

A pilot study of the Italian adaptation of the Session Evaluation Questionnaire fourth version

Diego Rocco,¹ Silvia Salcuni,¹ Elena Antonelli²

¹Department of Developmental and Social Psychology, University of Padua, Padua; ²Department of General Psychology, University of Padua, Padua, Italy

ABSTRACT

The Session Evaluation Questionnaire (SEQ) measures the impact of counselling and psychotherapy sessions; it may be conceived as a bridge between psychotherapy process and outcome. Even if the original American SEQ has been translated into many languages, only a few translations have been validated. This is a pilot study that attempted to replicate the five-dimensional structure of the fourth version of the Anglo-American SEQ, for the Italian population. The SEQ is a self-report tool asking patients about their experience with the clinical session just ended; it consists of 27 adjectives in semantic differential scale, divided into three thematic parts: evaluation of the session itself, feelings after the session, and evaluation of the therapist. Data were collected on 111 outpatients attending the Dynamic Psychological Service for University Students, after their first two clinical interviews. Exploratory factor analyses were performed on each of the three parts of the SEQ. Results confirmed the original factorial structure, for Depth, Smoothness, Positivity and Arousal dimensions; Good Therapist dimension overlapped perfectly with the original one. The Italian SEQ showed adequate internal consistency. Convergent validity measured with an index of perceived satisfaction in the counselling process showed significant positive correlations. This pilot study showed that the Italian SEQ is a reliable instrument to measure the impact of clinical sessions. Validation studies are needed, especially to replicate the factor structure of the instrument and to better assess its validity.

Key words: Session Evaluation Questionnaire; Italian sample; Session impact; Counselling outcome.

Introduction

In literature, as well as in therapeutic and counselling practice, the impacts that a clinical session process has on the patient are well known (Bergin & Lambert, 1978; Lambert, Bergin, & Collins, 1977). According to Stiles et al. (1994), impact is defined as the *immediate subjective effects, including clients' evaluation of the session, their*

assessment of the session's specific character, and their post-session affective state (p. 175). The impact can be considered as a mediator between process and outcome (Stiles, 1980). For the evaluation of the perceived impact, the best unit of analysis seems to be the single session, because it is more manageable for both the entire therapy and the single speaking turns and, at the same time, gives the possibility of bridging micro- and macro-analytic units (Stiles, 1980; Elliott & Wexler, 1994).

The evaluation of a single session's impact (that is strictly linked to the session's consequences, effects, and outcomes; Elliott, 1991) from a patient's point of view could allow a therapist to answer questions and solve dilemmas that patient develops after the session. For instance, a therapist can ask himself or herself which is the emotional reaction that the patient developed in respect with a certain topic (even if he or she does not have a clear idea of his or her problem); if the pressure that the clinician used (if he or she used it) in regard to a certain topic was too strong for the patient; if what was said in the session appears clear and useful to a therapist's personal experience and has the same features as a patient's experience (or, for instance, the opposite). These and other questions underline the importance of focusing on the general emotional quality of clients' reactions to sessions, which are not necessarily linked to a specific area or content.

The development of this perspective led to the creation of a questionnaire that was designed to point out the impact that a clinical session has provoked in patients: the

Correspondence: Diego Rocco, Department of Developmental and Social Psychology, University of Padua, Via Belzoni 84, 35122 Padua, Italy.
Email: diego.rocco@unipd.it

Citation: Rocco, D., Salcuni, S., & Antonelli, E. (2017). A pilot study of the Italian adaptation of the Session Evaluation Questionnaire fourth version. *Research in Psychotherapy: Psychopathology, Process and Outcome*, 20(2), 122-130. doi: 10.4081/ripppo.2017.269

Received for publication: 30 March 2017.
Accepted for publication: 26 June 2017.

This work is licensed under a Creative Commons Attribution Non-Commercial 4.0 License (CC BY-NC 4.0).

©Copyright D. Rocco et al., 2017
Licensee PAGEPress, Italy
Research in Psychotherapy: Psychopathology, Process and Outcome 2017; 20:122-130
doi:10.4081/ripppo.2017.269

Session Evaluation Questionnaire (SEQ; Stiles, 1980). The SEQ can be completed by both patients and psychologists/therapists, and has been applied to many types of individual therapy sessions, to group therapy and encounter group sessions, family and marital sessions, and supervision sessions (Stiles, Gordon, & Lani, 2002). It consists of a list of bipolar adjective scales presented in a seven-point (from 1 to 7), semantic differential format (Osgood, Suci, & Tannenbaum, 1957), and respondents are instructed: *Please circle the appropriate number to show how you feel about this session.* The number of items comprising the SEQ has been changed along the line, and they passed from 27 in the fourth version (Stiles et al., 1994) to the 21 of the current fifth version (Stiles, Gordon, & Lani, 2002). However, as a starting point of the Italian validation, we chose the fourth version of SEQ as proposed by Stiles et al. (1994), to guarantee a larger number of items for the factor analysis. This version is identical to Form 4 (Dill-Standiford, Stiles, & Rorer, 1988; Stiles et al., 1994), and the items are divided into three sections. The first section, which concerns the session evaluation, counts 12 items and is prefaced by the stem *This session was...*, while the second section, that concerns the post-session mood, counts 12 items and is prefaced by the stem *Right now I feel...* Finally, the third section, concerning a patient's evaluation of a therapist, includes three items and is prefaced by the stem *Today I feel my therapist was...* (this section was present only in the patient's SEQ version).

