction" OR "treatment satisfaction" OR "intervention satisfaction" OR "program satisfaction") AND (scale OR measurement OR questionnaire OR instrument OR evaluate OR evaluation OR assess OR assessment OR test OR measure OR reliability OR validity) AND ("social work" OR "mental health" OR psychology OR "social service" OR "social services")] NOT (medical OR medic\* OR business OR financial OR financ\* OR physical OR physic\* OR health OR commercial OR commerc\* OR customer OR patient)<sup>6</sup>

**Data collection procedures.** The search and data collection procedure is shown graphically in Figure 6.

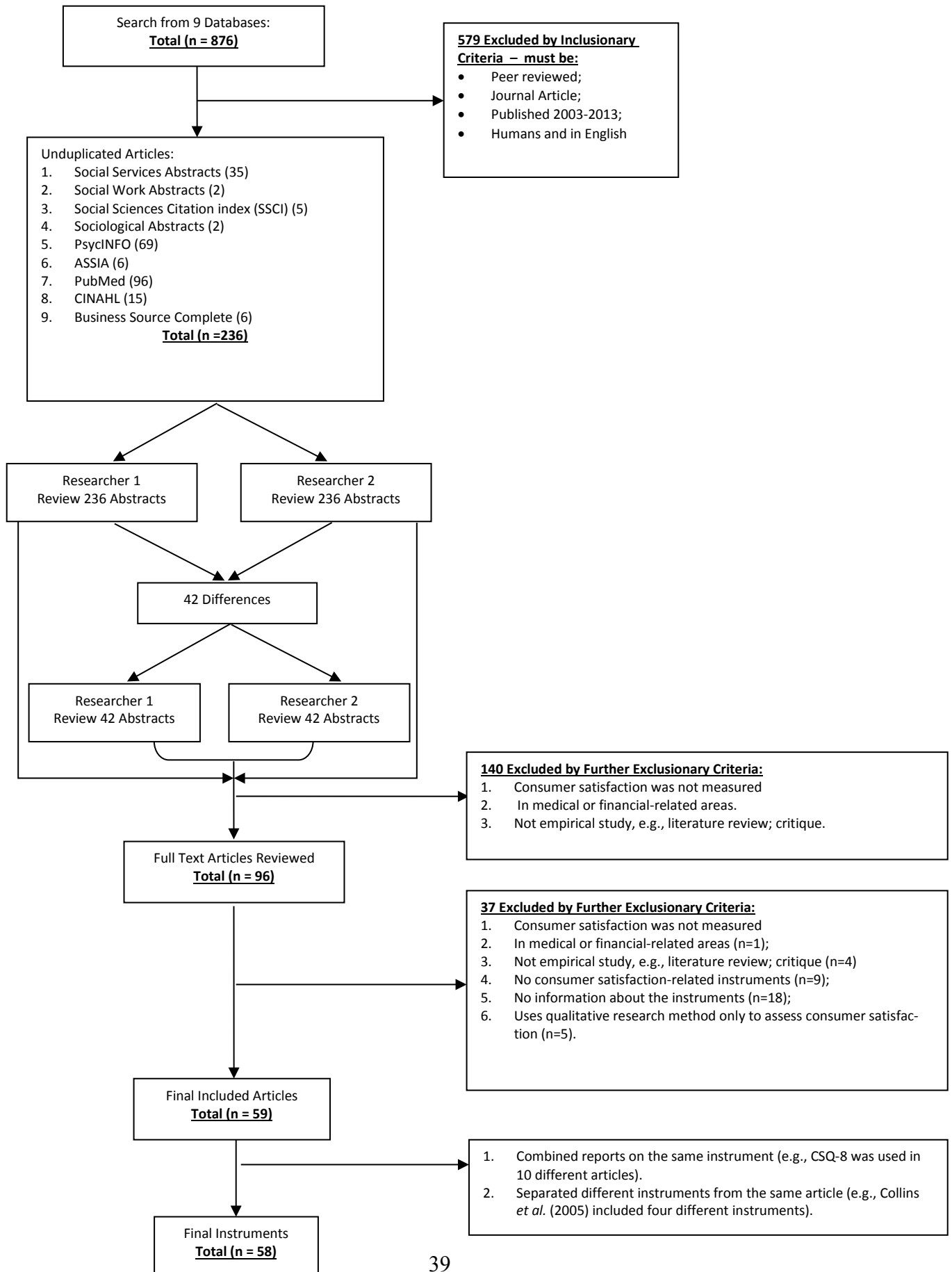
## Findings

Our review of the 59 articles that met the inclusion criteria yielded 58 satisfaction-related instruments. About 21% ( $n=12$ ) of the instruments used satisfaction-related items in an un-named scales. See Table 1 and the column "No Name." Of the 46 instruments that were named, 78% ( $n = 36$ ) included the word *satisfaction* in the instrument title. In Table 1, see the column entitled "Name Includes Satisfaction." About 17% of the studies provided a definition for or identified properties of satisfaction; however, the majority of the studies (83%) did not define satisfaction per se.

---

<sup>6</sup> For PubMed database: Search term changed a little because it cannot use "\*" for searching: medical OR medicine OR medically OR medication OR medic OR mediate OR business OR financial OR finances OR physical OR physic OR physics OR commercial OR commerce OR health OR customer OR patient OR patients

**Figure 6. Flow Chart for Systematic Review of Consumer Satisfaction Scales, 2003 – 2013**



*Table 1. Characteristics of Consumer Satisfaction Instruments, 2003-2013*

Description of Instrument (See Appendix A)	No Name	Name Includes Satisfaction	Items Includes Satisfaction	Non-English Version	Number of Reports
1. Treatment Evaluation Inventory-Short Form					
2. Children's Advocacy Center Non-offending Caregiver Satisfaction Survey		✓	✓		
3. Consumer Satisfaction Questionnaire-12 (CSQ-12)		✓	✓		
4. The Victim Satisfaction with Offender Dialogue Scale (VSODS)		✓	✓		
5. A Three-item Satisfaction Scale	✓		✓		
6. Client Satisfaction Survey		✓			
7. The Eight-item Satisfaction Scale	✓				
8. Resident Satisfaction Survey		✓	✓		
9. Parental Satisfaction Survey	✓		✓		
10. Client Satisfaction Inventory(CSI) The short-form version		✓			
11. Community satisfaction scale		✓	N/A		
12. Satisfaction with management scale		✓	N/A		
13. Satisfaction with employment scale, The short-form version		✓	N/A		
14. Satisfaction survey	✓		✓		
15. Behavior Intervention Rating Scale(BIRS)- Treatment Acceptability					2
16. 12-item Satisfaction Survey	✓				
17. Consumer Satisfaction Questionnaire-8 (CSQ-8) (English and Dutch)		✓	✓	✓	10
18. Working Alliance Inventory-Short version; (WAI-S)					2
19. Purdue live observation satisfaction scale (PLOSS)		✓	N/A		
20. Parenting Our Children to Excellence (PACE) Social Validity Survey			N/A		
21. Assisted Living Resident Satisfaction Scale (ALRSS)		✓	N/A		
22. Assisted Living Family Member Satisfaction Scale (ALFMSS)		✓	N/A		
23. The Family Satisfaction Instrument (final version-Section A-pretest vision)		✓	✓		
24. 3 satisfaction scales - Swedish to English	✓		✓	✓	
25. Counseling Evaluation Inventory-(The client satisfaction subscale version)		✓	✓		
26. Therapist satisfaction survey(1-item scale)		✓	✓		
27. Program Satisfaction Questionnaire (English to Chinese)		✓	✓	✓	
28. Making Better Career Decisions (MBCD)			✓		
29. Program Satisfaction Questionnaire	✓		✓		
30. Client Satisfaction: Case Management (CSAT-CM)		✓	✓		
31. Service Element Satisfaction Questionnaires		✓	✓		
32. Investigation Satisfaction Scale (ISS)-Caregivers		✓	✓		
33. Children's Satisfaction Survey-children		✓	✓		
34. The Parent Satisfaction with Foster Care Services Scale(PSFCSS)- Satisfaction items (Spanish and English)		✓	✓	✓	2
35. Treatment Satisfaction Survey(TSS)		✓	✓		
36. The School Opinion Survey- Parent form and student form			✓		
37. Youth Client Satisfaction Questionnaire (YCSQ) -revised version.		✓			
38. The Counselor Rating Form-Short (CRF-S)					
39. The Parent Satisfaction Questionnaire		✓			
40. Sex offender client treatment satisfaction survey	✓		✓		2
41. The Satisfaction With End-of-Life Care in Dementia Scale		✓			
42. The Multimodality Quality Assurance Instrument(MQA)			✓		
43. Parent satisfaction survey- with head start version	✓				
44. Client Satisfaction questionnaire		✓			
45. Client Satisfaction Survey(CSS)		✓			
46. Consumer Reports Effectiveness Score-4 items (CRES-4)-satisfaction		✓	✓		
47. Participants' Satisfaction With the Intervention	✓		✓		
48. Client satisfaction survey- English and Spanish	✓		✓	✓	
49. Chinese Subjective Outcome Scale (CSOS) -20 items					
50. Resident Satisfaction Index(RSI)- Short version		✓			
51. Overall Job Satisfaction Scale		✓	✓		
52. Clients' Overall Satisfaction Survey		✓	✓		
53. General satisfaction survey(Hebrew and plain English)		✓		✓	
54. Satisfaction with specific aspects of life at MA		✓	✓		
55. Post-Program Satisfaction Questionnaire		✓			
56. Client satisfaction measures	✓				
57. Student satisfaction survey		✓	✓		
58. The Consultation Evaluation Form (CEF)					
<b>Sum</b>	<b>12</b>	<b>36</b>	<b>31</b>	<b>6</b>	

Note. N/A = scale items are not available or reported in sufficient detail to make a determination



Only five instruments were used in more than one study, of which CSQ-8 was the most frequently used, appearing in 10 studies. Four instruments were each used in two studies: (a) the Behavior Intervention Rating Scale - Treatment Acceptability (BIRS; Cowan & Sheridan, 2003; Wilkinson, 2005); (b) Working Alliance Inventory-Short Version (WAI-S; Dearing et al., 2005; Fuertes et al., 2006); (c) Parent Satisfaction with Foster Care Services Scale-Satisfaction Items (PSFCSS; Kapp & Vela, 2003, 2004); and (d) Sex Offender Client Treatment Satisfaction Survey (Levenson et al., 2009; Levenson, Prescott & D'Amora, 2010). Given the number of studies that used the CSQ-8, we have summarized the research with this instrument in Table 2. A summary of the 58 instruments is presented in Appendix A. For instruments used in multiple studies, we have chosen one representative study.

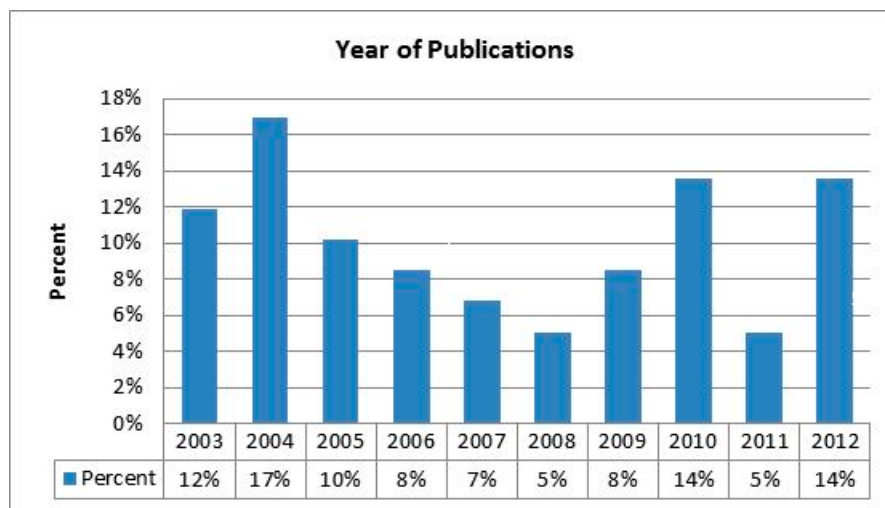
*Table 2. Studies Using the CSQ-8*

<b>Study</b>	<b>Behavioral Issue or Social Service Addressed</b>	<b>Reliability (<math>\alpha</math>)</b>
1. Dearing et al. (2005)	Alcohol treatment satisfaction	0.94
2. Denton, Nakonezny, & Burwell (2011)	Marriage and family therapy	0.86
3. Donker et al. (2009)	Depression and anxiety	0.91
4. Elledge et al. (2010)	Bullying	>.90
5. Hsieh & Guy (2009)	Caseworker performance	0.89
6. Murphy, Faulkner, & Behrens (2004)	Marriage and family therapy	0.86
7. Sorensen, Done, & Rhodes (2007)	Bipolar disorder	Not reported
8. Trute & Hiebert-Murphy (2007)	Disability services for children	0.96
9. Walsh & Lord (2004)	Hospital social work services	0.92
10. Yu (2005)	Anxiety	Not reported

Shown also in Table 1, a few instruments had been translated for data collections with non-English speaking populations. The CSQ-8 was translated for Dutch studies (Donker et al., 2009; Denton, Nakonezny & Burwell, 2011); (2) the General Satisfaction Survey was translated into Hebrew for an Israeli study (Spiro, Dekel & Peled, 2009); (3) the Client Satisfaction Survey (Schraufnagel & Li, 2010) and (4) the Parent Satisfaction with Foster Care Services Scale (PSFCSS) (Kapp & Vela, 2004) were translated into Spanish; (5) the Program Satisfaction Questionnaire was translated into Chinese (Gao, Luo & Chan, 2012); and (6) a scale with 3 satisfaction items was developed in Swedish (Friman, 2004). Additional details are provided in Appendix A.

Our literature search yielded 59 articles meeting inclusion criteria. One criterion for inclusion was publication in a peer-reviewed journal between 2003 and 2013. Shown in Figure 7, the largest percentage of the articles was published in 2004 (17%,  $n = 10$ ). The smallest percentages of reviewed articles were published in 2008 and 2011, with just 5% ( $n = 3$ ) of the reviewed articles published in each of those years.

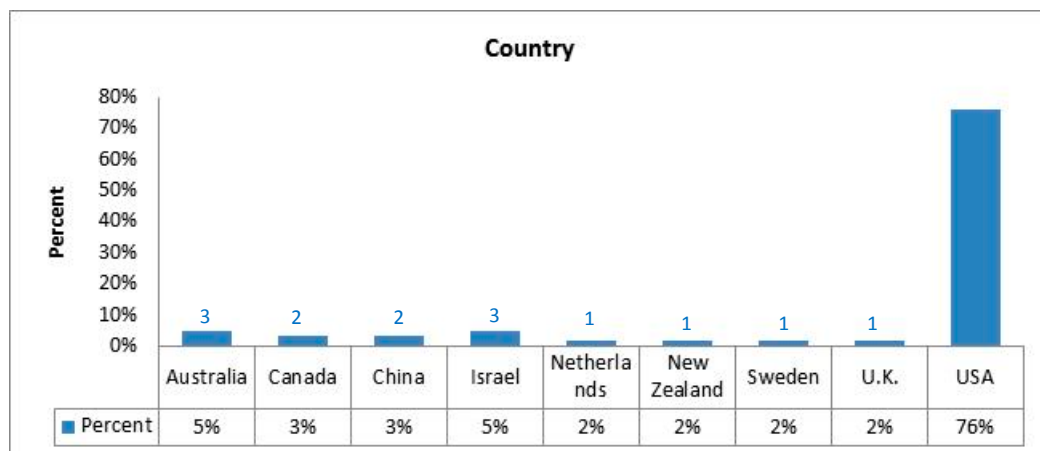
**Figure 7.** Distribution by Year of Publication



### Reports by Country

The 58 instruments were used to evaluate satisfaction with services in nine countries. The majority (76%) of studies were from the United States. Three studies were conducted in Australia (5%), three in Israel (5%), two in Canada (3%), and two in China (3%). Instruments were also identified from single studies based in each of the following countries: Netherlands, New Zealand, Sweden, and the United Kingdom (2%, respectively).

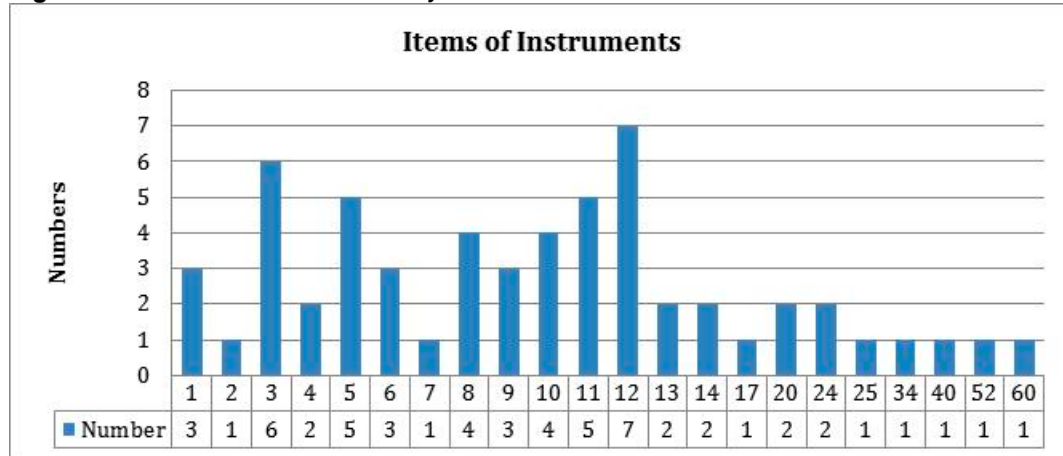
**Figure 8.** Distribution of Instruments by Country



### Instrument Size: Number of Items

The instruments varied considerably in the number of items, ranging from one item to 60 items. Of these 58 instruments, 7 (12%) used 12-item scales, 6 (10%) used 3-item scales. The modal number of items was 11. The length of longer scales may affect response rates.

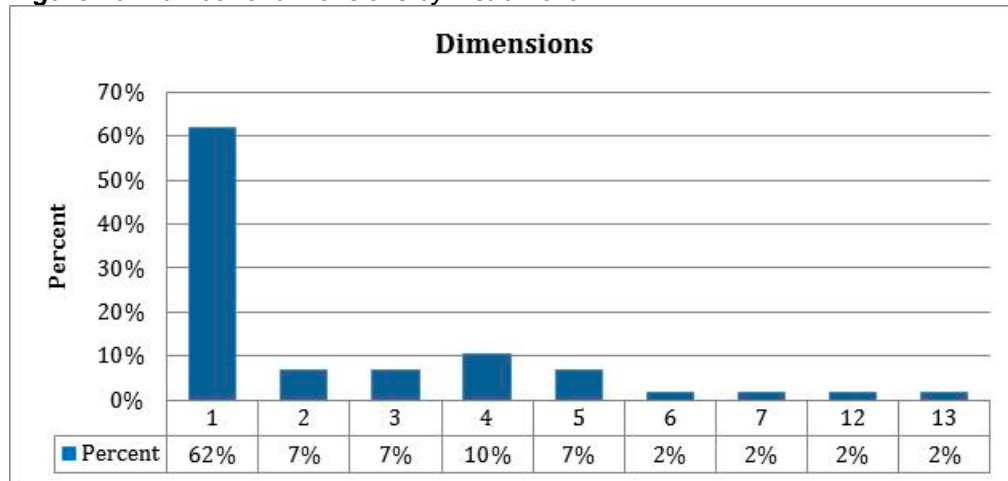
**Figure 9.** Distribution of the items by instrument



### Dimensions (Subscales): What Is the Dimensionality of Consumer Satisfaction?

A majority (62%, n=36) of the instruments had only had one dimension. However, shown in Figure 10, 22 scales had two or more subscales.

