



UNIVERSITY OF ÉVORA

SOCIAL SCIENCE SCHOOL

DEPARTMENT OF PSYCHOLOGY

**Why using individualized outcome
measures in mental health? A thematic
comparison of patient-generated items
in PQ with CORE-OM and PHQ-9**

Rita Carlota

Supervisor: Ph.D Célia Sales, University of Évora

Co-supervisor: Ph.D Robert Elliott, University of Strathclyde

Master in Psychology

Specialization area: *Health and Clinical Psychology*

Évora, 2016



UNIVERSIDADE DE ÉVORA

SOCIAL SCIENCE SCHOOL

Master in Psychology

Specialization in Health and Clinical Psychology

Why using individualized outcome measures in mental health? A thematic comparison of patient-generated items in PQ with CORE-OM and PHQ-9

Rita Carlota

Supervisor:

Ph.D. Célia Sales

Co-supervisor:

Ph.D. Robert Elliott

April, 2016

Acknowledgements

I want to thank all who contributed to the realization of my master's thesis:

First to my supervisor Dr. Celia Sales, for all the work and patience. Without her support none of this would have been possible. This experience allowed for me to grow, both professionally and personally. I would also like to thank my co-supervisor, Dr Robert Elliott, for accepting this collaboration and also for his contribution to this thesis.

The entire research team that enabled the data collection in the Hospital Espírito Santo in Évora; patients of the Department of Psychiatry and Mental Health who agreed to participate in the study and also the technicians and therapists who were always promptly available to help in the implementation of the data collection method in the department.

To my colleagues Claudia Brinquete and Inês Neves for all the support and help during this phase of our academic life.

Also to Dr. Paula Alves, researcher at ISCTE, for the data provided and help with treating these.

Finally, to all my friends and family for their support, especially Pedro for always believing in me.

Why using individualized outcome measures in mental health? A thematic comparison
of patient-generated items in PQ with CORE-OM and PHQ-9

Abstract

This study aims to investigate the utility of PQ (Personal Questionnaire) as an individualized measure. PQ is a tool built by the patient, with the purpose of measuring the patient's changes during the process of therapy. We intend to explore the ability of this tool to add information concerning the patient's problems when compared to standardized instruments. From a sample of 105 patients, gathered from a clinical population and from a population of drug abuse, 563 items were collected; these were categorized into 65 sub-themes and classified according to their quality. "Addiction" was the most indicated sub-theme by the sample, mentioned in 46 items (17.8%). About 38% of the items were not covered by the CORE-OM and about 71% were not covered by the PHQ-9. A large part of the sample (69%) showed at least a sub-theme not covered by the CORE-OM and practically the whole sample (97%) mentioned at least a sub-theme not represented by PHQ-9.

Keywords: individualized measures, standardized measures, outcome measures

Porquê usar medidas de resultado individualizadas em saúde mental? Uma comparação temática de itens gerados pelo paciente no PQ com o CORE-OM e o PHQ-9

Resumo

Este estudo pretende investigar a utilidade do PQ (Personal Questionnaire) como medida individualizada. O PQ é um instrumento construído pelo paciente, com a finalidade de medir mudanças do paciente durante o processo de terapia. Pretendemos explorar a capacidade deste instrumento em adicionar informação sobre os problemas dos pacientes quando comparado a instrumentos standartizados. De uma amostra de 105 pacientes, provenientes de uma de população clínica e de uma população de abuso de drogas, recolheram-se 563 itens; estes foram categorizados em 65 subtemas e classificados segundo a sua qualidade. O subtema mais indicado pela amostra foi “Dependência”, mencionado em 46 itens (8.17%). Cerca de 38% dos itens não foram cobertos pelo CORE-OM e cerca de 71% não foram cobertos pelo PHQ-9. Grande parte da amostra (69%) indicou pelo menos um subtema não coberto pelo CORE-OM e praticamente toda a amostra (97%) mencionou pelo menos um subtema não representado pelo PHQ-9.

Palavras-chave: medidas individualizadas, medidas standartizadas, medida de resultado

Contents

1. Introduction.....	1
2. Theoretical Background.....	3
2.1 Nomothetic approach of outcome assessment	3
2.2 Idiographic approach of outcome assessment.....	4
2.3 Patient-Generated Outcome Measures.....	5
2.4 Masurement properties of outcome tools: specificities of the individualized approach.....	7
3. Investigation proposal.....	13
4. Method	15
4.1 Sample 1	15
4.2 Sample 2	18
4.3 Instruments.....	20
4.4 Procedure for data collection	22
4.5 Data analysis procedure	23
5. Results	27
6. Discussion	37
7. References	41

List of Figures

Figure 1. Patient percentage of relative distribution according to the number of listed items.....	27
Figure 2. Sub-themes mentioned in the PQ.....	30
Figure 3. Sub-themes mentioned in the PQ, by each sample.....	31

List of Tables

Table 1. Summary table of psychometric criteria and clinical utility of PQ.....	12
Table 2. Sociodemographic variables of Sample 1.....	16
Table 3. Sociodemographic variables of Sample 2.....	19
Table 4. List of sub-themes.....	24
Table 5. PQ Item quality.....	28
Table 6. Item quality analysis.....	28
Table 7. Sub-themes not covered by CORE-OM and PHQ-9.....	32
Table 8. Number of responses and number of patients per sub-theme and its matching with the CORE-OM and PHQ-9.....	34

1. Introduction

In psychotherapy it is important to understand the changes that occur in patients during treatment. These changes are evaluated by outcome measures (Ogles, 2013). The most common form of assessment is made using standardized instruments in which patients select items that match their condition. However, there is a more sensitive alternative which takes into consideration the individual specification of each patient - individualized measures (Ashworth, Evans, & Clement, 2009). These allow patients to build items that correspond to the problems they want to work in therapy, instead of choosing preformatted answers (Elliott, et al., 2016).

Despite the increasing use of individualized methods little is known about its characteristics. In this sense, this study aims to explore some questions related to this type of instruments, particularly in regard to the information collected. Does the information of individualised outcome instruments add value to information collected by standardized instruments?

In order to answer these questions, we make a comparative analysis between the responses to these two types of instruments, examining whether the sub-themes listed in PGOM (patient-generated outcome measurement) are present or not in standardized instruments.