Items belonging to the first section of the SEQ fourth version are: bad–good, safe–dangerous, difficult–easy, valuable–worthless, shallow–deep, relaxed–tense, unpleasant–pleasant, full–empty, weak–powerful, special–ordinary, rough–smooth, and comfortable–uncomfortable. Items belonging to the second section of SEQ are: happy–sad, angry–pleased, moving–still, uncertain–definite, calm–excited, confident–afraid, wakeful–sleepy, friendly–unfriendly, slow–fast, energetic–peaceful, involved–detached, and quiet–aroused. Finally, the items present in the third section of SEQ are: skillful–unskillful, cold–warm, and trustworthy–untrustworthy.

The adjectives characterized by a positive and negative connotation balance the left and right polarities, with the aim to minimize response tendency, and the questionnaire presents the items in a mixed order within each section. Factor analysis (Stiles et al., 1994; Stiles, 1980; Reynolds et al., 1996; Stiles & Snow, 1984) shows that the adjective scales of the first section are organized in two dimensions, Depth and Smoothness, respectively, each consisting of five items, and the same is true also for the second section with the dimensions Positivity and Arousal. The internal consistency for all four dimensions, measured by Cronbach's alpha, was high (respectively, Depth=.87, Smoothness=.93, Positivity=.89 and Arousal=.78). Once a respondent fills out the SEQ, it is possible to calculate, for each dimension, an index by the

mean ratings of the items belonging to the considered dimension, with higher scores indicating greater Depth, Smoothness, Positivity, or Arousal. The index lies on the same seven-point scale as the individual items, making interpretation easier. The first polar scale, bad–good, was not included in these four dimensions, because for clients (but not for therapists) the item was split between two factors (Stiles & Snow, 1984; Stiles et al., 1994). The SEQ has been translated into many languages (for instance: Chinese, Spanish, Japanese, French; Stiles et al., 2002), however in the literature we found only two studies concerning the psychometric properties of the translations of the SEQ fifth version (Dutch translation; Hafkenscheid, 2009; German translation; Hartmann et al., 2013), and just one for the fourth version (Stiles et al., 1994).

In the literature is possible to find many works that use the SEQ. Dill-Standiford, Stiles and Rorer (1988) used it to evaluate the amount of agreement that counselors and patients have on their counseling sessions, showing that the agreement varied with the factor considered and the level of analysis. Stiles, Shapiro and Firth-Cozens (1988) used the SEQ to evaluate the different impact that an interpersonal-psychodynamic treatment has when compared to a cognitive-behavioural treatment. Friedlander, Siegel and Brenock (1989) analyzed the parallel processes activated during the clinical supervisions, and the SEQ was used together with the Working Alliance Inventory to measure the effect of *live supervision* compared to that of *videotape supervision* (Kivlighan, Angelone & Swafford, 1991). Samstag, Batchelder, Muran, Safran and Winston (1998) used the SEQ in a research design to determine the predictive validity of in-session measures. Duan and Kivlighan (2002) used it as session outcome measure, highlighting how a clinician's empathy seems to be in a relation with his or her SEQ's Depth dimension. Pesale and colleagues (Pesale, Hilsenroth, & Owen, 2012) considered the dimensions of the SEQ in relation to outcomes during psychodynamic psychotherapy, and they obtained meaningful results: even if patients' Depth rating was not found to be significantly related to outcome, higher patients' Smoothness dimension was linked to patients' improvement at the end of treatment in terms of global psychopathology and broad-range functioning, with a decrease in general symptoms. Finally, in Italy, Lingiardi, Colli, Gentile and Tanzilli (2011) used the translation of the SEQ Depth scale to assess the depth of psychotherapy sessions from an external rater's perspective. To our knowledge, the SEQ was neither translated nor validated in Italian in whole, even though the importance of the construct measured by the SEQ justifies an extensive use of it in current psychotherapy research.

For all these reasons we believe that the clinical value of a tool as the SEQ for our clinical activity, both as psychotherapists and as supervisors of young clinicians in their training, fully justifies the work of adaptation of its Italian version.

Purpose of the present study

The present study has two goals: the first - mainly methodological- is attempting to replicate the dimensional structure of the SEQ fourth version for its Italian translation. We translated the fourth version of the questionnaire with 27-item used by Stiles et al. (1994), which could reproduce the American SEQ's dimensionality (Stiles & Snow, 1984) — and not its most recent version, the fifth, 21-item version — to take advantage of a greater number of items for the factor analysis. The second goal is both methodological and clinical: in fact, to assess the SEQ convergent validity, we decided to utilize a questionnaire that measures patients' satisfaction about the consultation process. The discussion section will include methodological as well as the clinical implications of the results of the present study.