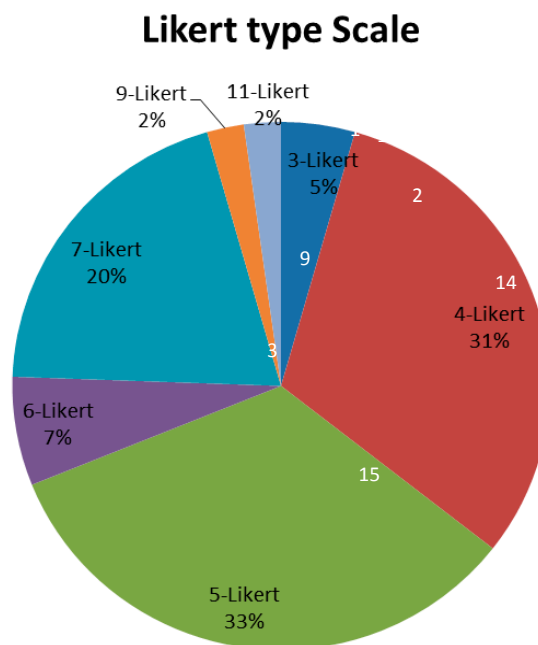
**Figure 10.** Number of dimensions by instrument



## Response Scales

The instruments used a variety of response scales. However, the majority (78%,  $n = 45$ ) employed a Likert-type response (i.e., *strongly disagree*, *disagree*, *neither disagree/agree*, *agree*, *strongly agree*). Five instruments (9%) scored items on ordinal scales. Four other instruments used either multiple-choice items (Castle et al., 2004) or 3-or 5-point scales (Levenson, 2009). The type of response scale was not reported for four instruments (Boyle et al., 2010; Edelman et al., 2006; Heinze, Hernandez, Jozefowicz, & Toro, 2010; Shek, 2010). Shown in Figure 11, of the 45 instruments that used Likert-type scales, 33% ( $n = 15$ ) used a 5-point Likert-type scale that recorded responses ranging from *strongly disagree* (1) to *strongly agree* (5) (e.g., Beavers, Kratochwill, & Braden, 2004; King & Bond, 2003). Another 31% ( $n = 14$ ) used a 4-point forced response Likert-type scale that ranged from *very dissatisfied* (1) to *very satisfied* (4); (e.g., Bonach, Mabry, & Potts-Henry, 2010; Liu, Guarino, & Lopez, 2012).

**Figure 11.** Likert-type response scales by instrument

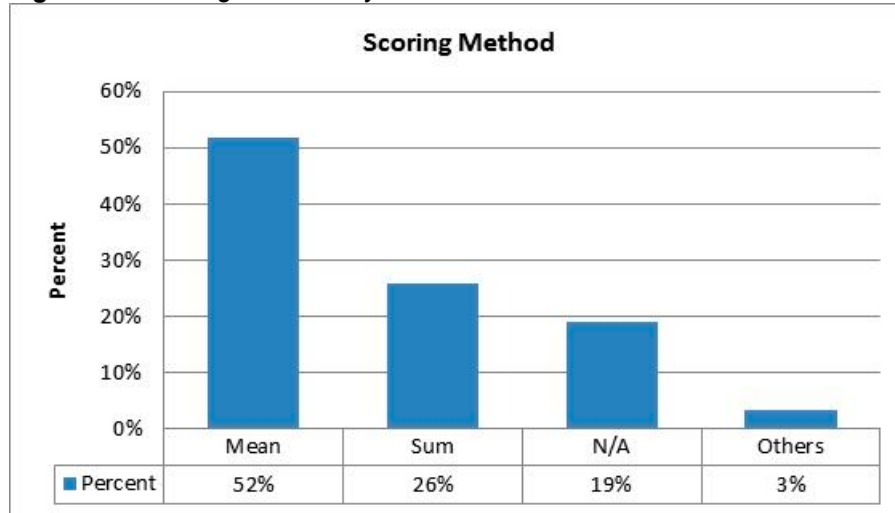


## Scoring Methods for Satisfaction

Satisfaction is usually reported as a mean value or a summed value of item scores. Shown in Figure 12, a majority of the 58 instruments (52%,  $n = 30$ ) averaged scores across scale items. For these instruments, the range of score val-

ues is formed by the low- and high-anchors of the scale. For example, if an instrument uses a 5-point Likert scale and uses the mean value scoring method, then the range of possible scores would be 1 to 5 (e.g., King & Bond, 2003). The second method, summing scores of all items, was used in about one quarter of the instruments (26%,  $n = 15$ ). With these instruments, the scoring ranges varied because the range depended on the number of items and the scoring scales of each instrument. The scoring method was not reported for 11 instruments (19%).

**Figure 12. Scoring methods by instrument**

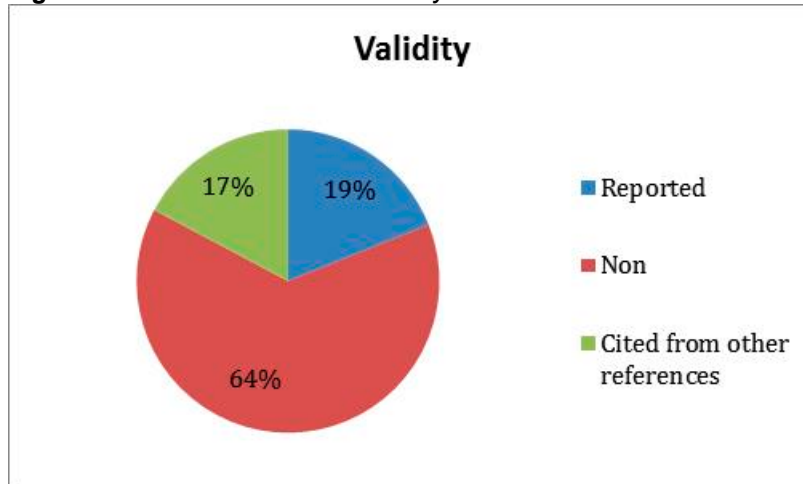


### Instrument Validity and Reliability

In large part, the process of refining and validating an instrument focuses on reducing the potential for measurement error. Estimates of an instrument's reliability assess the stability of the measures over time and the internal consistency of items (e.g., the average inter-item correlation). Although it has several dimensions, validity focuses on the degree to which an instrument is plausibly related to the construct of interest.

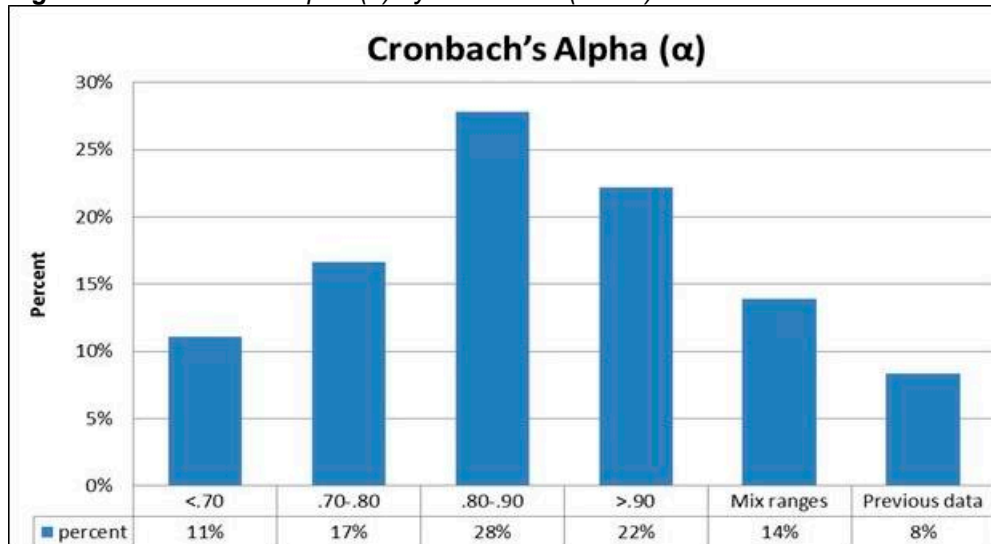
**Validity of instruments.** Of the 58 instruments, only 19% ( $n = 11$ ) conducted analyses to assess validity, with 10 reports (17%) citing previous validation studies (see Figure 13). For example, Denton et al. (2011) reported that the validity of the CSQ-8 was acceptable based on a previous study conducted by Attkisson and Zwick (1982). *Nearly two-thirds of the reports (64%,  $n = 37$ ) conducted no validation analyses.*

**Figure 13. Validation Assessment by Instrument**



**Reliability of instruments.** Most of the reports on the 58 instruments (69%,  $n = 40$ ) included a measure of reliability. Among the 40 instruments for which reliability was reported, 82% ( $n = 33$ ) reported the results of Cronbach's alpha ( $\alpha$ ), two instruments (5%) reported test-retest ( $r$ ) results, and two instruments (5%) reported reliability as both Cronbach's alpha and test-retest. See Figure 14. In addition, reliability for three instruments (8%) was reported by citing the estimates from previous studies.

**Figure 14. Cronbach's Alpha ( $\alpha$ ) by Instrument ( $n = 35$ )**



## Discussion

A variety of instruments has been developed to assess satisfaction with services in social welfare. These instruments were designed for a broad range of clients, including adults and children. Our review identified 58 satisfaction-related instruments developed or refined in the last 10 years. Although the CSQ-8 was used in 10 studies (17%), the variety of instruments suggests that no single scale currently encompasses all the aspects of satisfaction in which practitioners, policymakers, and researchers are interested. Further, the considerable efforts invested in developing study-specific measures of satisfaction suggest that global measures may have comparatively less utility for those who hope to understand the nuances of satisfaction and its influence on treatment adherence and outcomes.

That it has been a consistent focus of research during the past decade indicates sustained interest in satisfaction with services as a tool for evaluation and quality assurance. On average, scales were brief, with a 10-item format being most frequently used. The length of a scale is an important element because the burden of completing a survey may affect the quality of the data. That is, longer surveys with a greater number of items can reduce response rates (Royse et al., 2010). On balance, reliability was acceptable. The average reliability across all the reported alphas was .85, which falls in the acceptable range (Nunnally, 1978).

“Satisfaction” was operationalized in a variety of ways. Some scales do not use the word *satisfaction*, and others have subscales designed to assess specific feature of satisfaction. For example, some instruments used “Treatment Acceptability” as an indicator of satisfaction (Cowan & Sheridan, 2003; Wilkinson, 2005). Although most instruments adopted *satisfaction* in their titles, 83% failed to provide a definition or to discuss its properties.

The research on satisfaction with services has serious limitations. First, some two thirds (67%) of the reports ( $n = 39$ ) did not include any discussion of the limitations of satisfaction measures per se. To be sure, many reports discussed limitations imposed by designs (e.g., sampling methods, sample sizes). But they failed to consider limitations associated with satisfaction measures. A few articles noted instrument length. For example, Fuertes et al. (2006) and Gati et al. (2006) reported that using a single item to assess satisfaction might have produced measurement error. Second, many satisfaction scales appear insufficient in representing different aspects of satisfaction, such as net promotion (Coloma, Gibson, & Packard, 2012). Third, most reports (64%) contained no validity analyses. Fourth, items often appear subject to variable interpretations (Charbonneau & Van Ryzin, 2012). Fifth, the methods used to score instruments rarely account for different elements of service and the relative importance of service elements (Hsieh, 2012). On balance, these limitations suggest that the measurement of satisfaction in social welfare is less sophisticated and nuanced than the measurement of other constructs, such as social problems and mental health symptoms.

## Part 5. Summary and Recommendations

---

The use of satisfaction as an outcome in the evaluation of social welfare services is affected by a variety of knotty issues in design and measurement. The term *design* refers to the overall evaluation strategy. It includes the use of control or comparison groups, the mechanism for assigning participants to conditions, the selection of instruments, the means for data collection, the nature of hypotheses, the plans for assessing the properties of measures, and the methods of analysis. The “goodness” of designs is often gauged by threats to the validity of inferences associated with the features of different evaluation approaches. For example, because the use of a control group tends to eliminate alternative explanations, designs with control groups are considered stronger than designs without control groups.

In using satisfaction with services as a measure, two design-related issues arise: attrition and confoundedness. These issues are not unique to satisfaction. Indeed, they affect all measurement and inference. *Attrition* is a form of sampling bias that is introduced when missing information cannot be considered missing at random. That is, when there is a significant difference in the satisfaction scores and characteristics of program stayers and leavers (dropouts). *Confoundedness* refers to factors that may complicate making a causal inference about the relationship between satisfaction and program outcomes. These confounding factors include the influence of the organizational image, which may affect satisfaction ratings separately from the quality of services provided. Confounding factors often include unmeasured variables that may explain an apparent relationship. These unmeasured variables are sometimes called *unobserved heterogeneity*.

*Measurement* refers to the dimensionality, reliability, validity, and other properties of the scales that are intended to describe satisfaction with services. Over the years, satisfaction has been used in many evaluations of social welfare services, and it has been conceptualized in many ways. This variability is both a strength and a limitation in that instruments measure a wide variety of concepts. Finally, studies suggest that satisfaction has reactive properties. *Reactivity* refers to biases that are introduced when survey participants respond to items, the sequencing of items, or the formatting of items in a questionnaire. Reactivity can also be engendered from the conditions or the settings in which questionnaires are administered.



## Design: Attrition and Confoundedness

From a design perspective, measures of satisfaction are intended to provide information about the quality of services. To draw a conclusion about a service, satisfaction data must be available at the assignment level. The term *assignment* refers to the assignment of a client to a service condition. Policymakers are interested in learning the extent to which clients who are eligible for and who begin a service find the service satisfactory. In research studies, assignment is sometimes random, whereas in practice, assignment is usually administrative. That is, it is based on the expressed need of a client, the need as perceived by an intake worker, or, when services are court related, the need as perceived by an official with legal authority (e.g., a probation or parole officer). As an indicator of service quality, satisfaction data must be collected from all service participants who start a service and not merely from those participants who complete a service. To make an inference about a service and avoid bias, information is needed from all service participants, regardless of their length of participation.

**Is attrition biased?** Consumers who complete an episode of service are likely to be more satisfied than those who drop out (Gottlieb & Wachala, 2007). To have validity in making an inference about a service, satisfaction information is needed from all clients who begin services or, alternatively, attrition must be shown to be missing at random (MAR).<sup>7</sup> If missing information is MAR, attrition will not bias satisfaction estimates. Unfortunately, it is unlikely that attrition will be MAR. When satisfaction ratings have not been collected from dropouts, satisfaction scores will have validity only to the degree that attrition is unbiased.

**Confoundedness and unobserved heterogeneity.** In addition to attrition, a variety of factors may influence satisfaction ratings. It is clear that ratings are influenced by a nuts-and-bolts appraisal of the value of a service relative to expectations for service. However, at least three other potentially confounding factors – the image of the service provider (e.g., the reputation of the agency), the affective or utilitarian aspects of service (e.g., the courtesy of staff, the availability of parking, the attractiveness of the facility), and the consumer's sense of equity relative to the services received by others (e.g., whether services conform to the perception of services received by others) – may affect satisfaction ratings. These are considered confounding variables because they complicate using satisfaction to indicate the quality of services. Confounding variables may influence satisfaction independent of the effect of a service on outcomes.

Environmental and organizational factors also can confound comparison of satisfaction scores. If the eligibility criteria for a set of similar services vary

---

<sup>7</sup> Missing at Random (MAR) is a technical term used to describe patterns of missing data in statistics. Other such terms include Missing Completely at Random (MCAR) and Not Missing at Random (NMAR).

across municipalities, then the characteristics of people who receive those services are likely to vary. Suppose, for example, that all municipalities offer parenting training services. However, some municipalities restrict parent training to those parents of children who have been disruptive in school. Assuming a constant effect for services, satisfaction ratings, as well as behavioral outcomes, are likely to vary across municipalities because of differences in participant characteristics. This is a *selection bias*. To compare services across municipalities, eligibility criteria must be similar. Otherwise, they constitute a confound.

### **Measurement: Dimensionality, Reactivity, Subjective Causal Appraisal**

Satisfaction has many different formulations. Researchers have used factor analysis and confirmatory factor analysis to identify a variety of dimensions within satisfaction scales. Typically, these dimensions are defined as subscales of satisfaction, and they have acceptable reliability. These dimensions often include ratings of satisfaction per se, a performance or service quality appraisal, and willingness to make a word-of-mouth recommendation. Unlike in business, in which measuring customer satisfaction is separated from measuring performance, measurement in social welfare has tended to collapse subjective outcome appraisals (e.g., *To what degree did the services we provide help you solve your problems or meet your needs?*) into satisfaction scales.

**Dimensionality: What are the core elements of consumer satisfaction?** No common dimensionality emerged across the 58 consumer satisfaction measures. Some scales have one dimension, whereas others have two, three, or more dimensions. Some scales explicitly use the term *satisfaction* (e.g., *How satisfied were you with <x>?*), while others do not. Some scales include net promotion items (e.g., *Net of everything in your experience at <x>, would you recommend this service?*) or word-of-mouth recommendation (e.g., *If a friend were in need of similar help, would you recommend our program to him or her?*). Others do not. After more than 35 years, the dimensionality of satisfaction with services in social welfare remains unclear.

To be sure, many scales focus solely on satisfaction. For example, the Client Satisfaction – Case Management Scale begins with the stem question (Hsieh, 2006): *How satisfied are you with...*

- *your case manager's assessment of your needs?*
- *the plan of care your case manager developed?*
- *your case manager's knowledge regarding the services that are available?*
- *your case manager's ability to get services for you?*
- *the availability of your case manager?*

However, the content of many scales is no longer focused on satisfaction with a service experience. As indicators, satisfaction scales often include performance appraisals of program staff and willingness to make word-of-mouth referrals. Moreover, scales can include an invitation to make retrospective inferences about the effective-

ness of services in meeting needs or solving problems. In scales like the CSQ-8 (Nguyen, Attkisson, & Stegner, 1983), content focuses on satisfaction, quality, net promotion, and impact. Among other items, the CSQ-8 includes the following questions:

- How *satisfied* are you with the help you received?
- How would you rate the *quality* of the services you received?
- If a friend were in need of similar help, would you *recommend* our program to him or her?
- To what extent has our program *met your needs*?
- Have the services you received *helped you deal more effectively with your problems*?

These few items measure *satisfaction*, performance or service *quality*, willingness to *recommend* or refer (i.e., a measure of brand loyalty in business), and they invite a subjective appraisal of the impact of service on *needs* and *problems*. This brevity and breadth of measurement probably account for the widespread adoption of the CSQ-8.

The use of subjective appraisals of the impact of services on problems and needs has spurred some to wonder if *perceived change* and not satisfaction is being measured (Lunnen & Ogles, 1998). Does inviting a subjective appraisal of the effectiveness of a service have construct validity for *satisfaction*? Has scale content become so inclusive as to blur potentially important constructs, which might, as in business, be used independently in assessing outcomes? Arguably, perceived change is an informative measure whose value is masked when it is embedded in a satisfaction scale.

**Reactivity: “Thank You” effects.** Further complicating the use of consumer satisfaction as an evaluation measure is the fact that satisfaction ratings are known to be high for participation in nearly all social services, including attention-only services. Participants involved in reading groups, discussions, and social support groups, or who receive a placebo intervention often report high satisfaction scores (Ingram, & Chung, 1997). These high satisfaction ratings produce skewed distributions that are difficult to analyze. If all scores fall in an upper register, there may be little variability to explain differences in outcomes.

Social desirability biases typically arise from the appreciation of clients for any amount of attention. However, they may arise also from concern that an honest but negative evaluation might result in denial of future services and fear that responses might not be kept anonymous. Concerns about the confidentiality of responses may occur because of the setting in which satisfaction ratings are invited. For example, a client might not feel comfortable providing ratings in the waiting room of a clinic or office.

High scores on satisfaction instruments are sometime attributed to the *Thank-You* effect (Gottlieb & Wachala, 2007, p. 382). The thank-you effect derives from genuine appreciation that stems from participation in any service –

including study groups, seminars, and the like. These thank-you effects are non-trivial, and in assessing the impact of a service with a specific strategy (e.g., a service designed to build skills), they must be controlled. Thus, *where satisfaction is to be used in relation to a program with a specific theory of change, satisfaction ratings from an attention-only or support services control condition are desirable.*

**Masking to reduce reactivity.** To reduce reactivity, some instruments mask intent by avoiding the use of satisfaction-related terms. Masking words or phrases to which survey respondents might react is common in scale development. This masking is an attempt to reduce stereotypic and socially desirable but false responses. In juvenile and criminal justice, the words for more serious offenses – murder, rape, robbery, and burglary – are thought to be reactive, and therefore, masking is used. For example, rather than using the word *burglary* in self-report surveys, a scale might be constructed to ask, “Have you taken anything worth over X SEK that was not yours?” To have face and construct validity, masking has to be impeccably related to the construct (e.g., burglary).