2. Theoretical Background

The outcome in psychotherapy is a more complex concept than one might think, because there are many definitions for "what it is that changes" (Ogles, 2013, p. 136) during the therapeutic process. The measurement of outcome vary from symptoms or problems to psychological conflicts that arise in relational issues even including personality structure. The areas diversify depending on the theoretical basis that each therapist chooses to use, so there are many perspectives and issues that can be considered (Ogles, 2013).

2.1 Nomothetic approach of outcome assessment

Standardized measures were the first to be developed. Having a self-report format, they are the most common instruments to measure the changes that occur during psychotherapy (Ogles, 2013). This type of instruments, also called nomothetic measures, aim to investigate a large population, working overall dimensions that may apply to all individuals. They use sets of predetermined items based on the specific area in which it is intended the evaluation. Since the items are derived from problems and symptoms reported by general clinical population it is expected that all patients get a result in relation to the instrument (Ashworth, et al., 2007; Evans, Margison, & Barkham, 1998).

These instruments are easy to apply: consist in lists of responses from which patients choose the most similar to their condition; generally are paper-pen format; require no technical monitoring and are relatively fast to fill. They are used not only in psychological assessment, but also in research settings. Its psychometric characteristics allow a fairly rigorous assessment of change in psychotherapy. Once this type of instrument has a quantitative analysis, it becomes easy to compare patient's answers with normative samples, giving a better understanding of the results (Barkham, et al., 2001; Sales & Alves, in press; Sales, Gonçalves, Fragoeiro, Noronha, & Elliott, 2007; Overington & Ionita, 2012).

Standardized measures may vary depending on the concepts it assesses. Some are broad spectrum evaluating general concepts, such as the CORE-OM (Clinical Outcomes in Routine Evaluation - Evans, et al., 2000) that seeks to assess the psychological well-being. Others more specific focus only on certain concepts such as

the BDI (Beck Depression Inventory - Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) that evaluates depression.

Despite standardized measures are the most frequently used to evaluate outcome, some less positive aspects have been mentioned. One of the major criticisms against its use is that they tend to be less sensitive to the patient's change throughout the treatment because they contain few items relating to personal and private concerns of each patient (Hédinsson, Kristjánisdóttir, & Þór Ólason, 2013). Moreover, there is a wide range of problems such as loss or relationship problems, which turn out to be disregarded (Clark, Hook, & Stein, 1997). The fewer the items of the instrument, the more likely is the absence of specific items related with distress (Ashworth, et al., 2007). Patients also show some reluctance in its use saying that items can be vague, sets of response are limiting and have language problems, cultural differences and state-bias. (Crawford et al., 2002). For this reasons, patients are likely to have a non-representative score of their condition, which therefore influences psychotherapy results.

2.2 Idiographic approach of outcome assessment

In order to find an alternative to overcome the limits of nomothetic instruments, some authors argue an idiographic approach to measuring therapy outcome (Clark, Hook, & Stein, 1997; Elliott, 2010; Sales & Alves, 2012). An idiographic approach presupposes proximity with the patient and so a full involvement in the evaluation of their condition. Considering an involvement continuum, there is a range where only health professionals contribute to an assessment (minimum participation of patient) to a level where the health evaluation is done entirely by the patient (Sales & Alves, in press). These two opposing positions give patients different weights to assess their health state. Not involving the patient in this process, only the opinions and inferences made by the therapist on the patient's condition or about the treatment effect are valued. In a completely opposite position, patient's evaluation is purely based on their own opinion. This maximum patient's involment allows the development of individual measures. These "instruments that ask patients about their health" (Fitzpatrick, Davey, Buxton, & Jones, 1998, p. 4) are called patient-generated outcome measures (PGOM), or individualized outcome measures.

2.3 Patient-Generated Outcome Measures

In a different approach to nomothetic, PGOM follow an idiographic strategy that aims to maximize the patient involvement, by tailoring the outcome tool to the patient. These measures allow the patient to choose the symptoms and problems that are most important to work in therapy, taking into account their experiences, perspectives and narratives (Ashworth, et al., 2007; Robinson, Ashworth, Shepherd, & Evans, 2006). PGOM have an open-ended structure, which encourages the patients to express their problems or treatment goals, thus providing information on what they consider relevant to change (Sales & Alves, in press). Hence PGOM requires patients to define the contents of the specific instrument to be used, allowing a customized perspective. The Nacional Institute for Health and Care Excellence (2016) recommends the active involvement of patients in their treatment process, which is an indicator of quality of the therapeutic process. The use of PGOM becomes a good clinical practice, since it involves asking the patients to express in their own words the issues and concerns that they would like to work in therapy (Sales & Alves, in press).

Sales and Alves (in press) showed that the use of PGOM has been increasingly implemented and several advantages on the use of patient-generated measures have been reported: it allows patients to identify their own concerns; the patient's evolution is evaluated according to the relevance they give to their own problems (Ashworth et al, 2007; Robinson, Ashworth, Shepherd, & Evans, 2006) and there is no waste of time in the evaluation of symptoms that are not relevant to the patient (Wagner & Elliott, 2001). In addition, they enable the monitoring of the patient's progress on problems and goals, becoming thus an element of motivation. In interview-based instruments, it is possible to collect patient's idiosyncratic data, allowing for the establishment of a relationship between interviewer and interviewee (Turner-Stokes, 2011). Corroborating the purpose for which they were developed, Ashworth and colleagues (2007) concluded that patient-generated measures shown to be more sensitive to patients' clinical changes than standardized instruments, in a comparative study between PSYCHLOPS (Ashworth, et al., 2004) and CORE-OM (Evans, et al., 2000).

In a recent review, Sales and Alves (in press) found three PGOM - a goal attainment questionnaire and two target complaint questionnaires:

Goal Attainment Scaling (GAS)

GAS (Kiresuk & Sherman, 1968) is an interview-based procedure that focus on the objectives that patients want to achieve during therapy. In this instrument patients identify the most preoccupied problems and stablish a set of priority goals related to them. After this process, patients determine “the expected outcomes” for each goal corresponding to the “most probable result if the patient receives the expected treatment”. Finally, patients categorize from -2 (least expected outcome) to 2 (best possible outcome). Goals are supposed to be attained within a certain period of time.