Methods

This research received the approval of Padua University's Psychological Ethical Committee (Number 1550/2015), and the questionnaires administration was done from September 2015 to July 2016.

The research was conducted at the Dynamic Psychotherapy Service (SAP-DPS), a psychological assistance service of Padua University that offers to university students free clinical consultation sessions within a psychodynamic framework. Its goal is to help students in facing situations that cause psychological distress, being it relational (conflictual family life, problems with friends or partners), specific (eating disorders, addiction, complicated grieving, panic attacks, depression, self-esteem problems, difficulty in managing affects like anger, anxiety, sadness, and so on), or more general (confusion, difficult problem solving but not cognitive difficulties in facing exams or fears about entering the job market after graduation not linked to cognitive issues).

Participants

We included as potential participants in the study all the patients who started a consultation process at the Dynamic Psychotherapy Service for university students, from September 2015 to July 2016. All the patients, but 10 who refuse for personal reason to participate to the research, voluntarily accepted to participate. Participants were 111 students (87 females and 24 male) enrolled in Padua University following their courses in different Faculties; about 50% of the students was from Psychology and the other 50% spread out into the other Faculties (Law 9%, Political Sciences and Engineering about 5%, and so on). Their average age was 22.83 (range 19–35; SD 2.72).

Instruments

The replication of the dimensional structure of the SEQ in Italian was pursued taking into consideration what

Beaton, Bombardier, Guillemin and Bosi Ferraz (2000), Caro and Stiles (1997), and Hambleton and Patsula (1999) reported about difficulties and potentially misleading results that a translation of a personality test or of a questionnaire can produce, even if some strategies such as backward translation are adopted.

As Beaton et al. (2000) suggested, we followed a *cross-cultural adaptation* guideline to encompass a process that looks at both language (translation) and cultural adaptation (meanings) in preparing the SEQ in Italian. In fact, when measures might be used across cultures, the items must not only be translated well linguistically, but they must be adjusted culturally, if it is necessary, to maintain both content and conceptual validity of the tool (Beaton et al., 2000). The fourth version of the SEQ (Stiles et al., 1994) was translated into Italian by two translators independently, and after an agreement among them, it was sent to a native English-speaking proofreader with an excellent knowledge of Italian language for the backward translation. The three authors, therefore, reviewed all the translations and reached a consensus on any discrepancy, in language and in the contents of the items, to achieve equivalence between the source and SEQ Italian version. This process led to a perfect coincidence with the SEQ original version. The obtained result, which was considered the basis for our study, was comprised of 27 items divided into three sections. The first (session evaluation, introduced by *This session was...*) had 12 items, the second section (post-session mood, introduced by *Right now I feel...*) counted 12 items, and the third section (prefaced by the stem *Today I feel my therapist was...*) had three items (see Table 1 for the SEQ text).

A questionnaire about perceived satisfaction in the counseling process was also administered at the very end of the consultation process. Patients had to evaluate, on a 0 (nothing) to 100 (very much) scale, their perceptions about seven areas: the experience of being listened to and comprehended, the experience of being emotionally engaged, the experience of having a clearer definition of the clinical problem, the experience to have new perspectives on the problem, the experience to have bigger comprehension about themselves, the experience of general utility and finally the global satisfaction about the consultation process. This questionnaire is used in the SAP-DPS as a first indicator of the global outcome perceived by patients; moreover, it was used to analyze qualitatively the patients positive/negative attitude toward the clinician and his/her ability to understand and help. In this study, we utilized it to assess the convergent validity of the Italian SEQ, measuring patients' satisfaction.

Procedure

Each patient at the first contact with the secretary, who is a psychotherapist, received a screening battery, which included the Symptom Checklist 90 R (Derogatis, 1983), the Beck Depression Inventory-II (Beck, Steer, Ball, &

Ranieri, 1996), and the Millon Clinical Multiaxial Inventory III (Zennaro, Ferracuti, Lang, & Sanavio, 2008). Moreover, they received an informant consent form that explained the research project, including the administration of SEQ at the end of the first two sessions, and asked for their participation. Psychotherapists, trainees in psychotherapy, and psychiatrists composed the SAP-DPS team; all members received group supervision and had a collegial meeting once a week. The team monitors incoming patients and evaluates their symptoms and motivation to match patients with the appropriate clinician, looking at a patient's level of maladjustment and clinician's expertise. Then counseling is provided depending on single patient clinical characteristics and based on clinician's competence and availability.

Patients received their clinical interviews by 23 professionals (age $M=34.58$, $SD=6.82$): 19 psychologists in their professional training to become psychotherapists (all females, attending psychodynamic training institutes), and four experienced dynamic psychotherapists. Counselors had from 1 to 16 patients each and all of them met their clients at least in two sessions. Their clinical experience ranges from 1 to 3 years for psychologists in training, and from 7 to 19 years for the psychotherapists ($M=4.95$; $SD=4.01$).