For scales that do not use questions like “How satisfied are you...,” the conceptual tie of item content to satisfaction must be clear. Consider the 14-item Youth Client Satisfaction Scale (YCSQ: Shapiro, Welker, & Jacobson, 1997). Based on factor analyses, the YCSQ has two subscales: *relationship with therapist* and *benefits of therapy*. The first subscale focuses on whether youths *feel understood*, whether they *like their worker*, and whether the worker’s *suggestions seemed helpful*. The second subscale invites inferences about the effects of service. This subscale includes items asking whether youths *feel differently as a result of treatment* and whether service *helped resolve problems*. The specificity of the two subscales increases face validity. Indeed, in the absence of a Net Promotion item (e.g., *Would you recommend our program to others?*), the value of the YCSQ rests more – arguably – with its subjective appraisal of *worker alliance* and the *service impact* than with its measurement of satisfaction overall. Masking, in this case, appears to improve the program relevance of constructs, although neither construct mentions satisfaction.

**Subjective causal attributions about the effect of services.** The idea of inviting retrospective subjective evaluations of the impact of service experiences has been incorporated in many satisfaction scales in social welfare. The use of these subjective causal appraisals is somewhat different from business, where analytics are used to assess relationships between independent measures of expectations, perceived performance or quality, satisfaction, and outcomes. To be sure, subjective effect appraisals are common in business, but they are considered separately as an aspect of Perceived Performance. See Figure 1.

In social welfare, it is different. Performance appraisal has been included in satisfaction scales. For example, as a part of a measurement package using just three satisfaction items (i.e., *satisfaction with changes in fear level*, *satisfaction with avoidance*, and *satisfaction with interference of phobias following treat-*

ment) in a study of treatment for phobias, Ollendick and colleagues (2009) randomized 196 Swedish and American youths to a one-session brief intervention, an educational support group, or a waitlist control. The one-session intervention emerged superior in clinician ratings of phobic severity, clinician ratings of symptoms, and youth and parent ratings of satisfaction. In studies such as this one, where measures are tied closely to specific, program-relevant outcomes (e.g., changes in fear level), the subjective causal appraisals that are embedded in satisfaction ratings begin to approximate self-reports of behavioral change. Compared to global satisfaction scores, these kinds of satisfaction measures may have greater predictive validity for long-term outcomes.

**When service participation is investigative or required.** Finally, and largely as a caveat, use of the term *consumer* is conditional when services are not voluntary (Martin, Petr, & Kapp, 2003). In the child welfare, corrections, drug/alcohol, and mental health fields, services may be mandatory or involve restricted choices (e.g., removal of child from the home in lieu of parental participation in family preservation services). Unlike in the private sector, consumers in social welfare often have less agency and may be required to participate in programs that range from investigations of child maltreatment or intimate partner violence to court-ordered rehabilitation or treatment. Service providers usually do not compete for clients by accruing brand loyalty through advertising and customer satisfaction. That is, consumers in social welfare may have little choice, few resources, and nowhere else to shop. The Thank-You effects observed when service participation is voluntary are subdued, at best, when services are mandatory.

### **Need for Theoretical Models: Is Consumer Satisfaction a Mediator?**

Social welfare services are often designed to address particular risk factors (Fraser, Richman, Galinsky, & Day, 2009). Interventions for obesity, for example, are usually designed to address risk factors for obesity. These include inactivity and unhealthy diet. A theory of change specifies risk factors and the means for changing them.

Change theories may include satisfaction as a mediating variable that predicts longer-term outcomes. Testing mediation in a service designed to treat obesity might involve measuring changes in diet or exercise and monitoring outcomes such as changes in weight or body mass. It might also involve measuring client satisfaction. Studies that test for mediation have the potential to show that the original conceptualization of satisfaction was correct. That is, program participation may operate through satisfaction to produce behavioral changes.

Unfortunately, the evidence for mediation is scant. In a study of alcohol treatment, Dearing et al. (2005) found that satisfaction was related to abstinence. And in a study of the effectiveness of Assertive Community Treatment with homeless people with serious mental illnesses, Fletcher et al. (2008) found significant program effects for satisfaction and housing stability, but they observed

no effects for psychiatric symptoms and substance use, which should be principal outcomes for a mental health intervention.

Complicating the mediation argument, studies rarely control for other potential mediators. To show that satisfaction is an important mediator, the effects of other mediators have to be understood. These include, for example, the use of medications. Before it is possible to conclude that satisfaction is a necessary mediator (and not merely a covariate), other plausible explanations of change must be measured and included in mediation analyses. At best, the field is at the beginning of this enterprise.

## Recommendations

Many scales and measures of satisfaction are available. They range from simple single-item measures to multi-dimensional scales of the acceptability and the perceived effects of service participation. Reliability is generally in the acceptable range. However, the degree to which satisfaction predicts behavioral and other theoretically important outcomes remains uncertain. In this regard, the research is mixed.

From our review, five recommendations emerge:

- Satisfaction with services should complement the use of theoretically relevant outcomes in program evaluation and quality assurance.
- Satisfaction measures should include at least three kinds of questions:
  - (a) satisfaction with service elements (e.g., *To what degree were you satisfied with [service element x, y, or z]?);*
  - (b) promotion given an entire service experience (e.g., *Would you recommend this program a friend?);* and
  - (c) perceived change (e.g., *To what degree did participation in the program resolve your problems or meet your needs?).*

Because of the potential for these three elements of satisfaction to be differentially related to other outcomes, they should be considered separately as well as combined in data analysis.

- Satisfaction items that are related directly to the key elements of services and that invite a subjective causal appraisal of perceived changes may be preferable to global satisfaction ratings.
- When satisfaction with services is used, strategies must be developed to reduce attrition and to secure satisfaction ratings from program dropouts. These include conducting exit interviews with dropouts and collecting satisfaction ratings incrementally throughout service periods. If attrition biases are not controlled, satisfaction scores will have little validity as an indicator of the quality of a service. They will represent only the satisfaction of those clients who were so pleased with a service as to stay to the end. Under these circumstances, satisfaction scores can have validity only to the extent that information lost to attrition is missing at random.

- If consumer satisfaction data are to be used to compare services across providers, information on potential confounding variables must be collected. In particular, the image or reputation of service providers must be controlled in analyses. Independent of actual service outcomes, satisfaction ratings may be influenced by the prestige and prominence of the organization providing services.

Collecting information on the extent to which clients are satisfied with their experiences with social welfare services is consistent with both the New Public Management (NPM) and the Public Value (PV) perspectives in public administration and finance. NPM emphasizes accountability, privatized service provision, and consumer responsiveness (Ferlie et al., 1996; Hood, 1991), whereas the emerging PV perspective places emphasis less on business-related principles and more on citizen participation and societal benefit or value (Moore, 1995; Stoker, 2006). Considering satisfaction with services is consistent also with professional ethics and evidence-based practice, where consumer involvement is valued highly. Moreover, measuring consumer or client satisfaction is widely supported across professional, philanthropic, and advocacy organizations.

Satisfaction with services is an important outcome and an integral aspect of measuring the quality of social welfare programs (Royse et al., 2010; WHO, 2000). That is, measuring satisfaction gives agency and voice to service participants. In addition, the research suggests that satisfaction is a function of service engagement (e.g., Heinze et al., 2010). More engaged program participants are likely to report higher satisfaction and to observe greater benefit from their receipt of services. In this sense, satisfaction is an outcome that should – and sometimes does – predict behavioral and other outcomes. Assessing satisfaction is an essential aspect of evaluation and quality assurance. However, satisfaction should not be used alone. It should be used as an element of a measurement package that includes a range of theoretically relevant outcomes.

# References

---

- Allen, B., Brymer, M. J., Steinberg, A. M., Vernberg, E. M., Jacobs, A., Speier, A. H., & Pynoos, R. S. (2010). Perceptions of psychological first aid among providers responding to Hurricanes Gustav and Ike. *Journal of Traumatic Stress, 23*(4), 509-513. doi:10.1002/jts.20539
- Amirkhanyan, A. A. (2011). What is the effect of performance measurement on perceived accountability effectiveness in state and local government contracts? *Public Performance & Management Review, 35*(2), 303-339. doi:10.2753/PMR1530-9576350204
- Anderson, E. W., Fornell, C., & Mazvancheryl, S. K. (2004). Customer satisfaction and shareholder value. *Journal of Marketing, 68*(4), 172-185. doi:10.1509/jmkg.68.4.172.42723
- Anderson, R. E. (1973). Consumer dissatisfaction: The effect of disconfirmed expectancy on perceived product performance. *Journal of Marketing Research, 10*(1), 38-44. doi:10.2307/3149407
- Andreassen, T. W., & Lindestad, B. (1998). Customer loyalty and complex services: The impact of corporate image on quality, customer satisfaction, and loyalty for customers with varying degrees of service expertise. *International Journal of Service Industry Management, 9*, 7-23. doi:10.1108/09564239810199923
- Attkisson, C. C., & Zwick, R. (1982). The Client Satisfaction Questionnaire: Psychometric properties and correlations with service utilization and psychotherapy outcome. *Evaluation and Program Planning, 5*(3), 233-237. doi:10.1016/0149-7189(82)90074-X
- Beavers, K. F., Kratochwill, T. R., & Braden, J. P. (2004). Treatment utility of functional versus empiric assessment within consultation for reading problems. *School Psychology Quarterly, 19*(1), 29-49. doi:10.1521/scpq.19.1.29.29405
- Bonach, K., Mabry, J. B., & Potts-Henry, C. (2010). Exploring nonoffending caregiver satisfaction with a children's advocacy center. *Journal of Child Sexual Abuse, 19*(6), 687-708. doi:10.1080/10538712.2010.522495
- Boyle, C. L., Sanders, M. R., Lutzker, J. R., Prinz, R. J., Shapiro, C., & Whitaker, D. J. (2010). An analysis of training, generalization, and maintenance effects of primary care Triple P for parents of preschool-aged children with disruptive behavior. *Child Psychiatry and Human Development, 41*(1), 114-131. doi:10.1007/s10578-009-0156-7
- Bradshaw, W., & Umbreit, M. (2003). Assessing satisfaction with victim services: The development and use of the Victim Satisfaction with Offender Dialogue Scale (VSODS). *International Review of Victimology, 10*(1), 71-83. doi:10.1177/026975800301000104
- Brenninkmeijer, V., & Blonk, R. W. (2012). The effectiveness of the JOBS program among the long-term unemployed: A randomized experiment in the Netherlands. *Health Promotion International, 27*(2), 220-229. doi:10.1093/heapro/dar033
- Brooks, F., & Brown, E. (2005). A program evaluation of Los Angeles ACORN's welfare case advocacy. *Journal of Human Behavior in the Social Environment, 12*(2-3), 185-203. doi:10.1300/J137v12n02\_10



- Butler, S. S., Gomon, J., & Turner, W. (2004). Satisfaction and well-being among residents of a low-income, rural assisted living facility and the implications for social work practice. *Journal of Social Work in Long-Term Care*, 3(1), 61-84. doi:10.1300/J181v03n01\_06
- Cardozo, R. N. (1965). An experimental study of customer effort, expectation, and satisfaction. *Journal of Marketing Research*, 2(3), 244-249. doi:10.2307/3150182
- Castle, N. G., Lowe, T. J., Lucas, J. A., Robinson, J. P., & Crystal, S. (2004). Use of resident satisfaction surveys in New Jersey nursing homes and assisted living facilities. *Journal of Applied Gerontology*, 23(2), 156-171. doi:10.1177/0733464804265584
- Charbonneau, É., & Van Ryzin, G. G. (2012). Performance measures and parental satisfaction with New York City schools. *American Review of Public Administration*, 42(1), 54-65. doi:10.1177/0275074010397333
- Collins, M., Curley, A., Clay, C., & Lara, R. (2005). Evaluation of social services in a HOPE VI housing development: Resident and staff perceptions of successes and barriers. *Evaluation and Program Planning*, 28(1), 47-59. doi:10.1016/j.evalprogplan.2004.10.004
- Coloma, J., Gibson, C., & Packard, T. (2012). Participant outcomes of a leadership development initiative in eight human service organizations. *Administration in Social Work*, 36(1), 4-22. doi:10.1080/03643107.2011.614540
- Copeland, V., Koeske, G., & Greeno, C. G. (2004). Child and mother client satisfaction questionnaire scores regarding mental health services: Race, age, and gender correlates. *Research on Social Work Practice*, 14(6), 434-442. doi:10.1177/1049731504265839
- Corrigan, J. D., & Schmidt, L. D. (1983). Development and validation of revisions in the counselor rating form. *Journal of Counseling Psychology*, 30(1), 64-75. doi:10.1037/0022-0167.30.1.64
- Cowan, R. J., & Sheridan, S. M. (2003). Investigating the acceptability of behavioral interventions in applied conjoint behavioral consultation: Moving from analog conditions to naturalistic settings. *School Psychology Quarterly*, 18(1), 1-21. doi:10.1521/scpq.18.1.1.20877
- Cross, T. P., Jones, L. M., Walsh, W. A., Simone, M., Kolko, D. J., Szczepanski, J., ... Magnuson, S. (2008). *Evaluating children's advocacy centers' response to child sexual abuse*. Retrieved from <http://www.ncjrs.gov/pdffiles1/ojdp/218530.pdf>
- Dauenhauer, J. A., Cattat Mayer, K., & Mason, A. (2007). Evaluation of adult protective services: Perspectives of community professionals. *Journal of Elder Abuse & Neglect*, 19(3-4), 41-57. 10.1300/J084v19n03\_03
- Day, R. L. (1977). Toward a process model of consumer satisfaction. In H. K. Hunt (Ed.), *Conceptualization and measurement of consumer satisfaction and dissatisfaction*. Cambridge, MA: Marketing Science Institute.
- Dearing, R. L., Barrick, C., Dermen, K. H., & Walitzer, K. S. (2005). Indicators of client engagement: Influences on alcohol treatment, satisfaction, and outcomes. *Psychology of Addictive Behaviors*, 19(1), 71-78. doi:10.1037/0893-164X.19.1.71
- Denton, W. H., Nakonezny, P. A., & Burwell, S. R. (2011). The effects of meeting a family therapy supervision team on client satisfaction in an initial session. *Journal of Family Therapy*, 33(1), 85-97. doi:10.1111/j.1467-6427.2010.00518.x
- Department of Education, Queensland. (1996). *Partners for excellence: The strategic plan 1997-2001 for the Department of Education*. Brisbane, AU: Author.

- Donker, T., van Straten, A., Riper, H., Marks, I., Andersson, G., & Cuijpers, P. (2009). Implementation of Internet-based preventive interventions for depression and anxiety: Role of support? The design of a randomized controlled trial. *Trials*, 10(59). doi:10.1186/1745-6215-10-59
- Dufour, S., & Chamberland, C. (2004). The effectiveness of selected interventions for previous maltreatment: Enhancing the well-being of children who live at home. *Child & Family Social Work*, 9(1), 39-56. doi:10.1111/j.1365-2206.2004.00302.x
- Dumas, J. E., Arriaga, X. B., Begle, A. M., & Longoria, Z. N. (2011). Child and parental outcomes of a group parenting intervention for Latino families: A pilot study of the CANNE program. *Cultural Diversity and Ethnic Minority Psychology*, 17(1), 107-115. doi:10.1037/a0021972
- Edelman, P., Guihan, M., Bryant, F. B., & Munroe, D. J. (2006). Measuring resident and family member determinants of satisfaction with assisted living. *Gerontologist*, 46(5), 599-608. doi:10.1093/geront/46.5.599
- Ejaz, F. K., Straker, J. K., Fox, K., & Swami, S. (2003). Developing a satisfaction survey for families of Ohio's nursing home residents. *Gerontologist*, 43(4), 447-458. doi:10.1093/geront/43.4.447
- Elledge, L. C., Cavell, T. A., Ogle, N. T., & Newgent, R. A. (2010). School-based mentoring as selective prevention for bullied children: A preliminary test. *Journal of Primary Prevention*, 31(3), 171-187. doi:10.1007/s10935-010-0215-7
- Engel, J. F., Kollat, D. T., & Blackwell, R. D. (1973). *Consumer behavior* (2nd ed.). New York, NY: Holt, Rinehart, & Winston.
- Erchul, W. P. (1987). A relational communication analysis of control in school consultation. *Professional School Psychology*, 2(2), 113-124. doi:10.1037/h0090534
- Eyberg, S. M. (1993). Consumer satisfaction measures for assessing parent training programs. In L. Van de Creek, S. Knapp, & T. L. Jackson (Eds.), *Innovations in clinical practice: A source book* (Vol. 12; pp. 377-382). Sarasota, FL: Professional Resource Press.
- Ferlie, E., Ashburner, L., Fitzgerald, L., & Pettigrew, A. (1996). *New public management in action*. Oxford, UK: Oxford University Press. doi:10.1093/acprof:oso/9780198289029.001.0001
- Fields, D. L. (2002). *Taking the measure of work: A guide to validated scales for organizational research and diagnosis*. Thousand Oaks, CA: Sage.
- Fletcher, T. D., Cunningham, J. L., Calsyn, R. J., Morse, G. A., & Klinkenberg, W. (2008). Evaluation of treatment programs for dual disorder individuals: Modeling longitudinal and mediation effects. *Administration and Policy in Mental Health and Mental Health Services Research*, 35(4), 319-336. doi:10.1007/s10488-008-0170-2
- Fontana, A., Ford, J. D., & Rosenheck, R. (2003). A multivariate model of patients' satisfaction with treatment for posttraumatic stress disorder. *Journal of Traumatic Stress*, 16(1), 93-106. doi:10.1023/A:1022071613873
- Fornell, C. (1992). A national customer satisfaction barometer: The Swedish experience. *Journal of Marketing*, 56, 6-21. doi:10.2307/1252129
- Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., & Bryant, B. E. (1996). The American Customer Satisfaction Index: Nature, purpose, and findings. *Journal of Marketing*, 60, 7-18. doi:10.2307/1251898

- Fox, J. G., & Storms, D. M. (1981). A different approach to sociodemographic predictors of satisfaction with health care. *Social Science and Medicine*, 15(5), 557-564. doi: 10.1016/0271-7123(81)90079-1
- Fraser, M. W., Richman, J. M., Galinsky, M. J., & Day, S. H. (2009). *Intervention research: Developing social programs*. New York, NY: Oxford University Press.
- Freedman, N., Hoffenberg, J. D., Vorus, N., & Frosch, A. (1999). The effectiveness of psychoanalytic psychotherapy: The role of treatment duration, frequency of sessions, and the therapeutic relationship. *Journal of the American Psychoanalytic Association*, 47(3), 741-772. doi:10.1177/00030651990470031001
- Friman, M. (2004). The structure of affective reactions to critical incidents. *Journal of Economic Psychology*, 25(3), 331-353. doi:10.1016/S0167-4870(03)00012-6
- Fuertes, J. N., Stracuzzi, T. I., Bennett, J., Scheinholtz, J., Mislouack, A., Hersh, M., & Cheng, D. (2006). Therapist multicultural competency: A study of therapy dyads. *Psychotherapy: Theory, Research, Practice, Training*, 43(4), 480-490. doi:10.1037/0033-3204.43.4.480
- Gao, L. L., Luo, S. Y., & Chan, S. W. (2012). Interpersonal psychotherapy-oriented program for Chinese pregnant women: Delivery, content, and personal impact. *Nursing & Health Sciences*, 14(3), 318-324. doi:10.1111/j.1442-2018.2012.00722.x
- Garfield, S. L. (1983). Some comments on consumer satisfaction in behavior therapy. *Behavior Therapy*, 14(2), 237-241. doi:10.1016/S0005-7894(83)80113-0
- Garrett, T., Oliver, C., Wilcox, D. T., & Middleton, D. (2003). Who cares? The views of sexual offenders about the group treatment they receive. *Sexual Abuse: A Journal of Research and Treatment*, 15(4), 323-338. doi: 10.1177/107906320301500408
- Gati, I. (1996). Computer-assisted career counseling: Challenges and prospects. In M. L. Savickas & B. W. Walsh (Eds.), *Handbook of career theory and practice* (pp. 169-190). Palo Alto, CA: Davies-Black.
- Gati, I., Gadassi, R., & Shemesh, N. (2006). The predictive validity of a computer-assisted career decision-making system: A six-year follow-up. *Journal of Vocational Behavior*, 68(2), 205-219. doi:10.1016/j.jvb.2005.08.002
- Gottlieb, B. H., & Wachala, E. D. (2007). Cancer support groups: A critical review of empirical studies. *Psycho-Oncology*, 16, 379-400. doi:10.1002/pon.1078
- Hamidzadeh, M. R., Jazani, N., Hajikarimi, A., & Ebrahimi, A. (2011). A proposed model for explanation of influential factors on customer satisfaction in banking industry. *Interdisciplinary Journal of Contemporary Research in Business*, 3(4), 136-143. Retrieved from <http://journal-archives7.webs.com/Aug011.pdf>
- Harris, G., Poertner, J., & Joe, S. (2000). The Parents With Children In Foster Care Satisfaction Scale. *Administration in Social Work*, 24(2), 15-27. doi:10.1300/J147v24n02\_02
- Hawkins, E. J., Baer, J. S., & Kivlahan, D. R. (2008). Concurrent monitoring of psychological distress and satisfaction measures as predictors of addiction treatment retention. *Journal of Substance Abuse Treatment*, 35(2), 207-216. doi:10.1016/j.jsat.2007.10.001
- Heinze, H. J., Jozefowicz, D. M. H., & Toro, P.A. (2010). Taking the youth perspective: Assessment of program characteristics that promote positive development in homeless and at-risk youth. *Children and Youth Services Review*, 32(10), 1365-1372. doi:10.1016/j.chilyouth.2010.06.004
- Helson, H. (1964). *Adaptation-level theory*. New York, NY: Harper & Row.