Psychological Outcome Profiles (PSYCHLOPS)

PSYCHLOPS (Ashworth et al., 2004) evaluates the changes in patients' problems during treatment. It's a self-reported questionnaire that asks the patients to write down the two most worrying problems and what has been difficult to do because of those problems. Besides these questions, patients have to rate how much they've been affected by the problems during the past week (from 0 - “Not at all affected” to 5 - “Severely affected”) and how long have they been concerning them (from 0 – “Less than a month” to 4 – “More than 5 years”). Finally, there is an additional question about general well-being “How have you felt about yourself this last week?” that patients rate from 0 - “Very good” to 5 - “Very bad”.

There are three versions of PSYCHLOPS: pre-treatment, during treatment and end of treatment. The two latter versions are diferent. They ask if new problems have sprung, besides the ones already mentioned. The instrument used for the ending of the treatment asks the patient how he/she feels in relation to the beginning of treatment.

Personal Questionnaire (PQ)

According Sales and Alves (in press), the most popular patient-generated outcome measure is the PQ (Personal Questionnaire; Shapiro, 1961). Its first version allowed the comparison among different clients and their problems for specific psychological conditions. However, it was quite extensive and therefore Shapiro and others (McPherson & LeGassicke, 1965; Phillips, 1986; Shapiro, 1969) developed a new version. In this simplified version of PQ patients indicate just around 10 problems that they want to work in psicotherapy. There are five problem areas: symptoms, mood, specific performance, relationships and self-esteem. Application of PQ may be carried out both as a preparation for therapy or post-therapy for result discussion (Elliott, et al.,

2016). Since this instrument have been highlighted in the context of PGOM, we intend to explore its features, beyond the psychometric characteristics already validated by Elliott and colleagues (2016).

Patient-generated outcome measures are not widely used in mental health care, as shown in the Lambert and McRoberts' study in 1993, since time is a major concern in order to applying them (more than nomothetic instruments). In a literature review, Fitzpatrick, Davey, Buxton and Jones (1998) report that interview-based PGOMs show little feasibility when compared to self-completed instruments: time-consuming in application, complex, require a therapist to administer and seem to not generate normative data. However, research has been done showing that some of these facts are no longer true, as in the study of Elliott and colleagues (2016) which demonstrates and validates the psychometric characteristics of PQ. With increasing benefits in favor of a better knowledge and understanding of the patient, therapists consider that PQ is useful for both the pre-session and for the post-session, helping also in clinical decision making processes (Sales, Gonçalves, Fragoeiro, Noronha, & Elliott, 2007).

2.4 Measurement properties of outcome tools: specificities of the individualized approach

For a psychological scale to be used as an outcome measure, it needs to be evaluated according to psychometric characteristics and adequacy to the clinical context. We briefly present the following criteria: reliability, validity, sensitivity to change/responsiveness, precision, interpretability, appropriateness, acceptability and feasibility.

Reliability

Reliability concerns: reproducibility, assessing whether an instrument gets the same results in repeated applications when respondents have not changed in relation to the areas measured; and internal consistency, which checks whether the items are homogeneous and measure the same attribute. This criterion becomes essential when choosing an instrument since it verifies to what extent it is free of random errors, specifying whether any observed change is related or not with the problems of the

instrument itself (Fitzpatrick, Davey, Buxton, & Jones, 1998; Streiner & Norman, 2003). In a study by Elliott and colleagues (2016) involving five samples from three different countries, it could be verified that PQ shows moderate values of test-retest reliability ($r = .57$) and very high in internal consistency (range between .70 and .80); thus proves to be a useful tool for evaluation, showing no problems in its structure.

Validity

Validity intends to verify if the instrument really does measure what is intended to evaluate. There are several ways to confirm this: criterion validity - that is used when there is a proposal for a new instrument, comparing it to existing others already considered accurate; and construct validity – that checks whether the contents of the instrument is consistent with what is supposed to assess. The mental health measures assess constructs such as sadness, anxiety or isolation, which are different from other more 'visible' constructs. However, they represent aspects as important as other more directly observable. In this case, validity is assessed quantitatively by verifying its relation to other variables. It is, therefore, impossible to evaluate the validity of new instruments on their own. The more purposes included in an instrument, the more complex it is to verify its criterion of validity (Bergner & Rothman, 1987; Fitzpatrick, Davey, Buxton, & Jones 1998). PQ shows strong validity ($>.7$) when compared to other distress measures in different clinical populations, so we can conclude that it is equivalent to other instruments that purport to measure the same constructs (Elliott et al., 2016).

Sensitivity to change/Responsiveness

In a mental health instrument it becomes important to understand the changes of patients in relation to therapeutic process - it is necessary to take into account the responsiveness criterion, also known as sensitivity to change. A measure of change must be able to identify and detect differences as they are occurring in time. If significant differences show the alleged change, then it is considered that the instrument is sensitive to change (Ogles, 2013) The ability of an instrument to detect clinically important changes can be verified in several ways, being one of them the change scores. This method is based on calculating the change scores for some time as a longitudinal study or by checking for correlation changes of those scores

compared to other variables. Another way to study responsiveness is the effect size: calculate the size of change observed in a group of different applications, for example between an application made before treatment and other after treatment, and compare with the variability of values of the measure itself. Yet another way to verify responsiveness is by comparing the instrument with others considered effective, thus it is likely the occurrence of significant changes (Kirshner & Guyatt, 1985; Fitzpatrick, Davey, Buxton, & Jones, 1998; Kazis, Anderson, & Meenan, 1989). Elliott and colleagues (2016) have also shown that PQ shows sensitivity, since the analyzed sample showed enough differences before and after the session (even greater effects in patients with depression or anxiety disorders). Compared to other instruments, PQ was more sensitive to change.

Precision

One of the aspects that influence precision is how the instrument presents the response. At one extreme we have 'Yes' or 'No' answers, which do not allow the respondents to show more detailed difficulties. On the other, we can have a more graduated response, such as the Likert system. This last form of response increases in precision, the more categories are presented. The scores of each instrument items vary depending on the degree of coverage problems experienced by patients (Fitzpatrick, Davey, Buxton, & Jones, 1998). PQ (since the responses are free) enables each item to be as specific as the patient wants, making the instrument very detailed and representative of his condition (Elliott et al., 2016).