The patients were free to accept or refuse participation in the research without any penalization for refusing. Patients were aware that their psychologists/therapists did not have access to the filled SEQ questionnaires. If a patient accepted, his or her counsellor, after each of the first two consultation sessions, gave him or her the questionnaires to be filled. Once a patient filled out the questionnaire, they were treated as confidential (code/name) and left in a specific box.

Statistical analyses

In order to verify the dimensional structure of the Italian SEQ, we mainly relied on previous studies, especially the ones of Stiles et al. (1994), relating to the SEQ fourth version, and of Hafkenscheid (2009), concerning the Dutch version of the SEQ fifth form. Consequently, in this first pilot study of the structure of Italian SEQ, we implemented an exploratory factor analysis (EFA), rather than a confirmatory factor analysis (CFA). In fact, as Tinsley and Tinsley (1987) stated, hypothesis testing using CFA constitutes a less stringent test of the hypothesized structure than does performing EFA and then relating its results to a hypothesis (Stiles et al., 1994; Hafkenscheid, 2009). On the other hand, Gerbing and Hamilton (1996) maintain that EFA can be used prior to analysis techniques aiming to confirm hypotheses on the data structure. The EFA was conducted separately for the three parts of which the SEQ is composed.

Following the above-mentioned studies, we used principal components extraction and varimax rotation.

For the interpretation of the rotated factor loadings,

we adopted the rules proposed by Hafkenscheid (1993, 2009): (a) only items with factor loadings of at least +0.40 were considered, provided that (b) the next largest loading on the other factor(s) was at least 0.20 lower and under the condition that (c) there were at least four items fulfilling both inclusion criteria (a) and (b). We think that, as other measures of the therapeutic process, the SEQ is mainly intended as a session-by-session measure, thus the session was used as a unit of analysis for the factor analysis. Besides this, EFA is a descriptive rather than an inferential statistical method; therefore, following Elliot and Wexler (1994), we consider that the non-independence of sessions within cases is not critical. Thus, differing from Hafkenscheid (2009) and Stiles et al. (1994), who used deviation scores both at the session and at the client level, we carried out the factor analysis of the patients' raw ratings to both of the sessions in which they participated (Elliot & Wexler, 1994). It was conducted separately on the three sections of the SEQ (evaluation of the session itself, feelings after the session, and evaluation of the therapist), both including and excluding the item bad-good. This last item tends to be used differently by therapists and by clients but has been retained on Form 5, as a single, separate measure, because of its intrinsic interest as a global evaluation item (Stiles et al., 1994, 2002).

The internal structure of the SEQ in terms of intercorrelations between raw dimensional scores was also assessed. Reliability analysis for each dimension of the SEQ was performed calculating internal consistency coefficients (Cronbach's α). We also tried to ascertain the convergent validity of the Italian SEQ by correlating the scores obtained in its subscales with the scores obtained by the questionnaire on the perceived satisfaction about the consultation process given by a subsample of the patients.

Results

Dimensional structure

We conducted a principal components analysis followed by an orthogonal varimax rotation for patients' responses to the three parts of the SEQ: evaluation of the session itself, feelings after the session, and evaluation of the therapist, considering patients' responses to both the sessions in which they took part. Tests that the correlation matrix could be factor analysed were all satisfactory: determinants were higher than 0, that is .010, .012, and .50 respectively for the three sections (meaning that the variables were not linear dependent); Kaiser-Meyer-Olkin (KMO) tests were .85, .82, and .68 (meaning that the samples were adequate), and Bartlett sphericity tests were all statistically significant, with $P<.001$ (meaning that correlation matrices were different from identity matrices). Table 1 shows factor loadings for the components extracted from the three parts of the SEQ. We present the solution including the item bad-good because the structure

and the order of the factor loadings did not change when the item bad-good was omitted.

For the first section of the SEQ (*This session was...*), a two-factor solution was obtained, with both eigenvalues exceeding 1, also supported by Cattell's scree-test (Cattell & Vogelmann, 1977). The two factors accounted for 55.82% of the total variance.

The factors can be named Depth and Smoothness; in fact, the patterns of factor loadings of the Italian SEQ translation are very similar to those obtained by Stiles et al. (1994), both for the Depth dimension and the Smoothness dimension. For Depth, the only exception concerned the item bad-good, which in our study loaded only on this factor, while in Stiles et al. (1994) study the item was split between Depth and Smoothness. As far as Smoothness is concerned, in our analysis the item safe-dangerous loaded on both factors, while in Stiles et al. (1994) study only on Depth; however, they excluded this item from their following analyses (p. 178).

For the second section of the SEQ (*Right now I feel...*), two factors were extracted (with both eigenvalues exceeding 1, also supported by Cattell's scree-test), which accounted for 53.15% of the total variance. The Positivity dimension included five items which are present in the American (Stiles & Snow, 1984) and British versions of the SEQ (Stiles et al., 1994), plus the item wakeful-sleepy, split between Positivity and Arousal in Stiles et al. (1994) study. The Arousal dimension included four items, excluding the item involved-detached and energetic-peaceful, which loaded on the two factors. The former item was also split between Positivity and Arousal in Stiles et al. (1994) study.