- Homburg, C., Koschate, N., & Hoyer, W. D. (2005). Do satisfied customers really pay more? A study of the relationship between customer satisfaction and willingness to pay. *Journal of Marketing*, 69(2), 84-96. Retrieved from <http://www.jstor.org/stable/30162046>
- Hood, C. (1991). A public management for all seasons? *Public Administration*, 69(1), 3-19. doi:10.1111/j.1467-9299.1991.tb00779.x
- Horvath, A. O. (1994). Research on the alliance. In A. O. Horvath & L. S. Greenberg (Eds.), *The Working Alliance: Theory, Research, and Practice* (pp. 259-286). New York, NY: John Wiley & Sons.
- Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*, 36(2), 223-233. doi:10.1037/0022-0167.36.2.223
- Howard, J. A., & Sheth, J. N. (1969). *The theory of buyer behavior*. New York, NY: Wiley.
- Hsieh, C. M. (2006). Using client satisfaction to improve case management services for the elderly. *Research on Social Work Practice*, 16(6), 605-612. doi:10.1177/1087057106289360
- Hsieh, C. M. (2012). Incorporating perceived importance of service elements into client satisfaction measures. *Research on Social Work Practice*, 22(1) 93-99. doi:10.1177/1049731511416826
- Hsieh, C., & Guy, M. E. (2009). Performance outcomes: The relationship between managing the "heart" and managing client satisfaction. *Review of Public Personnel Administration*, 29(1), 41-57. doi:10.1177/0734371X08326767
- Hunt, H. K. (1977a). Overview and future research directions. In H. K. Hunt (Ed.), *Conceptualization and measurement of consumer satisfaction and dissatisfaction*. Cambridge, MA: Marketing Science Institute.
- Hunt, H. K. (1977b). Bits and pieces. In R. L. Day (Ed.), *Consumer satisfaction, dissatisfaction, and complaining behavior*. Bloomington, IN: Indiana University Press.
- Ince, E. E., Rubin, D., & Christian, C. W. (2010). Parental perceptions of hospital care in children with accidental or alleged non-accidental trauma. *Child Abuse & Neglect* 34, 403-406. doi:10.1016/j.chiabu.2009.10.008
- Ingram, B. L., & Chung, R. S. (1997). Client satisfaction data and quality improvement in managed mental health care organizations. *Health Care Management Review*, 22(3), 40-52.
- Jackson, A. C., Thomas, S. A., Thomason, N., Borrell, J., Crisp, B. R., Ho, W., et al (2000a) *Longitudinal evaluation of the effectiveness of problem gambling counseling services, community education strategies and information products – Volume 2: Evaluation of the effectiveness of the counselling interventions at BreakEven problem gambling counselling services*. Melbourne, AU: Victoria Department of Human Services.
- James, O. (2009). Evaluating the expectations disconfirmation and expectations anchoring approaches to citizen satisfaction with local public services. *Journal of Public Administration Research and Theory*, 19(1), 107-123. doi:10.1093/jopart/mum034
- Johnson, M. D., & Fornell, C. (1991). A framework for comparing customer satisfaction across individuals and product categories. *Journal of Economic Psychology*, 12(2), 267-286. doi:10.1016/0167-4870(91)90016-M

- Johnson, M. D., Gustafsson, A., Andreassen, T., Lervik, L., & Cha, J. (2001). The evolution and future of national customer satisfaction index models. *Journal of Economic Psychology*, 22(2), 217-245. doi:10.1016/S0167-4870(01)00030-7
- Jones, L. M., Cross, T. P., Walsh, W. A., & Simone, M. (2007). Do children's advocacy centers improve families' experiences of child sexual abuse investigations? *Child Abuse & Neglect*, 31(10), 1069-1085. doi:10.1016/j.chiabu.2007.07.003
- Kapp, S. A., & Vela, R. H. (2003). The Parent Satisfaction with Foster Care Services Scale. *Child Welfare*, 83(3), 263-287.
- Kapp, S. A., & Vela, R. H. (2004). The unheard client: Assessing the satisfaction of parents of children in foster care. *Child & Family Social Work*, 9(2), 197-206. doi:10.1111/j.1365-2206.2004.00323.x
- Kelley, M. L., Heffer, R. W., Gresham, F. M., & Elliott, S. N. (1989). Development of a modified treatment evaluation inventory. *Journal of Psychopathology and Behavioral Assessment*, 11(3), 235-247. doi:10.1007/BF00960495
- Kern, J. K., Fletcher, C. L., Garver, C. R., Mehta, J. A., Grannemann, B. D., Knox, K. R., . . . Trivedi, M. H. (2011). Prospective trial of equine-assisted activities in autism spectrum disorder. *Alternative Therapies in Health and Medicine*, 17(3), 14-20.
- Kiesler, C. A. (1983). Social psychological issues in studying consumer satisfaction with behavior therapy. *Behavior Therapy*, 14(2), 226-236. doi:10.1016/S0005-7894(83)80112-9
- King, J. A., & Bond, T. G. (2003). Measuring client satisfaction with public education I: Meeting competing demands in establishing state-wide benchmarks. *Journal of Applied Measurement*, 4(2), 111-123.
- Kivlighan Jr, D. M., & Tarrant, J. M. (2001). Does group climate mediate the group leadership – group member outcome relationship?: A test of Yalom's hypotheses about leadership priorities. *Group Dynamics: Theory, Research, and Practice*, 5(3), 220-234. doi:10.1037/1089-2699.5.3.220
- Kivlighan Jr., D. M., London, K., & Miles, J. R. (2012). Are two heads better than one? The relationship between number of group leaders and group members, and group climate and group member benefit from therapy. *Group Dynamics: Theory, Research, and Practice*, 16(1), 1-13. doi:10.1037/a0026242
- Koch, L. C., & Rumrill, P. D. (2008). Assessing consumer satisfaction in rehabilitation and allied health care settings. *Work: Journal of Prevention, Assessment & Rehabilitation*, 31(3), 357-363.
- Krampf, R., Ueltschy, L., & d'Amico, M. (2002). The contribution of emotion to consumer satisfaction in the service setting. *Marketing Management Journal*, 13(1), 32-52.
- Larsen, D. L., Attkisson, C. C., Hargreaves, W. A., & Nguyen, T. D. (1979). Assessment of client/patient satisfaction: Development of a general scale. *Evaluation and Program Planning*, 2(3), 197-207. doi:10.1016/0149-7189(79)90094-6
- Lawson, D. M., & Brossart, D. F. (2003). The relationship between counselor trainee family-of-origin structure and counseling effectiveness. *Clinical Supervisor*, 22(2), 21-36. doi:10.1300/J001v22n02\_03
- Lebow, J. L. (1982). Consumer satisfaction with mental health treatment. *Psychological Bulletin*, 91(2), 244-259. doi:10.1037/0033-2909.91.2.244
- Lebow, J. L. (1983). Research assessing consumer satisfaction with mental health treatment: A review of findings. *Evaluation and Program Planning*, 6, 211-236. doi:10.1016/0149-7189(83)90003-4

- Lee, L. Y., & Holroyd, E. (2009). Evaluating the effect of childbirth education class: Amixed-method study. *International Nursing Review*, 56(3), 361-368. doi:10.1111/j.1466-7657.2008.00701.x
- Leelakulthanit, O., & Hongcharu, B. (2011). Factors that impact customer satisfaction: Evidence from the Thailand mobile cellular network industry. *International Journal of Management and Marketing Research*, 4(2), 67-76.
- Lees, D. G., & Ronan, K. R. (2008). Engagement and effectiveness of parent management training (incredible years) for solo high-risk mothers: A multiple baseline evaluation. *Behaviour Change*, 25(2), 109-128. doi:10.1375/beh.25.2.109
- Levenson, J. S., MacGowan, M. J., Morin, J. W., & Cotter, L. P. (2009). Perceptions of sex offenders about treatment: Satisfaction and engagement in group therapy. *Sexual Abuse: A Journal of Research and Treatment*, 21(1), 35-56.
- Levenson, J. S., Prescott, D. S., & D'Amora, D. A. (2010). Sex offender treatment: Consumer satisfaction and engagement in therapy. *International Journal of Offender Therapy and Comparative Criminology*, 54(3), 307-326. doi:10.1177/0306624X08328752
- Linden, J. D., Stone, S. C., & Shertzer, B. (1965). Development and evaluation of an inventory for rating counseling. *Personnel and Guidance Journal*, 44(3), 267-276. doi:10.1002/j.2164-4918.1965.tb03516.x
- Linder-Pelz, S. (1982). Social psychological determinants of patient satisfaction: A test of five hypotheses. *Social Science & Medicine*, 16(5), 583-589. doi:10.1016/0277-9536(82)90312-4
- Liu, L. M., Guarino, A. J., & Lopez, R. (2012). Family satisfaction with care provided by nurse practitioners to nursing home residents with dementia at the end of life. *Clinical Nursing Research*, 21(3), 350-367. doi:10.1177/1054773811431883
- Locker, D., & Dunt, D. (1978). Theoretical and methodological issues in sociological studies of consumer satisfaction with medical care. *Social Science & Medicine*, 12(4-A), 283-292. doi:10.1016/0271-7123(78)90067-6
- Lunnen, K. M., & Ogles, B. M. (1998). A multiperspective, multivariable evaluation of reliable change. *Journal of Consulting and Clinical Psychology*, 66(2), 400-410. doi:10.1037/0022-006X.66.2.400
- Lunnen, K. M., Ogles, B. M., & Pappas, L. N. (2008). A multiperspective comparison of satisfaction, symptomatic change, perceived change, and end-point functioning. *Professional Psychology: Research and Practice*, 39(2), 145-152. doi:10.1037/0735-7028.39.2.145
- Mano, H., & Oliver, R. L. (1993). Assessing the dimensionality and structure of the consumption experience: Evaluation, feeling, and satisfaction. *Journal of Consumer Research*, 20, 451-466.
- Martensen, A., Gronholdt, L., & Kristensen, K. (2000). The drivers of customer satisfaction and loyalty: Cross-industry findings from Denmark. *Total Quality Management*, 11, 8544-8553. doi: 10.1080/09544120050007878
- Martin, J. S., Petr, C. G., & Kapp, S. A. (2003). Consumer satisfaction with children's mental health services. *Child & Adolescent Social Work Journal*, 20(3), 211-226. doi:10.1023/A:1023609912797
- McMurtry, S. L., & Hudson, W. W. (2000). The Client Satisfaction Inventory: Results of an initial validation study. *Research on Social Work Practice*, 10(5), 644-663.
- McNeill, T., Nicholas, D., Szechy K., & Lach, L. (1998). Perceived outcome of social work intervention: Beyond consumer satisfaction. *Social Work in Health Care*, 26(3), 1-18. doi:10.1300/J010v26n03\_01

- Melnick, G., Hawke, J., & Wexler, H. K. (2004). Client perceptions of prison-based therapeutic community drug treatment programs. *Prison Journal*, 84(1), 121-138. doi:10.1177/0032885503262459
- Melnick, G., & Pearson, F. (2000). A multimodality quality assurance instrument. New York, NY: National Development and Research Institutes.
- Mendez, J. L. (2010). How can parents get involved in preschool? Barriers and engagement in education by ethnic minority parents of children attending Head Start. *Cultural Diversity & Ethnic Minority Psychology*, 16(1), 26-36. doi:10.1037/a0016258
- Miller, J. K. (2008). Walk-in single session team therapy: A study of client satisfaction. *Journal of Systemic Therapies*, 27(3), 78-94.
- Miller, J. K., & Slive, A. (2004). Breaking down the barriers to clinical service delivery: Walk-in family therapy. *Journal of Marital and Family Therapy*, 30(1), 95-103. doi:10.1521/jsyt.2008.27.3.78
- Moore, M. (1995). *Creating public value*. Cambridge, MA: Harvard University Press.
- Morgeson, F. V., III, & Petrescu, C. (2011). Do they all perform alike? An examination of perceived performance, citizen satisfaction, and trust with U.S. federal agencies. *International Review of Administrative Sciences*, 77(3), 451-479. doi:10.1177/0020852311407278
- Morgeson III, F. V., (2013). Expectations, disconfirmation, and citizen satisfaction with the U.S. federal government: Testing and expanding the model. *Journal of Public Administration Research and Theory*, 23(2), 289-305. doi:10.1093/jopart/mus012
- Murphy, L., Parnass, P., Mitchell, D. L., Hallett, R., Cayley, P., & Seagram, S. (2009). Client satisfaction and outcome comparisons of online and face-to-face counseling methods. *British Journal of Social Work*, 39(4), 627-640. doi:10.1093/bjsw/bcp041
- Murphy, M. J., Faulkner, R. A., & Behrens, C. (2004). The effect of therapist-client racial similarity on client satisfaction and therapist evaluation of treatment. *Contemporary Family Therapy*, 26(3), 279-292. doi:10.1023/B:COFT.0000037915.95303.28
- Nelson, T. D., & Steele, R. G. (2006). Beyond efficacy and effectiveness: A multifaceted approach to treatment evaluation. *Professional Psychology: Research and Practice*, 37(4), 389-397. doi:10.1037/0735-7028.37.4.389
- New York City Department of Education. (2008). *2008 Parents survey*. Retrieved from <http://schools.nyc.gov/NR/rdonlyres/2E0CEA9C-79D6-4668-B277-1F1711ED08F2/33766/Parent20survey202008.pdf>
- Nguyen, T. D., Attkisson, C. C., & Stegner, B. L. (1983). Assessment of patient satisfaction: Development and refinement of a service evaluation questionnaire. *Evaluation and Program Planning*, 16, 109-118.
- Nielsen, S. L., Smart, D. W., Isakson, R. L., Worthen, V. E., Gregersen, A. T., & Lambert, M. J. (2004). The consumer reports effectiveness score: What did consumers report? *Journal of Counseling Psychology*, 51(1), 25-37. doi:10.1037/0022-0167.51.1.25
- Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.
- Ollendick, T. H., Öst, L., Reuterskiöld, L., Costa, N., Cederlund, R., Sirbu, C., ... Jarrett, M. A. (2009). One-session treatment of specific phobias in youth: A randomized clinical trial in the United States and Sweden. *Journal of Consulting and Clinical Psychology*, 77(3), 504-516. doi:10.1037/a0015158

- O'Loughlin, C., & Coenders, G. (2004). Estimation of the European Customer Satisfaction Index: Maximum likelihood versus partial least squares. Application to postal services. *Total Quality Management & Business Excellence*, 15(9/10), 1231-1255. doi:10.1080/1478336042000255604
- Olshavsky, R. W., & Miller, J. A. (1972). Consumer expectations, product performance, and perceived product quality. *Journal of Marketing Research*, 9(1), 19-21. doi:10.2307/3149600
- Parloff, M. B. (1983). Who will be satisfied by 'consumer satisfaction' evidence?. *Behavior Therapy*, 14(2), 242-246. doi:10.1016/S0005-7894(83)80114-2
- Pascoe, G. C. (1983). Patient satisfaction in primary health care: A literature review and analysis. *Evaluation and Program Planning*, 6(3-4), 185-210. doi:10.1016/0149-7189(83)90002-2
- Pascoe, G. C., & Attkisson, C. C. (1983). The evaluation ranking scale: A new methodology for assessing satisfaction. *Evaluation and Program Planning*, 6(3), 335-347. doi: 10.1016/0149-7189(83)90013-7
- Pekarik, G., & Wolff, C. B. (1996). Relationship of satisfaction to symptom change, follow-up adjustment, and clinical significance. *Professional Psychology: Research and Practice*, 27(2), 202-208. doi:10.1037/0735-7028.27.2.20
- Perschel, A. (2001). *Client satisfaction: The development of a survey utilizing client inputs*. Unpublished doctoral dissertation. Antioch New England Graduate School, Keene, NH.
- Poister, T. H., & Thomas, J. C. (2011). The effect of expectations and expectancy confirmation/disconfirmation on motorists' satisfaction with state highways. *Journal of Public Administration Research and Theory*, 21(4), 601-617. doi:10.1093/jopart/mur004
- Ponterotto, J. G., & Furlong, M. J. (1985). Evaluating counselor effectiveness: A critical review of rating scale instruments. *Journal of Counseling Psychology*, 32(4), 597-616. doi:10.1037/0022-0167.32.4.597
- Reid, P., & Gundlach, J. H. (1983). A scale for the measurement of consumer satisfaction with social services. *Journal of Social Service Research*, 7(1), 37-54. doi:10.1300/J079v07n01\_03
- Royse, D., Thyer, B. A., & Padgett, D. K. (2010). *Program evaluation: An introduction* (5th ed.). Belmont, CA: Wadsworth.
- Sackett, D. L. (1997, February 22). Evidence-based medicine and treatment choices. *Lancet*, 349(9051), 572-573. doi.org/10.1016/S0140-6736(97)80122-5
- Schiff, M., Witte, S. S., & El-Bassel, N. (2003). Client satisfaction and perceived helping components of an HIV/AIDS preventive intervention for urban couples. *Research on Social Work Practice*, 13(4), 468-492. doi:10.1177/1049731503013004004
- Schraufnagel, S., & Li, Q. (2010). Testing the suitability of mediation of child support orders in Title IV-D cases. *Research on Social Work Practice*, 20(2), 212-222. doi:10.1177/1049731509339029
- Seashore, S. E., Lawler, E. E., Mirvis, P., & Cammann, C. (Eds.). (1982). *Observing and measuring organizational change: A guide to field practice*. New York, NY: Wiley.
- Seligman, M. E. (1996). Science as an ally of practice. *American Psychologist*, 51(10), 1072-1079. doi:10.1037/0003-066X.51.10.1072



- Shapiro, J. P., Welker, C. J., & Jacobson, B. J. (1997). The Youth Client Satisfaction Questionnaire: Development, construct validation, and factor structure. *Journal of Clinical Child Psychology*, 26, 87-98. doi:10.1207/s15374424jccp2601\_9
- Sharma, V., Whitney, D., Kazarian, S. S., & Manchanda, R. (2000). Preferred terms for users of mental health services among service providers and recipients. *Psychiatric Services*, 51(2), 203-209. doi:10.1176/appi.ps.51.2.203
- Shaw, I. F. (1984). Literature review: Consumer evaluations of the personal social services. *British Journal of Social Work*, 14(3), 277-284.
- Shek, D. L. (2010). Subjective outcome and objective outcome evaluation findings: Insights from a Chinese context. *Research on Social Work Practice*, 20(3), 293-301. doi:10.1177/1049731509331951
- Shek, D. T., Siu, A. M., & Lee, T. Y. (2007). The Chinese Positive Youth Development Scale: A validation study. *Research on Social Work Practice*, 17(3), 380-391. doi:10.1177/1049731506296196
- Sikorska-Simmons, E. (2001). Development of an instrument to measure resident satisfaction with assisted living. *Journal of Applied Gerontology*, 20(1), 57-73. doi:10.1177/073346480102000104
- Sikorska-Simmons, E. (2006). Linking resident satisfaction to staff perceptions of the work environment in assisted living: A multilevel analysis. *Gerontologist*, 46(5), 590-598. doi:10.1093/geront/46.5.590
- Sitzia, J. (1999). How valid and reliable are patient satisfaction data? An analysis of 195 studies. *International Journal of Qualitative Health Care*, 11(4), 319-328. doi:10.1093/intqhc/11.4.319
- Smith, S. A., Thomas, S. A., & Jackson, A. C. (2004). An exploration of the therapeutic relationship and counselling outcomes in a problem gambling counselling service. *Journal of Social Work Practice*, 18(1), 99-112. doi:10.1080/0265053042000180581
- Sorensen, J., Done, D. J., & Rhodes, J. (2007). A case series evaluation of a brief, psycho-education approach intended for the prevention of relapse in bipolar disorder. *Behavioural and Cognitive Psychotherapy*, 35(1), 93-107. doi:10.1017/S1352465806003316
- Spiro, S. E., Dekel, R., & Peled, E. (2009). Dimensions and correlates of client satisfaction: An evaluation of a shelter for runaway and homeless youth. *Research on Social Work Practice*, 19(2), 261-270. doi:10.1177/1049731508329395
- Sprenkle, D. H., Constantine, J. A., & Piercy, F. P. (1982) *Purdue Live Observation Satisfaction Scale*. Unpublished instrument. Purdue University, West Lafayette, IN.
- Stoker, G. (2006). Public value management: A new narrative for networked governance? *American Review of Public Administration*, 36(1), 41-57. doi:10.1177/0275074005282583
- Strand, V. C., & Badger, L. (2005). Professionalizing child welfare: An evaluation of a clinical consultation model for supervisors. *Children and Youth Services Review*, 27(8), 865-880. doi:10.1016/j.childyouth.2004.12.001
- Stüntzner-Gibson, D., Koren, P. E., & DeChillo, N. (1995). The Youth Satisfaction Questionnaire: What kids think of services. *Families in Society*, 76, 616-624.
- Szymanski, D. M., & Henard, D. H. (2001). Customer satisfaction: A meta-analysis of the empirical evidence. *Journal of the Academy of Marketing Science*, 29(1), 16-35.