Interpretability

Interpretability has to do with how significant the results of an instrument are. The individualized measurements show less interpretability than other instruments, such as medical measures (tests of sugar levels or blood pressure). Fitzpatrick and colleges (1998) report that this also may be due to the less familiarity with the use of such instruments. However, this criteria also becomes important because it regards to the meaning of results that can be obtained with the instrument. In relation to the PQ, the interpretation of the data is not made in quantitative ways, but qualitative. Although the items are rated on a scale 1-7, the analysis is studying the issues highlighted by each patient and verifying if there are any changes in their classification and/or in their

problems. The interpretation of these changes gives meaning to the patients answers (Elliott et al., 2016).

Clinical utility

The criterion of clinical utility is usually not referred as essential for the choice of instrument to be used. However, Hunsley and Mash (2007) argue that the instrument's features must be consensual, meaning that the characteristics of their clinical use should be as important as the psychometric characteristics to be well accepted in clinical services. Thus, increasingly, researchers are interested both in efficiency and in practical issues of the instruments – there are ongoing studies about this area within our investigation group. Measures which are briefer, easy to understand and simple to score are more likely to be used either in clinical settings or in research settings (Ogles, 2013).

Appropriateness

This criterion relates to the adequacy of the instrument to the context in which it is supposed to be used. It is important to consider that the instrument should relate the best possible to the evaluation and intended purpose (Fitzpatrick, Davey, Buxton, & Jones, 1998). In the case of PQ, the choice of their use should take into account their application characteristics, since it is a time consuming instrument and requires a therapist in its administration. Due to it being conducted by interview and varying the time of application, their employability in a context of limited resources, technical and time, can be conditioned.

Acceptability

It is essential for an instrument to be well accepted by patients, which makes it a very important criterion when choosing an instrument. Generally, patients are already concerned about their problem, and the completion of questionnaires may increase stress. The acceptance of such measures has been less examined than other criteria and for that there are less consensus on what constitutes acceptability. If patients do not respond to an instrument or only respond to certain items, it can indicate that the instrument may be difficult to understand, cause distress or is unacceptable. High non-

response levels indicate that the instrument shows little validity. However, there may be other reasons for non-completion, such as the way how patients have access to instrument. Another way in which acceptability may be verified, is from the size of the measure: the longer it takes to complete, the less it will be acceptable (Ware, 1984; Fitzpatrick, Davey, Buxton, & Jones, 1998). In a study by Alves, Sales and Santos (2014) it was shown that when compared to two other instruments, PQ was the most accepted by patients. Notable for the fact that its interview format and establishment of goals had an extremely positive impact on patients (“I exposed the case of my drinking problem and if I had to write I would not say a thing”; “Helps to realize that even having made a mistake, if we are honest we have someone who can help us with our problems”, p. 25).

Feasibility

In addition to the already mentioned criteria, it is also important to understand the impact that the instrument has on professionals. Most of the time, the fulfillment of the assessment instruments is done in clinical setting, and the strain of application may jeopardize its completion (Fitzpatrick, Davey, Buxton, & Jones, 1998). For instance, measures that need a therapist to guide their fulfillment process, such as PQ. The fact that PQ is long and complex may influence adherence both on patients and professionals, albeit it showing very positive contributions to the therapeutic process.

Table 1. Summary table of psychometric criteria and clinical utility of PQ

Sensitivity to change		
Reliability	Reproducibility - moderate values of test-retest reliability (r=.57) Internal consistency - extremely high values (ranged between .70 and .80)	(Elliott et al., 2016)
Validity	Strong validity (>.7)	(Elliott et al., 2016)
Precision	Very detailed and representative of the condition of the patient	(Elliott et al., 2016)
Interpretability	Qualitative - problems and changes give meaning to the answers of patients	
Clinical Utility		
Appropriateness	Contexts with technical resources and application time	
Responsiveness	Show sensitivity to change	(Elliott et al., 2016)
Acceptability	Good acceptance - evidenced only positive aspects	(Alves, Sales, & Santos, 2014; Sales, Gonçalves, Fragoeiro, Noronha, & Elliott, 2007)
Feasibility	Adherence to their use can be affected by consuming time and resources	

3. Investigation proposal

Once known the psychometric characteristics of PQ mentioned in previous points, there isn't much work that presents results regarding the use of this measure, and in this sense this study comes to show its contribution. With the use and study of individualized measures it is possible to realize the benefits of its use, and there may be better understanding and interpretation of their results. In this context we intend to investigate: 1) the nature of the information provided by PQ and the extent to which adds content to standardized measures; 2) whether it is referred in PQ some theme that is not reported in standard measures; and 3) the methodological point of view - test analysis procedures of idiographic measures to complement the usual procedures for verification of the psychometric properties.

Realizing the disparities with regard to information obtained from the measures used in the study clarifies what characteristics influence the quality of data collected from both types of instruments. Thus, our research work concerns the comparison of results obtained by applying an idiographic measure (PQ; Elliot, Mack, & Shapiro, 1999) and two nomothetic (PHQ-9; Kroenke & Spitzer, 2002, and CORE-OM; Evans et al, 2000). This study is based upon work previously performed by Ashworth and colleagues in 2007, which carried out a comparison of themes from an idiographic measure (PSYCHLOPS; Ashworth, et al, 2004) and a nomothetic one also used in this work (CORE-OM).

4. Method

This study used two samples collected by different projects. The different samples are described below as well as the applied method.

4.1 Sample 1

Institutional context

The institutional context of the first sample is Hospital do Espírito Santo de Évora, specifically the Department of Psychiatry and Mental Health. This service is responsible for the prevention of mental health and its main mission is to promote mental health in Alentejo, dealing with psychiatric illnesses, whether outpatient or inpatient and taking responsibility for socio-professional and family rehabilitation and reintegration of patients. To meet its objectives, the department encompasses several services: Psychiatry Services, Nursing Services, Psychology Services, Psychiatric Unit for Children and Adolescents, Social Work, Speech Therapy Services, Occupational Therapy Services, Administrative Service, General Services and Internment Services.

Participants

Patients

The study participants are adults (age over 18 years) who were admitted for treatment at the Department of Psychiatry and Mental Health of Hospital Espírito Santo. The number of participants is the total number of patients who have agreed to collaborate in this study. The final sample of 57 patients were recruited between October 2013 and May 2014. The sample is constituted by 15 male patients (26.3%) and 42 (73.7%) female; ages are between 18 and 85 years ($M = 42.68$, $SD = 15:21$) and residence districts are Évora and Beja. The most frequent education level is 7th-9th grade (28.1%), with a variation from the 4th grade to Bachelor, Master or Doctorate. Most of the participants are full-time workers (45.6%), their marital status is married (43.9%) and has at least one child (71.4%). With regard to mental health history, 32 participants had psychological or psychiatric monitoring and 42 currently

take medication that aid to their well-being. Table 2 shows the sociodemographic variables of sample 1.