For the third section of the SEQ (*Today I felt my therapist was...*), only one factor emerged, named Good Therapist, which accounted for 66.25% of the total variance. This result replicates Stiles et al. (1994) finding.

Session Evaluation Questionnaire scores

Scores of the dimensions measured by the Italian SEQ were constructed on the basis of the factor analysis results. Each score was calculated as the mean of the items in italics in Table 1, excluding those with weak loadings or split between factors; two Depth scores were computed: one including bad-good (Depth1) and the other excluding it (Depth2).

Descriptive statistics and internal consistency are presented in Table 2. The scale range of the SEQ is 1–7, with 4 as the midpoint. All the mean scores are above 4, Good Therapist being the highest and Arousal the lowest.

Internal consistency of the dimensions was moderate or good, except for Arousal, for which it was sufficient ($\alpha=.65$).

The score distributions of four dimensions were normal, as attested by skewness and kurtosis values (included between -1 and +1). Instead, the distribution of Good Therapist score was negatively skewed, meaning that pos-

itive values of the response scale were the most frequent, while the high kurtosis index indicates that the distribution was narrower compared to the normal curve. To normalize the distribution, we applied a logarithmic transformation to the Good Therapist raw score (Barbaranelli, 2003; Barbaranelli & D'Olimpo, 2007).

Intercorrelations within the Session Evaluation Questionnaire

We computed Pearson's linear correlation coefficients among the five SEQ scores, with and without the item bad-good for Depth. They are shown in Table 3.

The intercorrelations of Depth with the other dimensions were very similar when bad-good was included (above the diagonal) and when it was not (below the diagonal). The only exception is with Smoothness: the low coefficient is statistically significant with the item bad-good and not significant when it was omitted.

The highest intercorrelation is between Depth and Good Therapist; a high intercorrelation was also found between Smoothness and Positivity. Moderate statistically significant intercorrelations emerged between Positivity and the Depth scores, and between Arousal and Depth. The only not significant correlation, close to zero, was between Arousal and Smoothness.

Considering the coefficients below the diagonal in Table 3, on the whole, this pattern of intercorrelations strongly resembles that obtained by Stiles et al. (1994) at the session level of analysis. The main differences concern the correlation between Depth and Good Therapist, which is much higher in our study (.66 vs .46) and the correlation between Arousal and Positivity, which also is higher in this study (.18 vs .10).

Correlations between Session Evaluation Questionnaire and the satisfaction score

Pearson's linear correlation coefficients were also calculated between the five SEQ scores obtained in the two sessions in which the patients took part and the score of perceived satisfaction in the consultation process given by 80 of them. Results showed that four coefficients were statistically significant, ranging from medium to low values: with Depth1 ($r=.40$; $P<.001$) and Depth2 ($r=.39$; $P<.001$), with Good Therapist ($r=.37$; $P<.001$), and with Positivity ($r=.24$; $P<.01$).

Discussion

This pilot study mainly sought to assess the SEQ's fourth version dimensionality in an Italian sample. Other analyses measured reliability, as internal consistency, and convergent validity.

The widely utilized distinction between Depth and Smoothness as separate evaluative dimensions of the session was supported. The Anglo-American Depth dimension

could be replicated quite perfectly in the Italian sample. Only the item bad-good, which in our study loaded only on the first factor, was split into Depth and Smoothness in Stiles et al. (1994). Nonetheless, this item has been retained in the SEQ because of its intrinsic interest as a global evaluation item, even if it is not included in the calculation of the dimensions scores (Stiles et al., 1994, 2002). On the whole, other results did not vary substantially when the item bad-good was included or not in the calculation of the Depth score. The Smoothness dimension could be reproduced perfectly in this study, also considering that the item safe-dangerous loaded both on Depth and on Smoothness, and was not included in the calculation of the Smoothness score. In fact, Stiles et al. (1994) neither used this item nor was it included in the SEQ's fifth version (Stiles et al., 2002).

The factorial structure of the second part of the SEQ, referring to patients' post-session moods, was replicated to

a substantial degree. The Positivity dimension comprised the five items of the Anglo-American version plus the item wakeful-sleepy. This finding can be easily explained: in the Italian language wakeful-sleepy, besides its usual meaning, also has a strong metaphorical meaning whereby it is used to refer to a smart person versus a dumb person. Thus it is not surprising that after a positive session, one can feel happier, more friendly, etc., but also more wakeful, in the sense of smart. Anyway, this item, which was split between the two factor in Stiles et al. (1994) study, was dropped from SEQ's Form 5 (Stiles et al., 2002).

The Arousal dimension comprised four of the six original Anglo-American items, excluding energetic-peaceful and involved-detached that were split between the two post-session mood factors. As far as the former item, we think that these adjectives, with low frequency of use in Italy, may have generated a problem in their translation,

Table 1. Principal components analysis of the Italian version of the Session Evaluation Questionnaire.