- Trotter, C. (2008). What does client satisfaction tell us about effectiveness?. *Child Abuse Review*, 17(4), 262-274. doi:10.1002/car.1038
- Trute, B., & Hiebert-Murphy, D. (2007). The implications of 'working alliance' for the measurement and evaluation of family-centered practice in childhood disability services. *Infants & Young Children*, 20(2), 109-119. doi:10.1097/01.IYC.0000264479.50817.4b
- Turco, T. L., & Elliott, S. N. (1986a). Assessment of students' acceptability ratings of teacher-initiated interventions for classroom misbehavior. *Journal of School Psychology*, 24(3), 277-283. doi:10.1016/0022-4405(86)90060-9
- Turco, T. L., & Elliott, S. N. (1986b). Students' acceptability ratings of interventions for classroom misbehaviors: A study of well-behaving and misbehaving youth. *Journal of Psychoeducational Assessment*, 4(4), 281-289. doi:10.1177/073428298600400404
- Van Doorn, J., Leeftang, P. S. H., & Tijs, M. (2013). Satisfaction as a predictor of future performance: A replication. *International Journal of Research in Marketing*, Advance online publication. doi:10.1016/j.ijresmar.2013.04.002
- Van Ryzin, G. G. (2006). Testing the expectancy disconfirmation model of citizen satisfaction with local government. *Journal of Public Administration Research & Theory*, 16(4), 599-611. doi:10.1093/jopart/mui058
- Västfjäll, D., Friman, M., Gärling, T., & Kleiner, M. (2002). The measurement of core affect: A Swedish self-report measure derived from the affect circumplex. *Scandinavian Journal of Psychology*, 43(1), 19-31. doi:10.1111/1467-9450.00265
- Volicer, L., Hurley, A. C., & Blasi, Z. V. (2001). Scales for evaluation of end-of-life care in dementia. *Alzheimer Disease & Associated Disorders*, 15(4), 194-200. doi:10.1097/00002093-200110000-00005
- Walsh, T., & Lord, B. (2004). Client satisfaction and empowerment through social work intervention. *Social Work in Health Care*, 38(4), 37-56. doi:10.1300/J010v38n04\_03
- Webster-Stratton, C. (1999). *The parent and child series handbook*. Seattle, WA: Seth Enterprises.
- Webster-Stratton, C., Reid, M. J., & Hammond, M. (2004). Treating children with early-onset conduct problems: Intervention outcomes for parent, child, and teacher training. *Journal of Clinical Child & Adolescent Psychology*, 33(1), 105-124. doi:10.1207/S15374424JCCP3301\_11
- Weinbach, R. W. (2005). *Evaluating social work services and programs*. Boston, MA: Allyn & Bacon.
- Westbrook, D., McManus, F., Clark, G., & Bennett-Levy, J. (2012). Preliminary evaluation of an online training package in cognitive behaviour therapy: Satisfaction ratings and impact on knowledge and confidence. *Behavioural and Cognitive Psychotherapy*, 40(4), 481-490. doi:10.1017/S1352465811000701
- Wilkinson, L. A. (2003). Using behavioral consultation to reduce challenging behavior in the classroom. *Preventing School Failure: Alternative Education for Children and Youth*, 47(3), 100-105. doi:10.1080/10459880309604437
- Wilkinson, L. A. (2005). Supporting the inclusion of a student with Asperger Syndrome: A case study using conjoint behavioural consultation and self-management. *Educational Psychology in Practice*, 21(4), 307-326. doi:10.1080/02667360500344914
- Williams, B. (1994). Patient satisfaction: A valid concept? *Social Science & Medicine*, 38(4), 509-516. doi:10.1016/0277-9536(94)90247-X

- World Health Organization. (2000). *Client satisfaction evaluation*. Retrieved from [www.emcdda.europa.eu/.../att\\_5868\\_EN\\_6\\_client\\_satisfaction\\_evaluation](http://www.emcdda.europa.eu/.../att_5868_EN_6_client_satisfaction_evaluation)
- Yi, Y. (1990). A critical review of consumer satisfaction. In V. A. Zeithaml (Ed.), *Review of marketing* (pp. 68-123). Chicago, IL: American Marketing Association.
- Young, S. C., 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(2), 219-238. doi:10.1007/BF02234097
- Yu, C. K. (2005). Application of cognitive-behavioural hypnotic treatment for anxiety management in the Chinese. *Contemporary Hypnosis*, 22(2), 104-114. doi:10.1002/ch.30
- Zhang, J., & Bloemer, J. M. M. (2008). The impact of value congruence on consumer-service brand relationships. *Journal of Service Research*, 11(2), 161-178. doi:10.1177/1094670508322561

## **Appendix A: Summary of Consumer Satisfaction Instruments, 2003 – 2013**

Instrument (‡ denotes exact name is not clear)	Author (*example study, if multiple reports) Original Citation Subscales	Purpose	Study Description	Validation Sample	a. No. of Items; b. Response scales c. Score range; d. Scoring	Validity	Reliability ( $\alpha$ or $r$ )
1. Treatment Evaluation Inventory-Short Form;  <b>TEI-SF</b>	Beavers, Kratochwill & Braden (2004) Kelly et al. (1989) Subscales: None described	Evaluate consumer perceptions of treatment acceptability and treatment effectiveness	Multi-site U.S.-based study. Data collected via face-to-face interviews with 18 teachers and 32 students with reading difficulties. Students received academic consultation services. Study sites were 6 elementary schools: 2 in Wisconsin and 4 in Tennessee. Analysis used a two-sample Wilcoxon test to examine the relationship between treatment dimensions and program satisfaction.	<b>Teacher:</b> Age ( $M$ ): Not available or reported (N/A) Gender: 94% female Race/ethnicity: N/A Income: N/A Other: 44% private school; 56% public school; Avg. teaching experience: 10 yrs. <b>Student:</b> Age: Elementary school age (6-12 yrs.) Gender: 38% female Race/ethnicity: 78% Caucasian; 12% African American. Income: range from low to high socioeconomic status (SES)	a. 9 items b. 5 points -- <i>strongly disagree</i> (1) to <i>strongly agree</i> (5) c. 9 – 45 d. summed	Authors report that, based on prior research, the TEI-SF is has acceptable construct validity (Kelly et al., 1989)	0.85 (cited from Kelly et al., 1989).
2. Children's Advocacy Center Nonoffending Caregiver Satisfaction Survey;  <b>CAC</b>	Bonach, Mabry & Potts-Henry (2010) Cross et al. (2008) <b>3 Subscales:</b> • Satisfaction with CAC Services • Satisfaction with the Multidisciplinary Team (MDT) • Overall Satisfaction with CAC Experience	Evaluate non-offending caregivers' satisfaction with services they and their children receive from CAC MDT program in response to allegations of child abuse, particularly sexual abuse	U.S.-based study. Data collected via mailed survey questionnaire sent to 26 non-offending caregivers in a rural community in the Eastern U.S. Analyses use bivariate correlation and linear regression to examine the relationship between program dimensions and satisfaction.	<b>Children:</b> Age ( $M$ ): 11 Gender: 58% female Race/ethnicity: N/A Income: N/A <b>Offender:</b> Age ( $M$ ): 31 Gender: 100% male Race/ethnicity: N/A Income: N/A Other: 46% referral source from child welfare; 38% from police/law enforcement.	a. 244 items b. 4 points -- <i>very satisfied</i> (=4) to <i>very dissatisfied</i> (=1) to <i>not applicable</i> (=0) c. 0 – 4 d. mean	N/A	0.83 to 0.93
3. Consumer Satisfaction Questionnaire-12;  <b>CSQ-12</b>	Boyle et al. (2010) Therapy Attitude Inventory: Eyberg (1993) Subscales: None described	Examine parental satisfaction with the quality, ease of use, and appropriateness of Primary Care Triple P.	U.S.-based study; 2 sites. Face-to-face interviews with 10 children and their families (9 families) who participated in the Triple P intervention in two cities in the U.S. South. Uses ANOVAs, t-tests, and MANOVA to examine different phases of intervention effects.	N/A	a. 12 b. N/A c. N/A d. summed	N/A	0.96

Instrument (‡ denotes exact name is not clear)	Author (*example study, if multiple reports) Original Citation Subscales	Purpose	Study Description	Validation Sample	a. No. of Items; b. Response scales c. Score range; d. Scoring	Validity	Reliability ( $\alpha$ or $r$ )
4. Victim Satisfaction with Offender Dialogue Scale  VSODS	Bradshaw & Umbreit (2003) <b>3 Subscales:</b> • Mediator skills • Experience with meeting the offender • Satisfaction with the restitution plan	To examine the general satisfaction with victim offender mediated dialogue services, and to develop a reliable instrument.	U.S.-based. Data were collected from a mailed survey questionnaire of 197 participants from 4 victim offender mediation services. Principal components factor analysis was used to estimate the dimensionality of the VSODS.	Age ( $M$ ): 39 Gender: 37% female Race/ethnicity: N/A Income: N/A Other: Avg. yrs. of education: 15 years	a. 11 b. 4 points <i>very satisfied</i> (4) to <i>very dissatisfied</i> (1) c. 11 – 44 d. summed	N/A	0.87
5. Three-item Satisfaction Scale‡	Brenninkmeijer & Blonk (2012) Subscales: None described	To examine satisfaction with the content, guidance, and helpfulness of a jobs program	Netherlands-based study. Face-to-face interviews of 118 participants (47 in JOBS condition, 33 in the voucher condition, 38 in the control condition) in a suburban city.	Age ( $M$ ): 38 Gender: 70% female Race/ethnicity: 55% Dutch; 12% Antillean; 16% Surinamese; 17% other Income: 85% got benefit from social welfare; Other: Avg. yrs. of education: 15 years. 46% single; 35% divorced; 16% married	a. 3 b. 5 points c. 1 – 5 d. mean	N/A	0.86 to 0.89
6. Client Satisfaction Survey	Brooks & Brown (2005) Subscales: None described	To examine satisfaction with ease of accessing ACORN and with outcomes	U.S.-based study. Data collected via telephone survey with 99 welfare clients participating in ACORN's case advocacy program in Los Angeles, CA.	N/A	a. 8 Likert items, 1 open-ended item b. 5-point -- <i>strongly disagree</i> (1) to <i>strongly agree</i> (5) c. 1-5 d. mean	N/A	0.89
7. Eight-item Satisfaction Scale ‡	Butler, Gomon, & Turner (2004) Subscales: Not described	To examine the extent of satisfaction with 14 aspects of an assisted living program	U.S.-based study. Data collected via face-to-face interviews with 23 low-income residents of an assisted living facility in a rural Northeastern state. Paired $t$ -tests were used to examine change in service satisfaction over time	Age ( $M$ ): 77 Gender: 79% female Race/ethnicity: 97% European American; 3% Native American. Other: 28% divorced; 3% married; 59% widowed; 28% had some college more.	a. 8 items b. 4 points -- <i>poor</i> =1, <i>fair</i> =2, <i>good</i> =3, <i>excellent</i> =4 c. 1 – 4 d. mean	N/A	N/A
8. Resident Satisfaction Survey	Castle et al. (2004) <b>2 Subscales:</b> • Use of Resident Satisfaction Surveys • Usefulness of Resi-	To examine the use and usefulness of resident satisfaction surveys in licensed nursing homes and assisted living	U.S.-based study. Data collected via from a mailed survey ( $N=363$ ) of licensed nursing Homes ( $n=266$ ) and assisted living facilities ( $n=97$ ) in New Jersey.	<b>Nursing Homes:</b> Age ( $M$ ): N/A Gender: 57% female Race/ethnicity: 4% African American; 3% Asian; 88%	a. 10 items (4 use; 6 usefulness) b. Multiple choice answers c. N/A d. N/A	N/A	N/A

Instrument (‡ denotes exact name is not clear)	Author (*example study, if multiple reports) Original Citation Subscales	Purpose	Study Description	Validation Sample	a. No. of Items; b. Response scales c. Score range; d. Scoring	Validity	Reliability ( $\alpha$ or $r$ )
	dent Satisfaction Surveys	facilities.		Caucasian; 3% Hispanic Income: N/A Other: 94% had an associated degree or more. 74% had long-term care insurance <b>Assisted Living:</b> Age ( $M$ ): N/A Gender: 70% female Race/ethnicity: 2% African American; 3% Asian; 95% Caucasian Income: N/A Other: 91% had an associate degree or more; 62% had long-term care insurance.			
9. Parental satisfaction survey‡	Charbonneau & Van Ryzin (2012) New York City Department of Education (2008) Subscales: Not described	To examine parental satisfaction with the New York City public schools.	U.S.-based study. Data collected via mailed surveys ( $N=937$ ) sent to parents or guardians of students enrolled in New York City public elementary or middle schools. Ordinary least square regression was used to explore the relationship of parental satisfaction to other variables.	N/A	a. 13 items b. 4-point scale -- <i>strongly disagree/very unsatisfied</i> = 1; <i>strongly agree/very satisfied</i> = 4 c. N/A d. N/A	Authors argue that the instrument has acceptable face validity.	0.95
10. Client Satisfaction Inventory: Short-Form <b>CSI-SF</b>	Collins et al. (2005) McMurtry & Hudson (2000) Subscales: Not described	To examine satisfaction with housing services.	U.S.-based study. Data collected via mailed surveys ( $N=76$ ) sent to residents and staff of the HOPE VI housing development services in Boston, MA.	<b>Year 1:</b> Age ( $M$ ): 47 years Gender: 88% female Race/ethnicity: 64% Latino; 30% Black; 7% other Income: 38% yearly income < \$7749 Other: 52% unemployed	a. 9 items b. 7-point scale -- <i>none of the time</i> = 1 to <i>all of the time</i> = 7 c. 1 – 7 d. mean	Authors report that the instrument has acceptable content and discriminant validity.	0.89 (Year 2 data is 0.92)
11. Community Satisfaction Scale <b>CSI</b>	Collins et al. (2005) Subscales: Not described	To examine satisfaction with housing services.	U.S.-based study. Data collected via mailed surveys ( $N=76$ ) sent to residents and staff of the HOPE VI housing	<b>Year 1:</b> Age ( $M$ ): 47 years Gender: 88% female Race/ethnicity: 64% Latino; 30% Black; 7% other	a. 17 items b. 18 (N/A) c. 0 – 17 d. mean	N/A	0.67

Instrument (‡ denotes exact name is not clear)	Author (*example study, if multiple reports) Original Citation Subscales	Purpose	Study Description	Validation Sample	a. No. of Items; b. Response scales c. Score range; d. Scoring	Validity	Reliability ( $\alpha$ or $r$ )
			development services in Boston, MA.	Income: 38% yearly income < \$7749 Other: 52% unemployed			
12. Satisfaction with Management Scale	Collins et al. (2005) Subscales: Not described	To examine satisfaction with housing services.	U.S.-based study. Data collected via mailed surveys (N=76) sent to residents and staff of the HOPE VI housing development services in Boston, MA.	<b>Year 1:</b> Age (M): 47 years Gender: 88% female Race/ethnicity: 64% Latino; 30% Black; 7% other Income: 38% yearly income < \$7749 Other: 52% unemployed	a. 6 items b. 5-point scale (N/A) c. 0 – 5 d. mean	N/A	0.83
13. Satisfaction with Employment Scale: Short Form	Collins, et al (2005) McMurtry & Hudson (2000) Subscales: Not described	To examine satisfaction with housing services y.	U.S.-based study. Data collected via mailed surveys (N=76) sent to residents and staff of the HOPE VI housing development services in Boston, MA.	<b>Year 1:</b> Age (M): 47 years Gender: 88% female Race/ethnicity: 64% Latino; 30% Black; 7% other Income: 38% yearly income < \$7749 Other: 52% unemployed	a. 9 item b. 5-point scale -- <i>very dissatisfied</i> = 1 to <i>very satisfied</i> = 5 c. 1 – 5 d. mean	N/A	0.79
14. Satisfaction Survey ‡	Coloma, Gibson, & Packard (2012) Subscales: Not described	To examine satisfaction with a leadership development program.	U.S.-based study. Data collected via face-to-face interviews (N=166) with residents and staff of HOPE VI housing development services in Boston, MA.	Age: 42%, 40-49 years; 37%, 50-59 years Gender: 59% female Race/ethnicity: 46% White,; 22% Hispanic/Latino; 21% African American; 10% Asian/Pacific Islander Income: N/A Other: 48% master's degree; 27% bachelor's degree; 13% some college education, 7% some graduate education, and 2% PhD	a. 2 items b. 5-point scale (5 was the highest rating) c. 1 – 5 d. mean	N/A	N/A
15. Behavior Intervention Rating Scale - Treatment Acceptability  <b>BIRS</b>	Cowan & Sheridan (2003)* Wilkinson (2005) Von Brock & Elliott (1987) <b>3 Subscales:</b> • Acceptability • Effectiveness • Time to Effect	To assess the acceptability of an educational program for children at risk of academic failure.	U.S.-based study. Secondary analysis of data collected from parents (n=45) with a child reported as at-risk for academic failure; data were also collected from the students' teachers (n=62). Study included 6 large school districts; 4 districts	<b>Parents:</b> Age (M): 37 years Gender: 76% female Race/ethnicity: 87% Caucasian; 13% others. Income: N/A <b>Teachers:</b> Age (M): 41 years Gender: 87% female	a. 24 items b. 6-point scale -- <i>strongly disagree</i> = 1; <i>strongly disagree</i> = 6 c. 1 – 6 d. mean	Authors report that construct validity was found acceptable in previous reports (Turco & Elliott, 1986a, 1986b).	Total scale: 0.97 Subscales: 0.97 for Acceptability; 0.92 for Effectiveness; .87 for Time to Effect