Research team

The sample was collected by a team of eight research assistants - six Master's students of Clinical and Health Psychology at the University of Évora and two psychologist trainees of the Department of Psychiatry and Mental Health of Évora Hospital.

Table 2. *Sociodemographic variables of Sample 1*

Variable	M	SD	n	%
<i>Gender</i>				
Female			42	73.7%
Male			15	26.3%
<i>Age</i>	42.68	15.12	57	
<i>Education Level</i>				
Up to 4th year of education			9	15.8%
5th to 6th year of education			11	19.3%
7th to 9th year of education			16	28.1%
10th to 12th year of education			12	21.1%
University attendance			5	8.8%
BSc/MSc/PhD			3	5.3%
Illiterate			1	1.8%
<i>District of Residence</i>				
Évora			56	98.2%
Beja			1	1.8%
<i>Marital Status</i>				
Married			25	43.9%
Divorced			9	15.8%
Single			16	28.1%
Life partners			5	8.8%
Widowed			2	3.5%

<i>Professional Status</i>			
Student		4	7.0%
Working student		2	3.5%
Working student seeking employment		1	1.8%
Full-time worker		26	45.6%
Part-time worker		2	3.5%
Unemployed		11	19.3%
Retired		11	19.3%
<i>Household Members</i>	2.66	1.30	56
<i>Number of Children</i>			
0		16	28.6%
1		13	23.2%
2		20	35.7%
3		6	10.7%
4		1	1.8%
<i>Previous Psychological/Psychiatric Support</i>			
Yes		32	56.1%
No		25	43.9
<i>Medication for psychological well-being</i>			
Yes		42	73.7%
No		15	26.3%
<i>Diagnosis</i>			
Anxiety/Depression		17	30.0%
Substance misuse		1	1.8%
Unknown		39	68.5%

4.2 Sample 2

Institutional context

Sample 2 comes from a research project (Alves, Sales & Ashworth, 2013) for a PhD thesis in psychology, funded by FCT (Foundation for Science and Technology) developed in ISCTE (Lisbon) in collaboration with King's College (London). Data were collected in four different institutions: Centro das Taipas, Lisbon; Unidade de Alcoologia, Lisbon; Centro de Respostas Integradas, Évora; and Comunidade Terapêutica de Esposende (for women with alcohol dependence). However, this study only used data from three of those institutions.

Participants

Patients

The study participants are adults (age over 18 years) who were admitted for treatment at Centro de Taipas, Centro de Respostas Integradas and at Unidade de Alcoologia. From a sample of 53 participants, five were excluded because they were not considered valid. Participants were recruited between April 2013 and April 2014.

The sample is constituted by 26 male patients (54.2%) and 14 (29.2%) female - this variable is missing for the remaining participants; ages are between 20 and 69 years ($M = 40.83$, $SD = 11:53$) and residence districts are Évora, Lisbon, Mora and Guarda. The most frequent education level is 7th-9th grade (28.2%), with variation from the 4th grade to Bachelor, Master or Doctorate. Most participants are unemployed (57.5%), single (41.7%) and has no children (37.5%). With regards to alcoholism or drug history, 12 subjects (31.6%) had experienced other treatments. Table 3 shows the sociodemographic variables of the sample 2.

Research team

The team for the data collection in this sample was consisted by four master's students in Psychology (three from the University of Évora and one of Psychology Faculty of the University of Lisbon), one PhD student of ISCTE - Lisbon University Institute and by a therapist of Comunidade Terapêutica of Esposende.

Table 3. Sociodemographic variables of Sample 2

Variable	M	SD	n	%
<i>Local Data Collection</i>				
Centro de Taipas (Lisboa)			12	25%
Centro de Respostas Integradas (Évora)			9	18.8%
Unidade de Alcoologia (Lisboa)			27	56.3%
<i>Gender</i>				
Female				
Male				
Age	40.83	11.53	40	
<i>Education Level</i>				
Up to 4th year of education			9	23.1%
5th to 6th year of education			7	17.9%
7th to 9th year of education			11	28.2%
10th to 12th year of education			7	17.9%
University attendance			4	10.3%
BSc/MSc/PhD			1	2.6%
<i>District of Residence</i>				
Évora			9	23.7%
Lisboa			27	71.1%
Mora			1	2.6%
Guarda			1	2.6%
<i>Marital Status</i>				
Married			10	25.0%
Divorced			9	22.5%
Single			20	50.0%
Widowed			1	2.5%
<i>Professional Status</i>				
Student			1	2.5%
Full-time worker			11	27.5%
Part-time worker			2	5.0%
Unemployed			23	57.5%

Retired			3	7.5%
<i>Household Members</i>	2.63	1.44	40	
<i>Number of Children</i>				
0			15	37.5%
1			12	30.0%
2			9	22.5%
3			2	5.0%
4			2	5.0%
<i>First Addiction Treatment</i>				
Yes			26	68.4%
No			12	31.6%

4.3 Instruments

4.3.1 PQ – Personal Questionnaire

PQ (Personal Questionnaire; Shapiro, 1961) is a patient-generated individualized outcome measure that has the purpose to measure changes in problem's patient throughout the therapy process. A trained interviewer asks patients about their difficulties and reviews the problems that were mentioned. The interviewer rewrites each problem to individual note cards, asking the client if he wants to include any others to the already given list and helps to clarify complex and ambiguous statements and phases out others, building a list of ten simple and nonredundant problems. Then, the patient ranks the problems from most important to least important using a 7-point anchored scale (1 - "not at all"; 2 - "very little"; 3 - "little"; 4 - "moderately"; 5 - "considerably"; 6 - "very considerably"; 7 - "maximum possible"). The patient also classifies the duration of each problem on a 7-point anchored scale (1 - "less than 1 month"; 2 - "1-5 months"; 3 - "6-11 months"; 4 - "1-2 years"; 5 - "3-5 years"; 6 - "6-10 years"; 7 - "more than 10 years"). Once PQ is completed, all the problems are typed and the patient has a blank space to add difficulties that could appear during the therapy process. On the following administrations, patients only classify severity (for procedure manual and blank forms, see Elliott et al., 1999).