	Component		Communality
	1	2	
First section SEQ (<i>This session was...</i>)			
<i>Depth</i>			
<i>Profonda-superficiale</i> (deep-shallow)	.80	.015	.64
<i>Speciale-ordinaria</i> (special-ordinary)	.79	-.071	.62
<i>Potente-debole</i> (powerful-weak)	.77	-.023	.59
<i>Piena-vuota</i> (full-empty)	.76	.15	.60
<i>Di valore-senza valore</i> (valuable-worthless)	.74	.11	.56
<i>Buona-cattiva</i> (good-bad)	.65	.39	.57
<i>Smoothness</i>			
<i>Rilassata-tesa</i> (relaxed-tense)	.001	.79	.62
<i>Confortevole-disagevole</i> (comfortable-uncomfortable)	.36	.71	.63
<i>Liscia-ruvida</i> (smooth-rough)	-.086	.69	.48
<i>Facile-difficile</i> (easy-difficult)	-.36	.69	.60
<i>Piacevole-spiacevole</i> (pleasant-unpleasant)	.24	.69	.53
<i>Sicura-pericolosa</i> (safe-dangerous)	.29	.42	.26
Second section SEQ (<i>Right now I feel...</i>)			
<i>Positivity</i>			
<i>Felice-triste</i> (happy-sad)	.80	.12	.66
<i>Amichevole-ostile</i> (friendly-unfriendly)	.78	.15	.63
<i>Fiducioso-pauroso</i> (confident-afraid)	.74	.093	.55
<i>Ben disposto-arrabbiato</i> (pleased-angry)	.71	.017	.50
<i>Certo-incerto</i> (definite-uncertain)	.64	-.023	.40
<i>Sveglio-assonnato</i> (wakeful-sleepy)	.58	.36	.47
<i>Arousal</i>			
<i>Attivato-quieto</i> (aroused-quiet)	-.17	.81	.68
<i>Eccitato-calmo</i> (excited-calm)	-.36	.69	.60
<i>Energetico-pacioso</i> (energetic-peaceful)	.49	.61	.61
<i>In movimento-fermo</i> (moving-still)	.24	.57	.38
<i>Veloce-lento</i> (fast-slow)	.36	.57	.46
<i>Coinvolto-distaccato</i> (involved-detached)	.42	.52	.45
Third section SEQ (<i>Today I felt my therapist was...</i>)			
<i>Good therapist</i>			
<i>Abile-non abile</i> (skilful-unskilful)	.84		.70
<i>Affidabile-non affidabile</i> (trustworthy-untrustworthy)	.83		.69
<i>Caldo-freddo</i> (warm-cold)	.77		.60

SEQ, Session Evaluation Questionnaire. n=222 sessions. Each score was calculated as the mean of the items in italics.

and/or may not be well understood by patients. On the other hand, involved-detached loaded on the two factors in the study of Stiles et al. (1994), too, and it does not appear in SEQ's fifth version.

The structure of the third part of the SEQ, concerning the evaluation of the therapist by patients, showed the factor Good Therapist, overlapped with the one found in Stiles et al. (1994) study.

Reliability of the dimensions as internal consistency varied from good to sufficient, with Cronbach's α being higher than .74, except for Arousal, with a coefficient of .65. Also in the Dutch version of the SEQ (Hafkenscheid, 2009), this dimension showed the lowest coefficient. The pattern of SEQ intercorrelations resembles the one found in Stiles et al. (1994) study. The highest coefficient was between Good Therapist and both versions of Depth. The intercorrelation between Positivity and Smoothness was very high, too. From a clinical perspective, these data seem to suggest that in this counselling context the patients express a good evaluation of counsellors/therapists when they not only gather information but also smoothly accompany the patients becoming more deeply in touch to their problems.

On the whole, these preliminary results appear noteworthy, especially bearing in mind that in our study, in order to facilitate a comparison of the Anglo-American and Italian results, for the main steps of the factor analysis we followed the methods used by Stiles et al. (1994), but we started from the raw scores (Elliott & Wexler, 1994) and

not from the deviation scores as they did. The final data we have obtained at the end of this process give us a SEQ Italian version that substantially overlapped the current English SEQ version (fifth version; Stiles et al., 2002). In fact, concerning the SEQ's first section (*This session was...*) only one item differs in the Italian version compared to Stiles et al. (2002) version: it is the item bad-good, which loaded on our Depth factor differently from theirs, in which the authors retained the item alone as a general evaluative measure. As far as the second section (*Right now I feel...*), the difference between the two versions concerns two items: wakeful-sleepy, which was included in our Smoothness dimension but not in SEQ fifth version (Stiles et al., 2002) same dimension, and energetic-peaceful, which even if it is present in SEQ fifth version's Arousal dimension, is not present in the same dimension of our version.

We assessed the first form of convergent validity of the Italian SEQ by correlating its scores with an index of the perceived satisfaction in the counselling process given by a subsample of the patients. The validity of the first version of the Italian SEQ was partially supported: significant, positive, and discrete correlations emerged with Good Therapist and both versions of Depth; anyway, a lower one emerged with Positivity dimension. This data could be interpreted reminding that patients compiled the SEQ after the first two counselling sessions but the questionnaire about perceived satisfaction at the very end of the consultation, after other 2 sessions. The whole consultation

Table 2. Means, standard deviations, skewness, kurtosis, internal consistency reliability (Cronbach's α) and their confidence intervals of the Italian Session Evaluation Questionnaire.