Instrument (‡ denotes exact name is not clear)	Author (*example study, if multiple reports) Original Citation Subscales	Purpose	Study Description	Validation Sample	a. No. of Items; b. Response scales c. Score range; d. Scoring	Validity	Reliability ( $\alpha$ or $r$ )
			were in a large Western U.S. city, and 2 were in a mid-size Midwestern city.	Race/ethnicity: 96% Caucasian; 4% others Income: N/A Other: Avg. yrs teaching experience: 12 years; 59% bachelor degree; 39% master degree; 2% PhD			
16. 12-item Satisfaction Survey‡	Dauenhauer et al. (2007) <b>4 Subscales:</b> • Intake Process • Assessment • Case Management • Overall Satisfaction	To examine satisfaction with adult protective services (APS)	U.S.-based study. Data collected via mailed surveys ( $N=58$ ) sent to community professionals who had used APS within the past year in Monroe County, NY.	Age ( $M$ ): N/A Gender: 83% female Race/ethnicity: N/A Income: N/A Other: 96% had a bachelor's degree or higher; 59% were social workers	a. 12 items b. 4-point scale w/ higher scores indicating greater satisfaction (i.e., <i>poor</i> to <i>excellent</i> , <i>never</i> to <i>always</i> , or <i>very dissatisfied</i> to <i>very satisfied</i> ) c. 1 – 4 d. mean	N/A	0.95  intake = 0.84 assessment = 0.80 case management = 0.84 overall satisfaction = 0.91
17. Consumer Satisfaction Questionnaire-8  <b>CSQ-8</b>	Denton, Nakonezny, & Burwell (2011)* Larsen et al. (1979)  CSQ-8 also used in 9 other articles (see Table 1) Subscales: Not described	To examine the general client satisfaction with marriage and family therapy	U.S.-based study. Data collected via face-to-face interviews with ( $N=86$ ) clients who participated the marriage and family therapy. Analyses using regression and ANCOVA were conducted to examine whether client satisfaction was related to meeting/not meeting the therapist supervision team.	<b>Meet team:</b> Age ( $M$ ): 35 years Gender: 54% female Race/ethnicity: 63% White; 37% non-White Income: 43% < \$19,900; 46% = \$20,000-\$59,900; 5% => \$60,000 Other: Avg. length of marriage: 6 years <b>Did Not Meet Team:</b> Age ( $M$ ): 36 years. Gender: 53% female Race/ethnicity: 80% White; 20% non-White Income: 25% < \$19,900; 55% = \$20,000-\$59,900; 8% = > \$60,000 Other: Avg. length of marriage: 9 years	a. 8 items b. 4-point scale, with higher scores indicating greater satisfaction anchor wording differs, by item) c. 8 – 32 d. summed	Authors report that the CSQ-8 has been found to have acceptable reliability and validity (cited from Attkisson & Zwick, 1982).	0.86
18. Working Alliance Inventory-Short version;	Fuertes et al. (2006)* Dearing et al. (2005) Horvath & Greenberg (1989)	To examine client-therapist bonding and goal agreement	U.S.-based study. Data collected via face-to-face interviews with 51 client – therapist dyads at three	<b>Clients:</b> Age ( $M$ ): 27 years Gender: 71% female Race/ethnicity: 24% Euro	a. 12 items b. 7- point scale <i>never</i> (= 1) to <i>always</i> (= 7)	Authors report that prior research has shown strong support for the	0.74

Instrument (‡ denotes exact name is not clear)	Author (*example study, if multiple reports) Original Citation Subscales	Purpose	Study Description	Validation Sample	a. No. of Items; b. Response scales c. Score range; d. Scoring	Validity	Reliability ( $\alpha$ or $r$ )
<b>WAI-S</b>	Subscales: Not described		university outpatient clinics in an urban Northeast city.	Americans; 33% Asian Americans; 27% African Americans; 16% Hispanic Americans Income: N/A <u>Therapists:</u> Age ( $M$ ): 32 years Gender: 53% female Race/ethnicity: 67% Euro Americans; 24% Asian Americans; 8% African Americans; 2% Hispanic Americans Income: N/A	c. 12 – 84 d. summed	psychometric properties of the WAI (e.g., Horvath, 1994), and that the WAI has been demonstrated to correlate with a variety of parallel outcome measures (Horvath & Greenberg, 1989).	
19. Purdue Live Observation Satisfaction Scale  <b>PLOSS</b>	Denton, Nakonezny, & Burwell (2011) Sprenkle et al. (1982) Subscales: Not described	To examine client – therapist bonding and goal agreement	U.S.-based study. Data collected via face-to-face interviews and mailed surveys with $n=86$ clients who received an initial therapy session with therapist trainee. Clients were assigned to 2 conditions; the <b>Meet</b> condition ( $n = 46$ ) met the trainee's supervisory team; the <b>Did Not Meet</b> ( $n=40$ ) condition did not meet the team. A mixed linear model of covariance was used to test client – therapist relationship among dyads in these two conditions and the extent of satisfaction.	<u>Meet team:</u> Age ( $M$ ): 35 years Gender: 54% female Race/ethnicity: 63% White; 37% Other. Yearly Income: 43% $\leq$ \$19,900; 11% $\geq$ \$60,000. Other: 80% married.	a. 11 b. 5-point scale -- <i>strongly agree</i> (=1) to <i>strongly disagree</i> (= 5) c. 1 – 5 d. mean	N/A	0.67
20. Parenting Our Children to Excellence  <b>PACE</b>	Dumas et al. (2011) <u>4 Subscales:</u> •Satisfied with group leaders •Accepted program goals •Found program beneficial; •Would recommend program	To examine satisfaction in a parenting education program	U.S.-based study. Data collected via 2 face-to-face interviews with ( $N=124$ ) parents with a child between 3-6 yrs. old. Conducted in 2 Midwestern cities in the U.S. Analyses used ANOVAs to estimate relationships of various program dimensions with parent satisfaction.	Age ( $M$ ): 32 years Gender: 94% female Race/ethnicity: 98% Hispanic, 2% Caucasian. Income: (Median annual): \$15,000 to \$19,999. Other: 57% married; 43% single. 22% had college experience.	a. 8 items b. 5-point scale -- (higher numbers indicate higher level of satisfaction) c. 1 – 5 d. mean	N/A	0.70 to 0.99; (avg. alpha 0.95)

Instrument (‡ denotes exact name is not clear)	Author (*example study, if multiple reports) Original Citation Subscales	Purpose	Study Description	Validation Sample	a. No. of Items; b. Response scales c. Score range; d. Scoring	Validity	Reliability ( $\alpha$ or $r$ )
21. Assisted Living Resident Satisfaction Scale  ALRSS	Edelman et al. (2006) <u>12 Subscales:</u> <ul style="list-style-type: none"> <li>• Safety/Peace of Mind (4 items)</li> <li>• Personal Attention (3 items)</li> <li>• General Satisfaction (4 items)</li> <li>• Staff (5 items)</li> <li>• Residents (5 items)</li> <li>• Knowledge (4 items)</li> <li>• Autonomy (4 items)</li> <li>• Aides (3 items)</li> <li>• Socialization With Family (2 items)</li> <li>• Transportation (2 items)</li> <li>• Privacy (2 items)</li> <li>• Activities (2 items)</li> </ul>	To examine the resident satisfaction with various aspects of assisted living	U.S.-based study. Data collected via mailed surveys (N=436) sent to 204 residents and 232 family members associated with 11 assisted living facilities in Illinois and Indiana. The root mean square error of approximation (RMSEA), the goodness-of-fit index (GFI), the comparative fit index (CFI), and the Tucker – Lewis coefficient (TLC) were used to test the fit of a 12-factor model of resident satisfaction.	Age (M): 86 years Gender: 77% female Race/ethnicity: 100% Non-Hispanic White Income: N/A Other: 72% widowed; 62% $\leq$ high school graduate.	a. 40 items b. 4-point scale -- <i>strongly agree</i> (1) to <i>strongly disagree</i> (4) c. 1 – 4 d. mean	Convergent validity and construct validity were examined.	Median $\alpha$ = 0.74, $\alpha$ of subscales range from 0.61 to 0.78.
22. Assisted Living Family Member Satisfaction Scale  ALFMSS	Edelman et al. (2006) <u>6 Subscales:</u> <ul style="list-style-type: none"> <li>• Staff Responsiveness (3 items)</li> <li>• Safety (5 items)</li> <li>• Transportation (3 items)</li> <li>• Activities (5 items)</li> <li>• Family Member Impact (5 items)</li> <li>• Resident Responsibilities (4 items)</li> </ul>	To examine the family members' satisfaction with various aspects of assisted living facilities and services.	U.S.-based study. Data collected via mailed surveys (N=436) sent to 204 residents and 232 family members associated with 11 assisted living facilities in Illinois and Indiana. The root mean square error of approximation (RMSEA), the goodness-of-fit index (GFI), the comparative fit index (CFI), and the Tucker – Lewis coefficient (TLC) were used to test the fit of a 12-factor model of resident satisfaction.	Age (M): 60 Gender: 61% female Race/ethnicity: 99% Non-Hispanic White Income: N/A Other: 12% widowed; 75% married; 61% had some college or more.	a. 25 items b. N/A c. N/A d. N/A	N/A	Median $\alpha$ = 0.86, $\alpha$ of subscales ranging from 0.60 to 0.83.
23. Family Satisfaction Instrument (Final Version-Section A-Pretest Vision)	Ejaz et al. (2003) <u>13 Service Subscales:</u> <ul style="list-style-type: none"> <li>• Admissions</li> </ul>	To develop a reliable and valid instrument to measure family satisfaction	U.S.-based study. Data were collected from both in-person surveys and mailed surveys (N=239). Respond-	<u>Family member:</u> Age (M): 61 years Gender: 61% female Race/ethnicity: 86% Cauca-	a. 60 -items b. Varied ("yes, definitely," "yes, I think so," "no, I	N/A	With the exception of the Choice subscale

Instrument (‡ denotes exact name is not clear)	Author (*example study, if multiple reports) Original Citation Subscales	Purpose	Study Description	Validation Sample	a. No. of Items; b. Response scales c. Score range; d. Scoring	Validity	Reliability ( $\alpha$ or $r$ )
	<ul style="list-style-type: none"> <li>• Social Services</li> <li>• Activities</li> <li>• Choice</li> <li>• Receptionist and Phone</li> <li>• Direct Care and Nurse Aides</li> <li>• Professional Nurses</li> <li>• Therapy</li> <li>• Administration</li> <li>• Meals &amp; Dining;</li> <li>• Laundry</li> <li>• Environment</li> <li>• Gen. Questions</li> </ul>	tion with care provided to family members living in a nursing home	ents were family members of nursing home residents in Ohio. Factor analyses were used to identify factors that might explain variation in satisfaction across the 13 service element subscales.	sian and 12% African American. Income: N/A Other: 36% receiving both Medicare and Medicaid, 19% receiving Medicaid only, 16% receiving Medicare only, 17% on private pay; 49% had some college or more.	<i>don't think so," "no, definitely not," and "don't know" – "not familiar with service."</i> c. Varied d. N/A		(0.66), all subscales had high alpha coefficients (0.78 and above) Test-retest reliability ranged from 0.49 to 0.88.
24. Three Satisfaction Scales - Swedish to English ‡	Friman (2004) Västfjäll et al. (2002) <u>Subscales:</u> <ul style="list-style-type: none"> <li>• Unpleasantness/ Pleasantness,</li> <li>• Deactivation/ Activation.</li> </ul>	To examine imagined affective reactions to public transportation.	Sweden-based study. Data were collected via face-to-face interviews ( $N = 41$ ) with students attending Karlstad University in Sweden.	Age ( $M$ ): 22 years Gender: 63% female Race/ethnicity: N/A Income: N/A	a. 12 items b. Bipolar c. 10 – 90 d. mean	N/A	0.71 to 0.93
25. Counseling Evaluation Inventory-(Client Satisfaction Subscale Version)  CEI	Fuertes et al. (2006) Linden, Stone, & Shertzer (1965) Subscales: Not described	To examine client satisfaction with therapy	U.S.-based study. Data collected via face-to-face interviews with 51 client – therapist dyads at three university therapy centers in an urban Northeast city.	<u>Clients:</u> Age ( $M$ ):27 years Gender:71% female Race/ethnicity: 24% Euro Americans; 33% Asian Americans; 27% African Americans; 16% Hispanic Americans Income: N/A	a. 5 items b. 6 – point scale <i>strongly disagree</i> (1) to <i>strongly agree</i> (6) c. 5 – 30 d. summed	The validity was reported as confirmed through factor analyses in prior studies. (cited from Pontorotto & Furlong, 1985)	0.95 Test-retest reliability of the CEI $r = 0.74$
26. Therapist Satisfaction Survey (1-Item Scale)	Fuertes et al. (2006) Subscales: Not described	To estimate therapist satisfaction	U.S.-based study. Data were collected from face-to-face interviews ( $N=51$ ) with clients and their therapists at three university therapy centers in an urban Northeast area in the U.S.	<u>Therapists:</u> Age ( $M$ ): 32 years Gender: 53% female Race/ethnicity: 67% Euro Americans; 24% Asian Americans; 8% African Americans; 2% Hispanic Americans Income: N/A	a. 1 item b. 6-point scale <i>strongly disagree</i> (1) to <i>strongly agree</i> (6) c. 1 – 6 d. summed	N/A	With one item, reliability cannot be calculated.
27. Program Satisfaction Questionnaire	Gao, Luo, & Chan (2012)	To examine program processes	China- based study. Data were collected from face-	Age ( $M$ ): 29 years Race/ethnicity: 100% Non-	a. 11 items b. 4- point scale --	N/A	N/A

Instrument (‡ denotes exact name is not clear)	Author (*example study, if multiple reports) Original Citation Subscales	Purpose	Study Description	Validation Sample	a. No. of Items; b. Response scales c. Score range; d. Scoring	Validity	Reliability ( $\alpha$ or $r$ )
	Lee & Holroyd (2009) <b>4 Subscales:</b> • Motivation to Attend Program • Positive Feedback • Helpful Aspects of Program • Improvement		to-face interviews ( $N=92$ ) with first-time mothers in a regional teaching hospital in China.	Hispanic White Income: 65% $>¥3000$ (monthly); 35% $<¥3000$ Other: 86% had a college degree or more.	<i>very dissatisfied; dissatisfied; satisfied; very satisfied</i> c. N/A d. N/A		
28. Making Better Career Decisions  <b>MBCD</b>	Gati, Gadassi, & Shemesh (2006) Gati (1996) Subscales: Not described	To examine satisfaction with occupational choices	Israel-based study. Data were collected from telephone interviews ( $N=73$ ) with clients who had participated in MBCD program during 1997.	Age ( $M$ ): 28 years Gender: 64% female Race/ethnicity: N/A Income: N/A	a. 1 item b. 9-point scale -- <i>low satisfaction</i> (1), to <i>high satisfaction</i> (9) c. 1 – 9 <i>3 categories</i> -High (8 – 9 pts) -Moderate (5-7) -Low (1-4) d. summed	N/A	N/A
29. Program Satisfaction Questionnaire ‡	Heinze et al. (2010) Subscales: Not described	To examine service satisfaction.	U.S.-based study. Data were collected from face-to-face interviews ( $N=133$ ) with homeless youth and youth at risk for homelessness receiving services from 6 community agencies in a Midwest metropolitan area. Regression analyses were used to estimate the relationship between program features and satisfaction.	Age ( $M$ ): 18 years Gender: 68% female Race/ethnicity: 25% White; 62% Black; 5% Latino; 8% other. Income: N/A	a. 11 items b. N/A c. N/A d. N/A	N/A	0.88
30. Client Satisfaction: Case Management  <b>CSAT-CM</b>	Hsieh (2006) <b>4 Subscales:</b> • Assessment of Client Needs; • Development of Plan of Care • Case Manager's Knowledge of Available Services • Availability of the	To examine the relationship between the component of element-specific satisfaction and global client satisfaction.	U.S.-based study. Data were collected via face-to-face interviews ( $N=112$ ) with clients of an agency providing case management services for older adults living in a large city in the Midwest.	Age ( $M$ ): 76 years Gender: 81% female Race/ethnicity: 92% African American. Income: 90% annual income $< \$15,000$ . Other: Avg. years of education: 10 years; 96% retired.	a. 5 items b. 7-point scale -- <i>completely dissatisfied</i> (1) to <i>completely satisfied</i> (7) c. N/A d. N/A	N/A	The test – retest reliability for the CSAT-CM in was $r = 0.81$ this study.

Instrument (‡ denotes exact name is not clear)	Author (*example study, if multiple reports) Original Citation Subscales	Purpose	Study Description	Validation Sample	a. No. of Items; b. Response scales c. Score range; d. Scoring	Validity	Reliability ( $\alpha$ or $r$ )
31. Service Element Satisfaction Questionnaires	Case Manager Hsieh (2012) <b>5 Subscales:</b> • Assessment of Client Needs; • Plan of Care Development, • Case Manager's Knowledge of Available Service • Case Manager's Ability to Obtain Services for Clients; • Availability of the Case Manager (see Hsieh, 2006 for details).	To examine the relationship between element-specific satisfaction and global satisfaction.	U.S.-based study. Data were collected via face-to-face interviews ( $N=112$ ) with clients of an agency providing case management services for older adults living in a large city in the Midwest.	Age ( $M$ ): 76 years Gender: 81% female Race/ethnicity: 92% African American. Income: 90% annual income < \$15,000. Other: Avg. years of education: 10 years; 96% retired..	a. 5 items b. 7-point scale <i>completely dissatisfied</i> (1) to <i>completely satisfied</i> (7)) c. 1 – 7 d. mean	N/A	Test-retest reliability of $r = 0.81$ .
32. Investigation Satisfaction Scale - Caregivers  ISS	Jones et al. (2007). <b>2 Subscales:</b> • Investigator Response (IR: 9 items) • Interview Experience (IE: 5 items)	To examine satisfaction with how the investigation was conducted and caregivers' satisfaction with how investigators interacted with the child.	U.S.-based, multi-site study. Data were collected from face-to-face interviews ( $N=229$ ) with caregivers of Children's Advocacy Center (CAC) cases at four sites across the U.S. Hierarchical regression analyses were used to examine the differences between CAC and comparison groups on satisfaction.	Age ( $M$ ): N/A Gender: 79% mothers; 6% fathers; 7% female relatives; 3% foster mothers. Race/ethnicity: N/A Income: N/A	a. 14 items b. 4-point scale -- <i>lowest satisfaction</i> (1) to <i>greatest satisfaction</i> (4) c. 1 – 4 d. mean	N/A	0.89 for IR subscale  0.81 for IE subscale
33. Children's Satisfaction Survey-Children	Jones et al. (2007) Subscales: Not described	To examine children's satisfaction with maltreatment investigation	U.S.-based, multi-site study. Data were collected from face-to-face interviews ( $N=90$ ) with youth involved in Children's Advocacy Center (CAC) cases at four U.S. Hierarchical regression analyses were used to examine the differences between CAC and comparison groups on satisfaction.	<u>CAC:</u> Age ( $M$ ): 9 years Gender: 78% female Race/ethnicity: 54% White; 30% African American; 9% Latino; 7% others. Income: N/A Other: 41% reported sexual abuse (penetration); 20% sustained physical injury.	a. 6 items b. 4-point scale -- <i>low satisfaction</i> (1) to <i>great satisfaction</i> (4) c. 1 – 4 d. mean	N/A	Analyses including reliability and factor analytic procedures indicated little shared variance between the items; therefore, each item was analyzed separately.