The psychometric criteria of PQ have shown that the use of the instrument can offer positive contributions to therapeutic process. In terms of reliability, PQ has moderate values of test-retest ($r=.57$) and extremely high values of internal consistency (ranged between .70 and .80); shows strong validity ($>.7$) and revealed to be very detailed and representative of the condition of the patient (Elliott et al., 2016). Therefore PQ is a very reliable instrument for measuring the sensitivity to change in psychotherapy. A Portuguese version developed by Sales and colleagues (2007) was used in this study.

4.3.2 Clinical Outcomes in Routine Evaluation – Outcome Measure

CORE-OM (Clinical Outcomes in Routine Evaluation – Outcome Measure; Evans et al., 2000) is an instrument with 34 items. It measures the psychological well-being and should be filled by adults who are able to do it by themselves, since it is a self report measure (Barkham et al., 2001). Items are divided in 4 dimensions – well-being (four items), social functioning (twelve items), problems/symptoms (twelve items) and risk (six items) – and rated in a 5-point Likert-scale (0 – “not at all”; 1 – “only occasionally”; 2 – “sometimes”; 3 – “often”; 4 – “most or all the time”). The patients are asked to answer according to how they felt during the last week. Usually CORE-OM is filled before the therapy as a diagnostic measure, but it can be filled during or even at the end of therapy to monitor psychological changes (Evans et al., 2000; Sales, Moleiro, Evans, & Alves, 2012). The Portuguese version of CORE-OM has already been studied and shows good internal reliability ($>.8$) - proving that the Portuguese version of CORE-OM is also a valid instrument to measure psychological changes in therapy (Sales, Moleiro, Evans, & Alves, 2012).

4.3.3. Patient Health Questionnaire – 9 items

Patient Health Questionnaire – 9 items (PHQ-9; Kroenke, & Spitzer, 2002) is an instrument that measures depression. PHQ-9 allows to know the level of severity of depression based in the score given to the items (0 – “not at all”, 1 – “several days”, 3 – “more than half the days”; 4 – “nearly every day”). If the total score of the items is between 1 to 4 the instrument indicates None Depression Severity; scores between 5 to 9 indicate Mild Depression Severity; scores between 10 to 14 reveal a Moderate Depression Severity; a total score between 15 to 19 shows a Moderately Severe Depression and scores between 20 to 27 indicate a Severe Depression. In this

instrument, there is also a final question about the patient's level of function with the intention to validate the severity of depression expressed in the 9 items scored before – patients select the answer which is more representative of their condition (“not difficult at all”; “somewhat difficult”; “very difficult” or “extremely difficult”) (Kroenke, & Spitzer, 2002; Kroenke, Spitzer, & Williams, 2001; Spitzer, et al., 1999).

Monteiro and colleagues (2013) developed the Portuguese version of PHQ-9 and found its psychometrics criteria. Comparing the PHQ-9 with the Hospital Anxiety and Depression Scale (HADS; Zigmond, & Snaith, 1983) and the Beck Depression Inventory (BDI; Beck, et al., 1961) the Portuguese version showed a satisfactory internal consistency (Cronbach's $\alpha=.86$); revealed a moderate convergent validity with HADS depression ($r=.59$; $p<.01$) and HADS anxiety ($r=.61$; $p<.01$) and showed high convergent validity with the BDI ($r=.85$; $p<.01$), proving that PHQ-9 is an adequate measure for depression (Monteiro, et al., 2013).

4.3.4 Socio-demographic questionnaire

The socio-demographic questionnaire aims to collect information about the participants, as gender, age, education level, residence, marital status, professional status, household members and number of children. Besides these informations, in case of sample 1 it is also asked if the participant had previous psychological or psychiatric support, if medication is taken for psychological well-being and if a diagnosis is known for his condition. In sample 2 it is asked if the present treatment is the first addiction treatment.

4.4 Procedure for data collection

The data collection procedure is common to the two samples from this study, only with the difference in randomisation of the application of instruments - each sample had its method.

Patients were notified by letter to arrive at the hospital one hour before their psychology consultation in order to be conducted a pre-treatment evaluation at that time. Two or three days before each consultation, patients were reminded, by phone call, of their appointment. On the day of consultation, the evaluation assistant makes the first contact with the patient and leads to pre-treatment session. Informs the patient

that the first moment is an evaluation of clinical condition, and only after he/she will be directed to the psychology consultation. The research protocol is distributed to the patient and included in it a sociodemographic questionnaire and three instruments measure – PQ, PHQ-9 and CORE-OM. The fill order of these three instruments was randomized previously for all study participants; sociodemographic questionnaire is the first to be completed by the patient. The self-report instruments, preferably, were completed by the patients themselves. However, in case of difficulty (sight, illiteracy or other) the research assistant helped the completion, leading the filling in oral form. After completing the evaluation protocol, patient is informed about the study in progress and invited to join the investigation. It is described to the patient that the service is conducting a study on the development of new evaluation methods in order to meet the expectations and goals of treatment. The consent form, that patients have to sign if they accept to participate in the research, describes all this information, as well as the confidentiality aspects of the data provided. With this procedure finalized and the protocol filled correctly, the patient is sampled and follows soon after for the previously scheduled psychology consultation.

4.5 Data analysis procedure

Since PQ is an instrument with freetext responses (each item is a free response from the patient, accurately transcribed by the investigator), the data analysis procedure followed four steps: 1) Quality of the freetext items; 2) Freetext coding; 3) Matching and 4) Frequency Distributions and Descriptive Statistics.

4.5.1 Quality of the freetext items

Items are a transcription from patient's words, so it is importante to evaluate the quality of each one. For this, the Item Rating System (Elliott, 2012) was used, wich classifies each freetext item in:

1. Well-formed - Specific, personal difficulty that is reasonably a focus for psychotherapy;
2. Vague personal difficulties (e.g., relationships);
3. Goal (e.g., get along better with people);
4. General societal problems (e.g., general economic situation);

5. Other item quality issues (please describe).

In order to clarify the items which did not correspond to the first four criteria, other two were created to describe the fifth criteria (5.1 – Other-Multiple problems; 5.2 – Other-Past problems). To facilitate the analysis, the criteria were divided into two groups: items classified with “1” show good quality; items classified with “2” to “5” exhibit poor quality.