	N. items	Mean	SD	Skewness	Kurtosis	Reliability	95% CI	
							Lower bound	Upper bound
Depth1	6	5.43	.92	-.68	.63	.86	.82	.88
Depth2	5	5.29	.98	-.52	.18	.85	.81	.86
Smoothness	5	4.66	1.05	.029	-.49	.77	.71	.81
Positivity	6	4.84	.97	-.32	-.21	.83	.79	.86
Arousal	4	4.21	.99	.34	-.17	.65	.56	.72
Good therapist	3	5.83	.95	-1.45	3.31	.74	.67	.79

SD, standard deviation; Depth1, score calculated including the item *bad-good*; Depth2, score calculated without the item *bad-good*; CI, confidence interval. n=222 sessions.

Table 3. Intercorrelations (Pearson's r) of the five dimensions of the Italian Session Evaluation Questionnaire.

	Depth	Smoothness	Positivity	Arousal	Good therapist
Depth	-	.13*	.38***	.23***	.69***
Smoothness	.08	-	.58***	-.06	.23**
Positivity	.34***	.58***	-	.18**	.42***
Arousal	.25***	-.06	.18**	-	.20**
Good therapist	.65***	.22**	.40**	.15**	-

n=222 sessions. Above the diagonal correlations of Depth 1, including *bad-good*; below the diagonal correlations of Depth 2, without *bad-good*. * $P < .05$; ** $P < .01$; *** $P < .001$.

process includes a motivational and clinical restitution to the patient about his/her problem and situation; probably, the total satisfaction perceived in the complexity of the counselling is just partially linked to the perceived sensation of Positivity related to the first two sessions of the consultation.

Our results allow us to adopt the SEQ as a reliable tool in our clinical SAP-DPS routine, as well as in our private practice: its good factorial structure as well as its clinical information gathered about patients' experience in the counselling sessions, are definitively useful for different reasons. Firstly, two of us are involved as supervisors in the Dynamic Psychotherapy Service (SAP-DPS), where the research was developed. In this activity, one of our goals is to monitor the impact of clinical sessions to help both psychologists/psychotherapists and patients to discover if the sessions are leading to some benefits or if they are ineffective or even harmful. Secondly, part of our interest in this questionnaire is due to a research project in which we are developing a new questionnaire concerning the global attitude that a therapist has toward a patient (Post-session Therapy Questionnaire, PSTQ): to this aim it is necessary to have the SEQ in its Italian version to verify our working hypothesis. Moreover, we will validate the Italian version of SEQ also for the therapists.

Conclusions

The present pilot study suffered from some limitations, the main concerning the small number of patients and sessions. This is one of the reasons why, together with the theoretical one illustrated above (see Statistical Analysis paragraph), we did not carry out a CFA, which requires greater numbers of subjects and sessions (Kahn, 2006). We will perform CFA in future validation studies. Further studies are needed also to better assess the validity of the Italian SEQ, and to consider the possibility of adapting some items to Italian respondents. Considering these limitations, we can say that the first version of the Italian SEQ has acceptable psychometric qualities.