Instrument (‡ denotes exact name is not clear)	Author (*example study, if multiple reports) Original Citation Subscales	Purpose	Study Description	Validation Sample	a. No. of Items; b. Response scales c. Score range; d. Scoring	Validity	Reliability ( $\alpha$ or $r$ )
34. Parent Satisfaction With Foster Care Services Scale-- Satisfaction Items (Spanish & English)  <b>PSFCSS</b>	Kapp & Vela (2004)* Kapp & Vela (2003) Harris, Poertner & Joe (2000) <b>5 Subscales:</b> • Contract Provider Worker Competency • State Worker Competency • Cultural Competency • Empowerment/ Client Rights • Agency Quality and Outcomes	To examine the parents' overall satisfaction with foster care services.	U.S.-based study. Data were collected via a telephone survey of ( $N=184$ ) parents whose children received foster care services from private contract providers in Kansas. Logistic regression was used to explore the determinants of satisfaction.	Age ( $M$ ): N/A Gender: 79% mothers; 6% fathers; 7% female relatives; 3% foster mothers. Race/ethnicity: N/A Income: N/A	a. 34 items b. 3-point scale -- <i>agree</i> (1), <i>unsure</i> (2), <i>disagree</i> (3) c. 1 – 3 d. mean	N/A	0.94
35. Treatment Satisfaction Survey  <b>TSS</b>	Kern et al. (2011)  Subscales: Not described	To examine caregiver satisfaction	U.S.-based study. Data were collected via face-to-face surveys ( $N=18$ ) with parents whose children were diagnosed with autism spectrum disorder.	N/A	a. 4 items b. 5-point scale -- <i>low satisfaction</i> (1) to <i>great satisfaction</i> (5) c. 1 – 5 d. mean	N/A	N/A
36. School Opinion Survey - Parent Form & Student Form	King & Bond (2003) Department of Education, Queensland (1996) Subscales: Not described	To evaluate satisfaction with public education services.	Australia-based study. Data were collected using a survey of 714 parents and 1,143 students at 10 schools in Queensland, Australia. Rasch analysis of Likert-type response scales was used to establish benchmark values for client satisfaction with public education.	N/A	a. 20 items b. 5-point scale -- <i>very dissatisfied</i> (1) – <i>very satisfied</i> (5)  c. 1 – 5 d. mean	N/A	N/A
37. Youth Client Satisfaction Questionnaire - Revised Version.  <b>YCSQ</b>	Kivlighan, London, & Miles (2012) Shapiro, Welker, & Jacobson (1997) Kivlighan & Tarrant (2001) revised. <b>2 Subscales:</b> • Relationship with group/therapist	To examine client satisfaction with group therapy	U.S.-based, multi-site study. Data collected via face-to-face surveys ( $N=176$ ) with youth at 4 sites of Children's Advocacy Centers (CAC) in the U.S. Hierarchical linear modeling was used to examine the relationships	Age ( $M$ ): 15 years Gender: 51% female Race/ethnicity: 48% White; 37% African American, and 15% Native Americans, Latinos, or Asian Americans. Income: N/A	a. 14 items b. 4-point scale -- <i>not at all satisfied</i> (0) to <i>a great deal</i> (3) c. 0 – 3 d. mean	Authors report that the validity of YCSQ was established in prior work by Shapiro et al. (1997).	0.85 for original instrument: Revised version: 0.83 for the relationship with the group and 0.84 for

Instrument (‡ denotes exact name is not clear)	Author (*example study, if multiple reports) Original Citation Subscales	Purpose	Study Description	Validation Sample	a. No. of Items; b. Response scales c. Score range; d. Scoring	Validity	Reliability ( $\alpha$ or $r$ )
	• Benefits of therapy		between group leadership structure, group size, and two group satisfaction variables.				the benefits of therapy subscales.
38. Counselor Rating Form-Short  <b>CRF-S</b>	Lawson & Brossart (2003) Corrigan & Schmidt (1983) <b>3 Subscales:</b> • Attractiveness • Expertness • Trustworthiness	To examine satisfaction with counseling	U.S.-based study. Data were collected from face-to-face interviews with 20 clients in a community-based counseling center in the U.S.	Age ( $M$ ): 30 years Gender: 65% female Race/ethnicity: 5% Hispanic/Latino; 5% Asian American, 80% Caucasian; and 5% African Americans; 5% unspecified. Income: N/A	a. 12 items b. 7-point scale -- not very (1) to very (7) c. 12 – 84 d. summed	N/A	0.88
39. Parent Satisfaction Questionnaire	Lees & Ronan (2008) Webster-Stratton (1999) Subscales: Not described	To examine the satisfaction of parents of children with ADHD diagnoses who were involved in a public clinic program	New Zealand-based study. Data collected via face-to-face interviews with 4 high-risk single mothers whose children were diagnosed with attention deficit hyperactivity disorder (ADHD). Study site was a public clinic setting.	N/A	a. 11 items b. 7-point scale -- <i>least satisfied</i> (1) to <i>most satisfied</i> (7). c. 1 – 7 d. mean	N/A	N/A
40. Sex Offender Client Treatment Satisfaction Survey ‡	Levenson et al. (2009)* Levenson, Prescott & D'Amora (2010) (Drawn from: Garrett et al. (2003) <b>7 Subscales (# items):</b> • Treatment Content (15); • Group Process (5); • Group Therapy (7) • Individual Therapy (4) • Group Therapists (9) • Program Policy (7) • Client Overall Satisfaction With Program (5 items)	To investigate the satisfaction of participants in a sex offender treatment program	U.S.-based, multi-site study. Data were collected via face-to-face interviews ( $N=228$ ) with male sex offenders from 3 outpatient sex offender counseling centers in Florida and Minnesota.	Age: 80% 26 – 64 years Gender: 100% male Race/ethnicity: 79% White; 21% minority race Income: 57% <\$ 30,000 per year Others: 33% never been married; 28% divorced. 59% completed high school or more.	a. 52 items b. Varied (3- or 5-point Likert scales) c. varied d. summed	N/A	Treatment Content Subscale: 0.89 Group Process: 0.81 Group Therapy: 0.70 Individual Therapy: 0.40 Group Therapists: 0.91 Program Policy: 0.82 Overall Program Satisfaction: 0.85.
41. Satisfaction With End-Of-Life Care In Dementia Scale	Liu, Guarino, & Lopez (2012) Volicer, Hurley, & Blasi (2001)	To examine family members overall satisfaction with end-of-life care	U.S.-based study. Data were collected via a mailed survey ( $N=239$ ) sent to family members of nursing	Age ( $M$ ): 65 years Gender: 71% female Race/ethnicity: 98% Caucasian; 2% others	a. 10 items b. 4-point scale <i>strongly disagree</i> (1) to <i>strongly</i>	Authors indicate that convergent validity was demonstrated in	0.90



Instrument (‡ denotes exact name is not clear)	Author (*example study, if multiple reports) Original Citation Subscales	Purpose	Study Description	Validation Sample	a. No. of Items; b. Response scales c. Score range; d. Scoring	Validity	Reliability ( $\alpha$ or $r$ )
	Subscales: Not described	services	home residents who died with dementia in Massachusetts and Rhode Island.	Income: N/A; Other: 13% widowed; 65% married; 72% had some college or more.	<i>agree</i> (4) c. 10 – 40 d. summed	previous reports.	
42. Multimodality Quality Assurance Instrument  <b>MQA</b>	Melnick, Hawke, & Wexler (2004) Melnick & Pearson (2000) Subscales: Not described	To examine treatment satisfaction	U.S.-based, multi-site study. Data collected via face-to-face interview ( $N=1,059$ ) with participants and staff in 13 prison-based drug treatment programs across the U.S. Exploratory factor analyses (EFAs) were used to determine the optimal way to combine items into conceptually based scales.	Age: 40% 30 – 40 years Gender: N/A Race/ethnicity: 67% minorities Income: N/A Other: marijuana (32%), alcohol (28%), and cocaine/crack (23%) were the most prevalent drugs of abuse.	a. 12 items b. 5-point scale -- <i>very dissatisfied</i> (1) to <i>very satisfied</i> (5) c. N/A d. N/A	N/A	0.88
43. Parent Satisfaction Survey- With Head Start Version‡	Mendez (2010) Instrument items derived from a 2003 national survey of family members of Head Start children, called the Family and Children's Experiences (FACES) study. <b>2 Subscales:</b> • Child (2 items); • Family (2 items)	To examine the satisfaction with Head Start services and with parenting interventions.	U.S.-based study. Data collected via face-to-face interview ( $N=288$ : 177 intervention and 111 control) of families from 3 cohorts in 4 Head Start programs serving African American communities in a Southern U.S. city.	Age ( $M$ ): N/A Gender: 97% female Race/ethnicity: 94% African American. Income: N/A Other: 94% biological mother; 3% adoptive mother; 3% fathers; 67% single; 48% full-time employed. 40% high school diploma, 34% had some college experience.	a. 4 items b. 4-point scale -- <i>low satisfaction</i> (1) to <i>high satisfaction</i> (4) c. 1 – 4 d. mean	N/A	N/A
44. Client Satisfaction Questionnaire	Miller (2008)  Miller & Slive (2004)  Subscales: Not described	To examine client satisfaction with a walk-in single session therapy service	Canada-based study. Data collected via an in-person survey either immediately following a walk-in session or via mailed survey within 1 week of the session. Respondents ( $N=403$ ) were clients and therapists involved in services at Eastside Family Center Advisory Counsel in Calgary, Alberta.	Age ( $M$ ): N/A  Gender: 56% female  Race/ethnicity: 86% Caucasian; 14% Asian, Japanese, Chinese, or Native American  Income: Avg. monthly income = \$ 1,400.	a. 5 items b. 5-point scale -- <i>very dissatisfied</i> (1) to <i>very satisfied</i> (5)  c. 1 – 5 d. mean	N/A	N/A
45. Client Satisfaction Survey	Murphy et al. (2009)  <b>2 Subscales:</b>	To examine the client satisfaction in a counseling pro-	Canada-based study. Data were collected from online survey ( $N=45$ ) and face-to-face interview ( $N=43$ )	Age ( $M$ ): 42 yrs. online group; 44 yrs. face-to-face group;  Gender: 73% female for	a. 13 items b. 5-point scale -- <i>strongly disagree</i>	N/A	N/A

Instrument (‡ denotes exact name is not clear)	Author (*example study, if multiple reports) Original Citation Subscales	Purpose	Study Description	Validation Sample	a. No. of Items; b. Response scales c. Score range; d. Scoring	Validity	Reliability ( $\alpha$ or $r$ )
<b>CSS</b>	<ul style="list-style-type: none"> <li>Quality of Your Experience (8 items)</li> <li>Program Effectiveness (5 items)</li> </ul>	gram	surveys with clients who received either face to face or online counseling in Canada. ANOVA was used to examine the differences between counseling modalities.	online; 76% for face-to-face;  Race/ethnicity: N/A  Income: N/A	(1) to <i>strongly agree</i> (5)  c. 1 – 5  d. mean		
46. Consumer Reports Effectiveness Score - 4 Items - Satisfaction <b>CRES-4</b>	Nielsen et al. (2004) Seligman (1995) Freedman et al. (1999) Subscales: Not described	To examine satisfaction in a counseling center	U.S.-based study. Data collected via mailed survey (N=302) sent to former clients of the Counseling & Career Center, Brigham Young University in Provo, Utah.	Age (M): 24 years Gender: 71% female Race/ethnicity: 94% White; Income: N/A Other: 33% married	a. 3 items b. Varied (N/A) c. 0 – 300 d. composite	Authors claimed that the CRES-4 has well-established validity.	0.63
47. Participants' Satisfaction With Intervention ‡	Schiff, Witte, & El-Bassel (2003) Subscales: Not described	To examine satisfaction with an HIV/STD intervention.	U.S.-based study. Data collected via mailed survey (N=107) of participants who attended the HIV/STD relationship-based preventive intervention in a primary health care setting in a low-income, inner-city neighborhood in Bronx, New York City.	Age (M): 37 years Gender: 68% female Race/ethnicity: 52% African American; 41% Hispanic. Income: 64% < \$5,000 per year Other: 68% had less than a high school education; 61% single; 28% HIV positive.	a. 3 items b. 5-point scale -- <i>not at all satisfied/ honest</i> (1) to <i>very satisfied/ honest</i> (5) c. 1 – 5 d. mean	N/A	0.77
48. Client Satisfaction Survey- English & Spanish versions ‡	Schraufnagel & Li (2010). Subscales: Not described	To examine the satisfaction with mediation services.	U.S.-based study. Data collected via random sample face-to-face survey (N=65) of clients who went through a mediation process or experienced the court process in Florida. Sensitivity analysis was used to test the explanations for clients' satisfaction.	N/A	a. 3 items b. 5-point scale -- <i>very dissatisfied</i> (1) to <i>very satisfied</i> (5) c. 0.2-1.0 (3/15-15/15) d. Sum	N/A	N/A
49. Chinese Subjective Outcome Scale -20 Items <b>CSOS</b>	Shek (2010). Shek et al. (2007) Subscales: Not described	To assess the relationship between satisfaction and program processes and outcomes	China-based study. Data collected from a random sample survey (N=3,298) of students from 22 schools in Hong Kong, China. A non-orthogonal factor extraction procedure (alpha factoring)	N/A	a. 20 items b. N/A c. N/A d. N/A	N/A	0.97

Instrument (‡ denotes exact name is not clear)	Author (*example study, if multiple reports) Original Citation Subscales	Purpose	Study Description	Validation Sample	a. No. of Items; b. Response scales c. Score range; d. Scoring	Validity	Reliability ( $\alpha$ or $r$ )
			was used to analyze the responses of the students to the scale items. Pearson correlation was used to test the relationship between objective and subjective outcomes and perceived program effectiveness.				
50. Resident Satisfaction Index -- Short Version  <b>RSI</b>	Sikorska-Simmons (2006) Sikorska-Simmons (2001) Subscales: Not described	To examine resident satisfaction in assisted living facilities	U.S.-based study. Data were collected from face-to-face interviews with residents ( $N=335$ ) and staff (298) in 43 assisted living facilities.	Age ( $M$ ): 83 years Gender: 74% female Race/ethnicity: 94% White Income: N/A Other: 70% widowed; 24% had some college or more.	a. 6 items b. 4-point scale -- <i>never</i> (1) to <i>always</i> (4) c. 6 – 24 d. summed	N/A	0.67
51. Overall Job Satisfaction Scale	Sikorska-Simmons (2006) Seashore et al. (1982) Subscales: Not described	To examine staff job satisfaction with the work environment.	U.S.-based study. Data collected via face-to-face interviews with residents ( $n=335$ ) and staff ( $n=298$ ) in 43 assisted living facilities in the U.S.	Age ( $M$ ): 43 years Gender: 91% female Race/ethnicity: 57% White; 33% Black Income: N/A Other: 68% not married; 52% had some college or more.	a. 3 items b. 7 point scale -- <i>strongly disagree</i> (1) to <i>strongly agree</i> (7) c. 3 – 21 d. summed	The authors indicate that the validity of this scale was established in prior studies (cited Fields, 2002).	0.79
52. Clients' Overall Satisfaction Survey	Smith, Thomas, & Jackson (2004) Jackson et al. (2000a) Subscales: Not described	To examine satisfaction with the problem gambling counseling service	Australia-based study. Data collected via face-to-face interview survey ( $N=150$ ) with clients who undertook short-term counseling in the government funded <i>Gambler's Help</i> Problem-Gambling Counseling Services in Victoria, Australia.	N/A	a. 10 items b. 4-point scale -- <i>dissatisfied; neither; satisfied; very satisfied</i> c. N/A d. N/A	N/A	N/A
53. General Satisfaction Survey (Hebrew & English)	Spiro, Dekel, & Peled (2009) From CSQ (Larsen et al., 1979), the YSQ (Stuntzner-Gibson et al., 1995), and the Youth Client Satisfaction Questionnaire (YCSQ) (Shapiro et al., 1997) Subscales: Not described	To examine satisfaction at Makom Acher (MA).	Israel-based study. Data collected via telephone survey of ( $N=102$ ) adolescents and young adults who left a youth shelter in Tel Aviv. Pearson's product-moment correlation was used to explore the relationships among various client satisfaction scales. A multiple regression analysis was used to develop a model predicting general satisfaction with 7 aspects of MA.	Age: 12% 13 – 14 years 40% 15 – 16 years 32% 17 years 16% 18 – 20 years  Gender: 53% female Race/ethnicity: N/A Income: N/A Other: After leaving shelter, 69% returned to their families, 13% were placed in group or foster care, and 18% departed	a. 1 b. 4-point scale -- (very good, good, about average, and not so good.) c. 1-4 d. sum	N/A	N/A

Instrument (‡ denotes exact name is not clear)	Author (*example study, if multiple reports) Original Citation Subscales	Purpose	Study Description	Validation Sample	a. No. of Items; b. Response scales c. Score range; d. Scoring	Validity	Reliability ( $\alpha$ or $r$ )
				to an independent, non-normative or unknown living arrangement.			
54. Satisfaction With Specific Aspects of Life at Makom Acher Youth Shelter	Spiro et al. (2009) From CSQ (Larsen et al., 1979), the YSQ (Stuntzner-Gibson et al., 1995), and the Youth Client Satisfaction Questionnaire (YCSQ) (Shapiro et al., 1997) Subscales: Not described	To examine satisfaction of 7 aspects of life in the Makom Acher Youth Shelter.	Israel-based study. Data were collected from a telephone survey of ( $N=102$ ) adolescents and young adults who left a youth shelter in Tel Aviv, Israel. Pearson's product-moment correlation was used to explore the relationships among the various client satisfaction scales. A multiple regression analysis was used to develop a model predicting general satisfaction with 7 aspects of MA.	Age: 12% 13 – 14 years 40% 15 – 16 years 32% 17 years 16% 18 – 20 years Gender: 53% female Race/ethnicity: N/A Income: N/A Other: After leaving from the shelter, 69% went to their families, 13% were placed in group or foster care, and 18% departed to an independent, non-normative or unknown living arrangement.	a. 7 items b. 4-point scale <i>very good, good, about average, and not so good</i> c. 1 – 4 d. mean	N/A	N/A
55. Post-Program Satisfaction Questionnaire	Strand & Badger (2005) Subscales: Not described	To examine satisfaction with child welfare program	U.S.-based study. Data were collected from face to face interviews with ( $N=158$ ) supervisors in child welfare agencies in New York City.	N/A	a. 20 items total; 10 items w/ 3-point Likert scale; 10 use other scoring b. 3-point scale <i>not really (1), somewhat (2), a lot (3)</i> c. 10 – 30 d. summed	N/A	0.79
56. Client Satisfaction Measures ‡	Trotter (2008) Dufour & Chamberland (2004) Subscales: Not described	To examine satisfaction with child protective services	Australia-based study. Data were collected from face to face interviews of ( $N=247$ ) family members who received child protection services in Victoria, Australia.	N/A	a. 3 items b. 7-point scale -- <i>very poor progress (1) to very good progress (7)</i> c. N/A d. N/A	Author indicates that prior reports have established face validity (Dufour & Chamberland, 2004).	N/A
57. Student Satisfaction Survey	Westbrook et al. (2012) Subscales: Not described	To assess satisfaction with CBT training	U.K.-based study. Data were collected from an online survey of ( $N=94$ ) students who participated in an online cognitive behavior therapy (CBT) at Oxford Cognitive Therapy Centre in the U.K.	N/A	a. 5 items b. 11-point scale (0-10) c. 0 – 10 d. mean	N/A	N/A

<b>Instrument</b> (‡ denotes exact name is not clear)	<b>Author</b> (*example study, if multiple reports) <b>Original Citation</b> <b>Subscales</b>	<b>Purpose</b>	<b>Study Description</b>	<b>Validation Sample</b>	<b>a. No. of Items; b. Response scales c. Score range; d. Scoring</b>	<b>Validity</b>	<b>Reliability (<math>\alpha</math> or <math>r</math>)</b>
58. Consultation Evaluation Form  <b>CEF</b>	Wilkinson (2005) Erchul (1987) Subscales: Not described	To assess satisfaction with conjoint behavioral consultation services	U.S.-based study. Data were collected from a case study by using face-to face parent and teacher interviews.	<b>Case study:</b> A 9-year- old Caucasian boy diagnosed with Asperger syndrome and attention-deficit/ hyperactivity disorder.	a. 12 items b. 7 – point scale -- <i>strongly disagree</i> (1) to <i>strongly agree</i> (7) c. 1 – 7 d. mean	N/A	0.94.(Cited from previous work by Wilkinson, 2003)