Each item was classified by two independent judges (Master degree students in Clinical Psychology at the University of Évora). Differences were discussed in order to reach agreement; when agreement didn't occur, a third judge (another Master degree student in Clinical Psychology at the University of Évora) was consulted.

4.5.2 Freetext coding

In order to compare problems indicated by the patients in PQ to the items of CORE-OM and PHQ-9 it was necessary to code the freetext answers based on their implicit theme. For this step, a classification system was used (Robinson, Ashworth, Sheperd, & Evans, 2006), which was also used in a study from Ashworth and partners in 2007. This system has 61 sub-themes (table 4); because some items did not fit into the sub-themes already constructed, we added 4 more (62, 63, 64 and 65) – the validation of the new sub-themes was verified by 3 independent judges (three master degree students in Clinical Psychology of the University of Évora – as in previous steps).

Table 4. *List of sub-themes*

1. Depression/Anxiety	2. Self image/self worth	3. Achievement
4. Work-related problems	5. Concentration	6. Moving on
7. Relationships – geral	8. Bereavement	8. Fears/panics
10. Relationships difficulties: family general	11. Sleep problems	12. Coping: general
13. Aggression/irritability	14. Relationships difficulties: partner – breaking up	15. Relationships difficulties: partner – development
16. Relationships difficulties: partner – general	17. Relaxing	18. Worries about health
19. Being happy	20. Socialising	21. Loneliness/being

		alone
22.Future	23.Having positive Outlook	24. Sexual problems
25.Somatic symptoms	26.Traumatic event	27. Relationships difficulties: family conflict
28.Self-acceptance	29.Relationship difficulties: partner – conflict	30. Addiction
31.Another person’s illness	32.Motivation	33. Relationships difficulties: family – breaking up
34.Victim of abuse/sexual violence	35.Coping: daily living	36. Money worries
37.Relationships difficulties: family – development	38.Understanding self/events	39. Making decisions
40.Relationships difficulties: family – worry about another	41.Communication	42.Emotions – unspecified
43.Going out/travelling	44.Guilt	45.Outlook on life
46.Dependence on other people	47.Having time	48. Housing worries
49.OCD (Obsessive-compulsive disorder)	50.Relationship difficulties: partner – forming	51. Avoiding issues
52.Coping: feelings	53.Eating problems	54.Personal development
55.Existence/existial	56.Global	57.Relationship difficulties: family – caring
58.Relationship difficulties: partner – worry about another	59.Suicidal thoughts	60.Thinking rationally
61.Thoughts	62.Attempted suicide	63.Self-harm
64.Academic-related problems	65.Justice-related problems	

4.5.3 Matching

The last step of qualitative analysis was to compare the content of PQ items with the items of CORE-OM and PHQ-9. Since PQ items were classified with sub-themes,

we needed to verify if these sub-themes map or not the items of the other two instruments. The classification used is described below:

1. Definite yes - When there is a direct and clear matching on the content of the item;
2. Possible yes - When the sub-theme reports a problem that could have much probably been caused by a problem reported on CORE-OM or PHQ-9;
3. Possible no - Vague sub-themes, or general, that might be or not associated to the CORE-OM or PHQ-9 items;
4. No - Different content, no clear matching.

A third judge was consulted when the independent judges could not reach a consensus on the classification – in that case, the original responses of PQ were compared with the CORE-OM or PHQ-9 items to obtain a more accurate classification.

4.5.4 Frequency Distributions and Descriptive Statistics

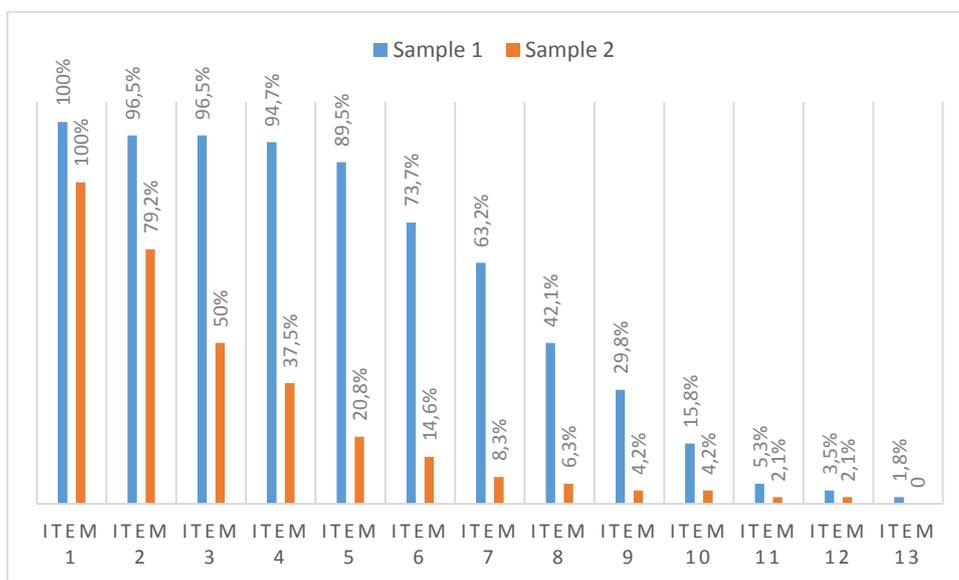
The latter process was quantitative analysis. At this stage we verify, through IBM SPSS Statistics 21 program:

- 1) the frequency of the sub-themes in the PQ items;
- 2) how many patients indicate each sub-theme in the PQ items;
- 3) how many patients indicate at least one non-mapped sub-theme in the CORE-OM and PHQ-9;
- 4) which sub-themes not mapped in the CORE-OM and in the PHQ-9 are more frequent.

5. Results

The total number of PQ items indicated by the 105 participants was 563 (sample 1= 406; sample 2= 157). The mean of items from total sample is 5.37 (SD=3.05), wherein the number of responses ranging from 1 to 13 (Figure 1). In the sample of drug abuse population, the number of items is significantly lower (M=3.29; SD=2.36) than in the psychiatric population sample (M=7.12; DP=2.39), which shows the significant differences between the two samples (U=344.5; W=1520.5; p=0.000).

Figure 1. Patient percentage of relative distribution according to the number of listed items



5.1 Quality of PQ items

During item quality analysis, it was necessary to specify one criteris: multiple problems and past problems, using 6 categories to rank the quality.