References

- Barbaranelli, C. (2003). Analisi dei dati. *Tecniche multivariate per la ricerca psicologica e sociale* [Data analysis. Multivariate techniques for psychological and social research]. Milano: LED.
- Barbaranelli, C., & D'Olimpio, F. (2007). *Analisi dei dati con SPSS. I. Le analisi di base* [Data analysis with SPSS. I. Basic analysis]. Milano: LED.
- Beaton, D.E., Bombardier, C., Guillemin, F., & Bosi Ferraz, M. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *PINE*, 24-25, 3186-3191.
- Beck, A.T., Steer, R.A., Ball, R., & Ranieri, W. (1996). Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients. *Journal of Personality Assessment*, 63, 588-597.
- Bergin, A.E., & Lambert, M.J. (1978). The evaluation of therapeutic outcome. In: S.L. Garfield & A.E. Bergin (Eds.), *Handbook of psychotherapy and behaviour change: An empirical analysis* (2nd ed.) (pp. 357-361). New York: Wiley.
- Caro, I., & Stiles, W.B. (1997). Vamos a raducir los MRV (let's translate the VRM): Linguistic and cultural inferences drawn from translating a verbal coding system from English into Spanish. *Psychiatry*, 60, 233-247.
- Cattell, R.B., & Vogelmann, S. (1977). A comprehensive trial of the scree and KG criteria for determining the number of factors. *Multivariate Behavioral Research*, 12, 3-30.
- Derogatis, L.R. (1983). *SCL-90-R: Administration, scoring, and procedures manual II*. Baltimore: Clinical Psychometric Research.
- Dill-Standiford, T.J., Stiles, W.B., & Rorer, L.G. (1988). Counselor-client agreement on session impact. *Journal of Counseling Psychology*, 35, 47-55.
- Duan, C., & Kivlighan, D.M. (2002). Relationships among therapist pre-session mood, therapist empathy, and session evaluation. *Psychotherapy Research*, 12, 23-37.
- Elliott, R. (1991). Five dimensions of therapy process. *Psychotherapy Research*, 1, 92-103.
- Elliott, R., & Wexler, M.M. (1994). Measuring the impact of sessions in process-experiential therapy of depression: the Session Impacts Scale. *Journal of Counseling Psychology*, 41, 166-174.
- Friedlander, M.L., Siegel, S.M., & Brenock, K. (1989). Parallel processes in counseling and supervision: a case study. *Journal of Counseling Psychology*, 36, 149-157.
- Gerbing, D.W., & Hamilton, J.G. (1996). Viability of exploratory factor analysis as a precursor to confirmatory factor analysis. *Structural Equation Modeling*, 3(1), 62-72.
- Hafkenscheid, A. (1993). Psychometric evaluation of the symptom checklist in psychiatric inpatients. *Personality and Individual Differences*, 14, 751-756.
- Hafkenscheid, A. (2009). The impact of psychotherapy sessions: internal structure of the Dutch Session Evaluation Questionnaire (SEQ). *Psychology and Psychotherapy: Theory, Research and Practice*, 82, 99-111.
- Hambleton, R.K., & Patsula, L. (1999). Increasing the validity of adapted tests: myths to be avoided and guidelines for improving test adaptation practices. *Journal of Applied Testing Technology*, 1, 1-30.
- Hartmann, A., Leonhart, R., Hermann, S., Joos, A., Stiles, W.B., & Zeeck, A. (2013). Die Evaluation von Therapiesitzungen durch Patienten und Therapeuten. Faktorstruktur und Interpretation des SEQ-D. *Diagnostica*, 59(1), 45-59.
- Kahn, J.H. (2006). Factor analysis in counseling psychology research, training, and practice: principles, advances, and applications. *The Counseling Psychologist*, 34(5), 684-718.
- Kivlighan, M., Angelone, E.O., & Swafford, K.G. (1991). Live supervision in individual psychotherapy: effects on therapist's intention use and client's evaluation of session effect and working alliance. *Professional Psychology: Research and Practice*, 22, 489-495.
- Lambert, M.J., Bergin, A.E., & Collins, J.L. (1977). Therapist-induced deterioration in psychotherapy. In: A.S. Gurman & A.M. Razin (Eds.), *Effective psychotherapy; a handbook of research*. Oxford, UK: Pergamon Press.
- Lingiardi, V., Colli, A., Gentile, D., & Tanzilli, A. (2011). Exploration of session process: relationship to depth and alliance. *Psychotherapy*, 48, 391-400.
- Osgood, C.E., Suci, G.J., & Tannenbaum, P.H. (1957). *The meas-*

- urement of meaning. Urbana, IL: University of Illinois Press.
- Pesale, F.P., Hilsenroth, M.J., & Owen, J.J. (2012). Patient early session experience and treatment outcome. *Psychotherapy Research, 22*, 417-425.
- Reynolds, S., Stiles, W.B., Barkham, M., Shapiro, D.A., Hardy, G.E., & Rees, A. (1996). Acceleration of changes in session impact during contrasting time-limited psychotherapies. *Journal of Consulting Clinical Psychology, 64*, 577-586.
- Samstag, L.W., Batchelder, S.T., Muran, J.C., Safran, J.D., & Winston, A. (1998). Early identification of treatment failures in short-term psychotherapy. *Journal of Psychotherapy Practice and Research, 7*, 126-143.
- Stiles, W.B. (1980). Measurement of the impact of psychotherapy sessions. *Journal of Consulting and Clinical Psychology, 48*, 176-185.
- Stiles, W.B., Gordon, L.E., & Lani, J.A. (2002). Session evaluation and the Session Evaluation Questionnaire. In: G.S. Tryon (Ed.), *Counseling based on process research: applying what we know* (pp. 325-343). Boston, MA: Allyn & Bacon.
- Stiles, W.B., Reynolds, S., Hardy, G.E., Rees, A., Barkman, M., & Shapiro, D.A. (1994). Evaluation and description of psychotherapy sessions by clients using the session evaluation questionnaire and the session impacts scale. *Journal of Counseling Psychology, 41*, 175-185.
- Stiles, W.B., Shapiro, D.A., & Firth-Cozens, J.A. (1988). Do sessions of different treatments have different impacts? *Journal of Counseling Psychology, 35*, 391-396.
- Stiles, W.B., & Snow, J.S. (1984). Dimensions of psychotherapy session impact across sessions and across clients. *British Journal of Clinical Psychology, 23*, 59-63.
- Tinsley, E.A., & Tinsley, D.J. (1987). Use of factor analysis in counseling psychology research. *Journal of Counseling Psychology, 34*(4), 414-424.
- Zennaro, A., Ferracuti, S., Lang, M., & Sanavio, E. (2008). *L'adattamento italiano del MCMI-III. Studi di validazione*. Firenze: Giunti O.S.