## References - Appendix A

- Attkisson, C. C., & Zwick, R. (1982). The Client Satisfaction Questionnaire: Psychometric properties and correlations with service utilization and psychotherapy outcome. *Evaluation and Program Planning*, 5(3), 233-237. doi: 10.1016/0149-7189(82)90074-X
- Beavers, K. F., Kratochwill, T. R., & Braden, J. P. (2004). Treatment utility of functional versus empiric assessment within consultation for reading problems. *School Psychology Quarterly*, 19(1), 29-49. doi:10.1521/scpq.19.1.29.29405
- Bonach, K., Mabry, J. B., & Potts-Henry, C. (2010). Exploring nonoffending caregiver satisfaction with a children's advocacy center. *Journal of Child Sexual Abuse*, 19(6), 687-708. doi: 10.1080/10538712.2010.522495
- Boyle, C. L., Sanders, M. R., Lutzker, J. R., Prinz, R. J., Shapiro, C., & Whitaker, D. J. (2010). An analysis of training, generalization, and maintenance effects of primary care Triple P for parents of preschool-aged children with disruptive behavior. *Child Psychiatry and Human Development*, 41(1), 114-131. doi: 10.1007/s10578-009-0156-7
- Bradshaw, W., & Umbreit, M. (2003). Assessing satisfaction with victim services: The development and use of the victim satisfaction with offender dialogue scale (VSODS). *International Review of Victimology*, 10(1), 71-83. doi: 10.1177/026975800301000104
- Brenninkmeijer, V., & Blonk, R. W. (2012). The effectiveness of the JOBS program among the long-term unemployed: A randomized experiment in the Netherlands. *Health Promotion International*, 27(2), 220-229. doi: 10.1093/heapro/dar033
- Brooks, F., & Brown, E. (2005). A program evaluation of Los Angeles ACORN's welfare case advocacy. *Journal of Human Behavior in the Social Environment*, 12(2-3), 185-203. doi: 10.1300/J137v12n02\_10
- Butler, S. S., Gomon, J., & Turner, W. (2004). Satisfaction and well-being among residents of a low-income, rural assisted living facility and the implications for social work practice. *Journal of Social Work in Long-Term Care*, 3(1), 61-84. doi: 10.1300/J181v03n01\_06
- Castle, N. G., Lowe, T. J., Lucas, J. A., Robinson, J. P., & Crystal, S. (2004). Use of resident satisfaction surveys in New Jersey nursing homes and assisted living facilities. *Journal of Applied Gerontology*, 23(2), 156-171. doi: 10.1177/0733464804265584
- Charbonneau, É., & Van Ryzin, G. G. (2012). Performance measures and parental satisfaction with New York City schools. *American Review of Public Administration*, 42(1), 54-65. doi: 10.1177/0275074010397333
- Collins, M., Curley, A., Clay, C., & Lara, R. (2005). Evaluation of social services in a HOPE VI housing development: Resident and staff perceptions of successes and barriers. *Evaluation and Program Planning*, 28(1), 47-59. doi:10.1016/j.evalprogplan.2004.10.004
- Coloma, J., Gibson, C., & Packard, T. (2012). Participant outcomes of a leadership development initiative in eight human service organizations. *Administration in Social Work*, 36(1), 4-22. doi: 10.1080/03643107.2011.614540
- Corrigan, J. D., & Schmidt, L. D. (1983). Development and validation of revisions in the counselor rating form. *Journal of Counseling Psychology*, 30(1), 64. doi: 10.1037/0022-0167.30.1.64
- Cowan, R. J., & Sheridan, S. M. (2003). Investigating the acceptability of behavioral interventions in applied conjoint behavioral consultation: Moving from analog conditions to naturalistic settings. *School Psychology Quarterly*, 18(1), 1-21. doi:10.1521/scpq.18.1.1.20877

- Cross, T. P., Jones, L. M., Walsh, W. A., Simone, M., Kolko, D. J., Szczepanski, J., et al. (2008). *Evaluating children's advocacy centers' response to child sexual abuse*. Retrieved from <http://www.ncjrs.gov/pdffiles1/ojdp/218530.pdf>
- Dauenhauer, J. A., Cattat Mayer, K., & Mason, A. (2007). Evaluation of adult protective services: Perspectives of community professionals. *Journal of Elder Abuse & Neglect*, 19(3-4), 41-57. doi:10.1300/J084v19n03\_03
- Dearing, R. L., Barrick, C., Dermen, K. H., & Walitzer, K. S. (2005). Indicators of client engagement: Influences on alcohol treatment satisfaction and outcomes. *Psychology of Addictive Behaviors*, 19(1), 71-78. doi:10.1037/0893-164X.19.1.71
- Denton, W. H., Nakonezny, P. A., & Burwell, S. R. (2011). The effects of meeting a family therapy supervision team on client satisfaction in an initial session. *Journal of Family Therapy*, 33(1), 85-97. doi: 10.1111/j.1467-6427.2010.00518.x
- Department of Education, Queensland. (1996). *Partners for excellence: The strategic plan 1997-2001 for the Department of Education*. Brisbane, AU: Author.
- Dufour, S., & Chamberland, C. (2004). The effectiveness of selected interventions for previous maltreatment: Enhancing the well-being of children who live at home. *Child & Family Social Work*, 9(1), 39-56. doi: 10.1111/j.1365-2206.2004.00302.x
- Dumas, J. E., Arriaga, X. B., Begle, A. M., & Longoria, Z. N. (2011). Child and parental outcomes of a group parenting intervention for Latino families: A pilot study of the CANNE program. *Cultural Diversity and Ethnic Minority Psychology*, 17(1), 107-115. doi: 10.1037/a0021972
- Edelman, P., Guihan, M., Bryant, F. B., & Munroe, D. J. (2006). Measuring resident and family member determinants of satisfaction with assisted living. *Gerontologist*, 46(5), 599-608. doi: 10.1093/geront/46.5.599
- Ejaz, F. K., Straker, J. K., Fox, K., & Swami, S. (2003). Developing a satisfaction survey for families of Ohio's nursing home residents. *Gerontologist*, 43(4), 447-458. doi: 10.1093/geront/43.4.447
- Erchul, W. P. (1987). A relational communication analysis of control in school consultation. *Professional School Psychology*, 2(2), 113-124. doi:10.1037/h0090534
- Eyberg, S. M. (1993). Consumer satisfaction measures for assessing parent training programs. In L. Van de Creek, S. Knapp, & T. L. Jackson (Eds.), *Innovations in clinical practice: A source book* (Vol. 12; pp. 377-382). Sarasota, FL: Professional Resource Press.
- Freedman, N., Hoffenberg, J. D., Vorus, N., & Frosch, A. (1999). The effectiveness of psychoanalytic psychotherapy: The role of treatment duration, frequency of sessions, and the therapeutic relationship. *Journal of the American Psychoanalytic Association*, 47(3), 741-772. doi:10.1177/00030651990470031001
- Friman, M. (2004). The structure of affective reactions to critical incidents. *Journal of Economic Psychology*, 25(3), 331-353. doi:10.1016/S0167-4870(03)00012-6
- Fuertes, J. N., Stracuzzi, T. I., Bennett, J., Scheinholtz, J., Mislouack, A., Hersh, M., & Cheng, D. (2006). Therapist multicultural competency: A study of therapy dyads. *Psychotherapy: Theory, Research, Practice, Training*, 43(4), 480-490. doi:10.1037/0033-3204.43.4.480
- Gao, L. L., Luo, S. Y., & Chan, S. W. (2012). Interpersonal psychotherapy-oriented program for Chinese pregnant women: Delivery, content, and personal impact. *Nursing & Health Sciences*, 14(3), 318-324. doi:10.1111/j.1442-2018.2012.00722.x

- Garrett, T., Oliver, C., Wilcox, D. T., & Middleton, D. (2003). Who cares? The views of sexual offenders about the group treatment they receive. *Sexual Abuse: A Journal of Research and Treatment*, 15(4), 323-338.
- Gati, I. (1996). Computer-assisted career counseling: Challenges and prospects. In M. L. Savickas & B. W. Walsh (Eds.), *Handbook of career theory and practice* (pp. 169-190). Palo Alto, CA: Davies-Black.
- Gati, I., Gadassi, R., & Shemesh, N. (2006). The predictive validity of a computer-assisted career decision-making system: A six-year follow-up. *Journal of Vocational Behavior*, 68(2), 205-219. doi:10.1016/j.jvb.2005.08.002
- Harris, G., Poertner, J., & Joe, S. (2000). The parents with children in foster care satisfaction scale. *Administration in Social Work*, 24(2), 15-27. doi:10.1300/J147v24n02\_02
- Heinze, H. J., Hernandez Jozefowicz, D. M., & Toro, P. A. (2010). Taking the youth perspective: Assessment of program characteristics that promote positive development in homeless and at-risk youth. *Children and Youth Services Review*, 32(10), 1365-1372. doi:10.1016/j.childyouth.2010.06.004
- Horvath, A. O. (1994). Research on the alliance. In A.O. Horvath & L.S. Greenberg (Eds.), *The working alliance: Theory, research, and practice* (pp. 259-286). New York, NY: John Wiley & Sons.
- Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*, 36(2), 223-233. doi:10.1037/0022-0167.36.2.223
- Hsieh, C. (2006). Using client satisfaction to improve case management services for the elderly. *Research on Social Work Practice*, 16(6), 605-612. doi:10.1177/1087057106289360
- Hsieh, C. (2012). Incorporating perceived importance of service elements into client satisfaction measures. *Research on Social Work Practice*, 22(1), 93-99. doi:10.1177/1049731511416826
- Jackson, A. C., Thomas, S. A., Thomason, N., Borrell, J., Crisp, B. R., Ho, W., ... Smith, S. (2000) *Longitudinal evaluation of the effectiveness of problem gambling counselling services, community education strategies and information products – Volume 2: Evaluation of the effectiveness of the counselling interventions at BreakEven Problem Gambling Counselling Services*. Melbourne, AU: Victoria Department of Human Services. Retrieved from [http://education.unimelb.edu.au/\\_\\_data/assets/pdf\\_file/0003/632784/pg\\_volume2.pdf](http://education.unimelb.edu.au/__data/assets/pdf_file/0003/632784/pg_volume2.pdf)
- Jones, L. M., Cross, T. P., Walsh, W. A., & Simone, M. (2007). Do children's advocacy centers improve families' experiences of child sexual abuse investigations? *Child Abuse & Neglect*, 31(10), 1069-1085. doi: 10.1016/j.chiabu.2007.07.003
- Kapp, S. A., & Vela, R. H. (2003). The Parent Satisfaction with Foster Care Services Scale. *Child Welfare*, 83(3), 263-287.
- Kapp, S. A., & Vela, R. H. (2004). The unheard client: Assessing the satisfaction of parents of children in foster care. *Child & Family Social Work*, 9(2), 197-206. doi:10.1111/j.1365-2206.2004.00323.x
- Kelley, M. L., Heffer, R. W., Gresham, F. M., & Elliott, S. N. (1989). Development of a modified treatment evaluation inventory. *Journal of Psychopathology and Behavioral Assessment*, 11(3), 235-247. doi: 10.1007/BF00960495
- Kern, J. K., Fletcher, C. L., Garver, C. R., Mehta, J. A., Grannemann, B. D., Knox, K. R., . . . Trivedi, M. H. (2011). Prospective trial of equine-assisted activities in autism spectrum disorder. *Alternative Therapies in Health and Medicine*, 17(3), 14-20.



- King, J. A., & Bond, T. G. (2003). Measuring client satisfaction with public education I: Meeting competing demands in establishing state-wide benchmarks. *Journal of Applied Measurement, 4*(2), 111-123.
- Kivlighan Jr., D. M., & Tarrant, J. M. (2001). Does group climate mediate the group leadership – group member outcome relationship?: A test of Yalom's hypotheses about leadership priorities. *Group Dynamics: Theory, Research, and Practice, 5*(3), 220. doi:10.1037/1089-2699.5.3.220
- Kivlighan Jr., D. M., London, K., & Miles, J. R. (2012). Are two heads better than one? The relationship between number of group leaders and group members, and group climate and group member benefit from therapy. *Group Dynamics: Theory, Research, and Practice, 16*(1), 1-13. doi: 10.1037/a0026242
- Larsen, D. L., Attkisson, C. C., Hargreaves, W. A., & Nguyen, T. D. (1979). Assessment of client/patient satisfaction: Development of a general scale. *Evaluation and Program Planning, 2*(3), 197-207. doi: 10.1016/0149-7189(79)90094-6
- Lawson, D. M., & Brossart, D. F. (2003). The relationship between counselor trainee family-of-origin structure and counseling effectiveness. *Clinical Supervisor, 22*(2), 21-36. doi: 10.1300/J001v22n02\_03
- Lee, L. Y., & Holroyd, E. (2009). Evaluating the effect of childbirth education class: A mixed-method study. *International Nursing Review, 56*(3), 361-368. doi:10.1111/j.1466-7657.2008.00701.x
- Lees, D. G., & Ronan, K. R. (2008). Engagement and effectiveness of parent management training (incredible years) for solo high-risk mothers: A multiple baseline evaluation. *Behaviour Change, 25*(2), 109-128. doi: 10.1375/behc.25.2.109
- Levenson, J. S., MacGowan, M. J., Morin, J. W., & Cotter, L. P. (2009). Perceptions of sex offenders about treatment: Satisfaction and engagement in group therapy. *Sexual Abuse: A Journal of Research and Treatment, 21*(1), 35-56.
- Levenson, J. S., Prescott, D. S., & D'Amora, D. A. (2010). Sex offender treatment. *International Journal of Offender Therapy and Comparative Criminology, 54*(3), 307-326. doi:10.1177/0306624X08328752
- Linden, J. D., Stone, S. C., & Shertzer, B. (1965). Development and evaluation of an inventory for rating counseling. *Personnel and Guidance Journal, 44*(3), 267-276. doi:10.1002/j.2164-4918.1965.tb03516.x
- Liu, L. M., Guarino, A. J., & Lopez, R. P. (2012). Family satisfaction with care provided by nurse practitioners to nursing home residents with dementia at the end of life. *Clinical Nursing Research, 21*(3), 350-367. doi:10.1177/1054773811431883
- McMurtry, S. L., & Hudson, W. W. (2000). The Client Satisfaction Inventory: Results of an initial validation study. *Research on Social Work Practice, 10*(5), 644-663.
- Melnick, G., Hawke, J., & Wexler, H. K. (2004). Client perceptions of prison-based therapeutic community drug treatment programs. *Prison Journal, 84*(1), 121-138. doi:10.1177/0032885503262459
- Melnick, G., & Pearson, F. (2000). *A multimodality quality assurance instrument*. New York, NY: National Development and Research Institutes.
- Mendez, J. L. (2010). How can parents get involved in preschool? Barriers and engagement in education by ethnic minority parents of children attending Head Start. *Cultural Diversity & Ethnic Minority Psychology, 16*(1), 26-36. doi:10.1037/a0016258
- Miller, J. K. (2008). Walk-in single session team therapy: A study of client satisfaction. *Journal of Systemic Therapies, 27*(3), 78-94. doi:10.1521/jsyt.2008.27.3.78

- Miller, J. K., & Slive, A. (2004). Breaking down the barriers to clinical service delivery: Walk-in family therapy. *Journal of Marital and Family Therapy*, 30(1), 95-103. doi:10.1111/j.1752-0606.2004.tb01225.
- Murphy, L., Parnass, P., Mitchell, D. L., Hallett, R., Cayley, P., & Seagram, S. (2009). Client satisfaction and outcome comparisons of online and face-to-face counselling methods. *British Journal of Social Work*, 39(4), 627-640. doi:10.1093/bjsw/bcp041
- New York City Department of Education. (2008). *2008 Parents survey*. Retrieved from <http://schools.nyc.gov/NR/ronlyres/2E0CEA9C-79D6-4668-B277-1F1711ED08F2/33766/Parent20survey202008.pdf>
- Nielsen, S. L., Smart, D. W., Isakson, R. L., Worthen, V. E., Gregersen, A. T., & Lambert, M. J. (2004). The consumer reports effectiveness score: What did consumers report? *Journal of Counseling Psychology*, 51(1), 25-37. doi:10.1037/0022-0167.51.1.25
- Ponterotto, J. G., & Furlong, M. J. (1985). Evaluating counselor effectiveness: A critical review of rating scale instruments. *Journal of Counseling Psychology*, 32(4), 597-616. doi:10.1037/0022-0167.32.4.597
- Schiff, M., Witte, S. S., & El-Bassel, N. (2003). Client satisfaction and perceived helping components of an HIV/AIDS preventive intervention for urban couples. *Research on Social Work Practice*, 13(4), 468-492. doi:10.1177/1049731503013004004
- Schraufnagel, S., & Li, Q. (2010). Testing the suitability of mediation of child support orders in Title IV-D cases. *Research on Social Work Practice*, 20(2), 212-222. doi:10.1177/1049731509339029
- Seashore, S. E., Lawler, E. E., Mirvis, P., & Cammann, C. (Eds.). (1982). *Observing and measuring organizational change: A guide to field practice*. New York, NY: Wiley.
- Seligman, M. E. (1996). Science as an ally of practice. *American Psychologist*, 51(10), 1072-1079. doi:10.1037/0003-066X.51.10.1072
- Shapiro, J. P., Welker, C. J., & Jacobson, B. J. (1997). The Youth Client Satisfaction Questionnaire: Development, construct validation, and factor structure. *Journal of Clinical Child Psychology*, 26(1), 87-98. doi:10.1207/s15374424jccp2601\_9
- Shek, D. T. L. (2010). Subjective outcome and objective outcome evaluation findings: Insights from a Chinese context. *Research on Social Work Practice*, 20(3), 293-301. doi:10.1177/1049731509331951
- Shek, D. T., Siu, A. M., & Lee, T. Y. (2007). The Chinese Positive Youth Development Scale: A validation study. *Research on Social Work Practice*, 17(3), 380-391. doi:10.1177/1049731506296196
- Sikorska-Simmons, E. (2001). Development of an instrument to measure resident satisfaction with assisted living. *Journal of Applied Gerontology*, 20(1), 57-73. doi:10.1177/073346480102000104
- Sikorska-Simmons, E. (2006). Linking resident satisfaction to staff perceptions of the work environment in assisted living: A multilevel analysis. *Gerontologist*, 46(5), 590-598. doi:10.1093/geront/46.5.590
- Smith, S. A., Thomas, S. A., & Jackson, A. C. (2004). An exploration of the therapeutic relationship and counselling outcomes in a problem gambling counselling service. *Journal of Social Work Practice*, 18(1), 99-112. doi:10.1080/0265053042000180581
- Spiro, S. E., Dekel, R., & Peled, E. (2009). Dimensions and correlates of client satisfaction: An evaluation of a shelter for runaway and homeless youth. *Research on Social Work Practice*, 19(2), 261-270. doi:10.1177/1049731508329395

- Sprenkle, D. H., Constantine, J. A. and Piercy, F. P. (1982) *Purdue Live Observation Satisfaction Scale*. Unpublished instrument [PMid:7067793]. Purdue University, West Lafayette, IN.
- Strand, V. C., & Badger, L. (2005). Professionalizing child welfare: An evaluation of a clinical consultation model for supervisors. *Children and Youth Services Review*, 27(8), 865-880. doi:10.1016/j.childyouth.2004.12.001
- Stuntzner-Gibson, D., Koren, P. E., & DeChillo, N. (1995). The Youth Satisfaction Questionnaire: What kids think of services. *Families in Society*, 76(10), 616-624. doi:10.1606/1044-3894.1291
- Trotter, C. (2008). What does client satisfaction tell us about effectiveness? *Child Abuse Review*, 17(4), 262-274. doi:10.1002/car.1038
- Turco, T. L., & Elliott, S. N. (1986a). Assessment of students' acceptability ratings of teacher-initiated interventions for classroom misbehavior. *Journal of School Psychology*, 24(3), 277-283. doi:10.1016/0022-4405(86)90060-9
- Turco, T. L., & Elliott, S. N. (1986b). Students' acceptability ratings of interventions for classroom misbehaviors: A study of well-behaving and misbehaving youth. *Journal of Psychoeducational Assessment*, 4(4), 281-289. doi:10.1177/073428298600400404
- Västfjäll, D., Friman, M., Gärling, T., & Kleiner, M. (2002). The measurement of core affect: A Swedish self-report measure derived from the affect circumplex. *Scandinavian Journal of Psychology*, 43(1), 19-31. doi:10.1111/1467-9450.00265
- Volicer, L., Hurley, A. C., & Blasi, Z. V. (2001). Scales for evaluation of end-of-life care in dementia. *Alzheimer Disease & Associated Disorders*, 15(4), 194-200. doi:10.1097/00002093-200110000-00005
- Von Brock, M., & Elliott, S. N. (1987). Influence and treatment effectiveness information on the acceptability of classroom interventions. *Journal of School Psychology*, 25, 131-144. doi:10.1016/0022-4405(87)90022-7
- Webster-Stratton, C. (1999). *The parent and child series handbook*. Seattle, WA: Seth Enterprises.
- Westbrook, D., McManus, F., Clark, G., & Bennett-Levy, J. (2012). Preliminary evaluation of an online training package in cognitive behaviour therapy: Satisfaction ratings and impact on knowledge and confidence. *Behavioural and Cognitive Psychotherapy*, 40(4), 481-490. doi:10.1017/S1352465811000701
- Wilkinson, L. A. (2003). Using behavioral consultation to reduce challenging behavior in the classroom. *Preventing School Failure: Alternative Education for Children and Youth*, 47(3), 100-105. doi:10.1080/10459880309604437
- Wilkinson, L. A. (2005). Supporting the inclusion of a student with Asperger Syndrome: A case study using conjoint behavioural consultation and self-management. *Educational Psychology in Practice*, 21(4), 307-326. doi:10.1080/02667360500344914