Quality items may be divided into two groups: well-formed and low quality - the latter encompasses Vague personal difficulties items, Goal, General societal problems, Other-multiple problems and Other-past problems.

Slightly more than half of the responses given by patients were considered Well-formed (n=291, 51.69%). The remaining presented quality issues, around 30% (n=174) of the items, describe vague personal difficulties (see Table 5).

Since the study is based on two samples, it would be interesting not only to analyze the sample in total, as well as individually – psychiatric population and drug abuse population.

Looking at the sample separately, the number of items of low quality is higher in sample 2 (n=114, 72.61%) and the opposite applies in relation to sample 1 - there is a higher number of good quality items (n=248, 61.08%), showing once again statistically significant differences ($X^2(2)=51.48$; $p=0.000$; $N=563$) (table 6).

Table 5. *PQ Item quality*

Item quality	Total sample n=105		Sample 1 n=57		Sample 2 n=48	
	n	%	n	%	n	%
Well-formed	291	51,69	248	61,08	43	27,39
Vague personal difficulties	174	30,91	101	24,88	73	46,50
Goal	11	1,95	4	0,99	7	4,46
General societal problems	6	1,07	3	0,74	3	1,91
Other - Multiple problems	57	10,12	31	7,64	26	16,56
Other - Past problems	24	4,26	19	4,68	5	3,18
	563	100	406	100	157	100

Table 6. *Item quality analysis*

Item quality	Sample 1	Sample 2	X^2	p	d.f
Well-formed	248	43	51,48	0.000	1
Low quality	158	114			

5.2 Sub-themes of PQ items

Four new sub-themes were found and added to the coding system used. In total, the items described 61 sub-themes, and the most common were: Addiction (n=46,

8.17%), Work-related problems (n=36, 6.39%), Depression/Anxiety (n=31, 5.51%), Loneliness/Being alone (n=27, 4.79%), Relationships difficulties: family - worry about another and Self-image/self-worth (n=25, 4.44%) (see figure 2).

In the psychiatric population, the most indicated sub-themes were "Family Difficulties - worry about another" (n = 27, 6.65%), "Depression / Anxiety" (n = 24, 5.91%), Fears/Panics (n=21, 5.17%) and Loneliness/being alone (n=21, 5.17%). "Addiction" was the most indicated by sample 2 (n=45, 28.66%), since the population is related to alcohol abuse and drug addition contexts (Figure 3).

Figure 2. Sub-themes mentioned in the PQ

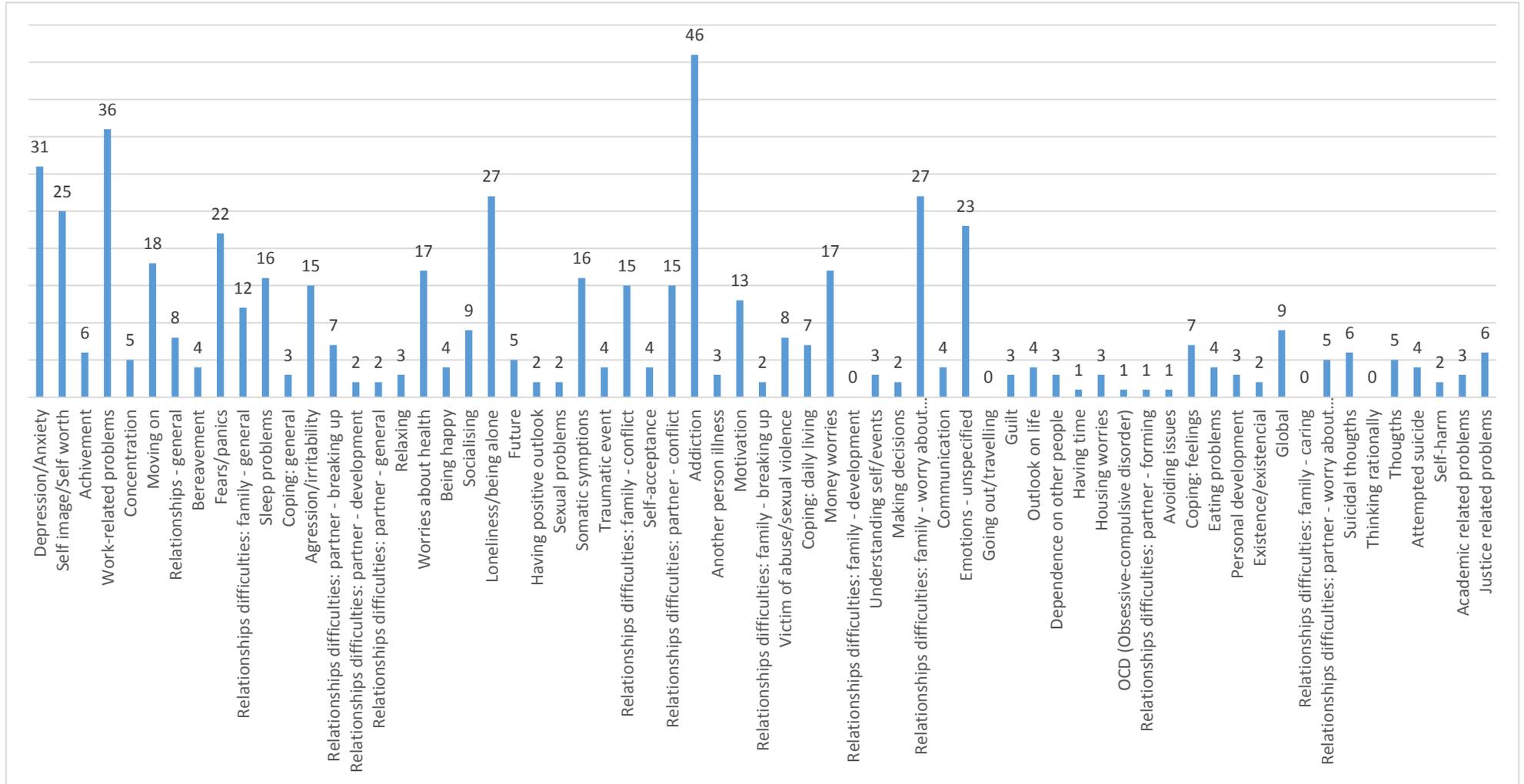
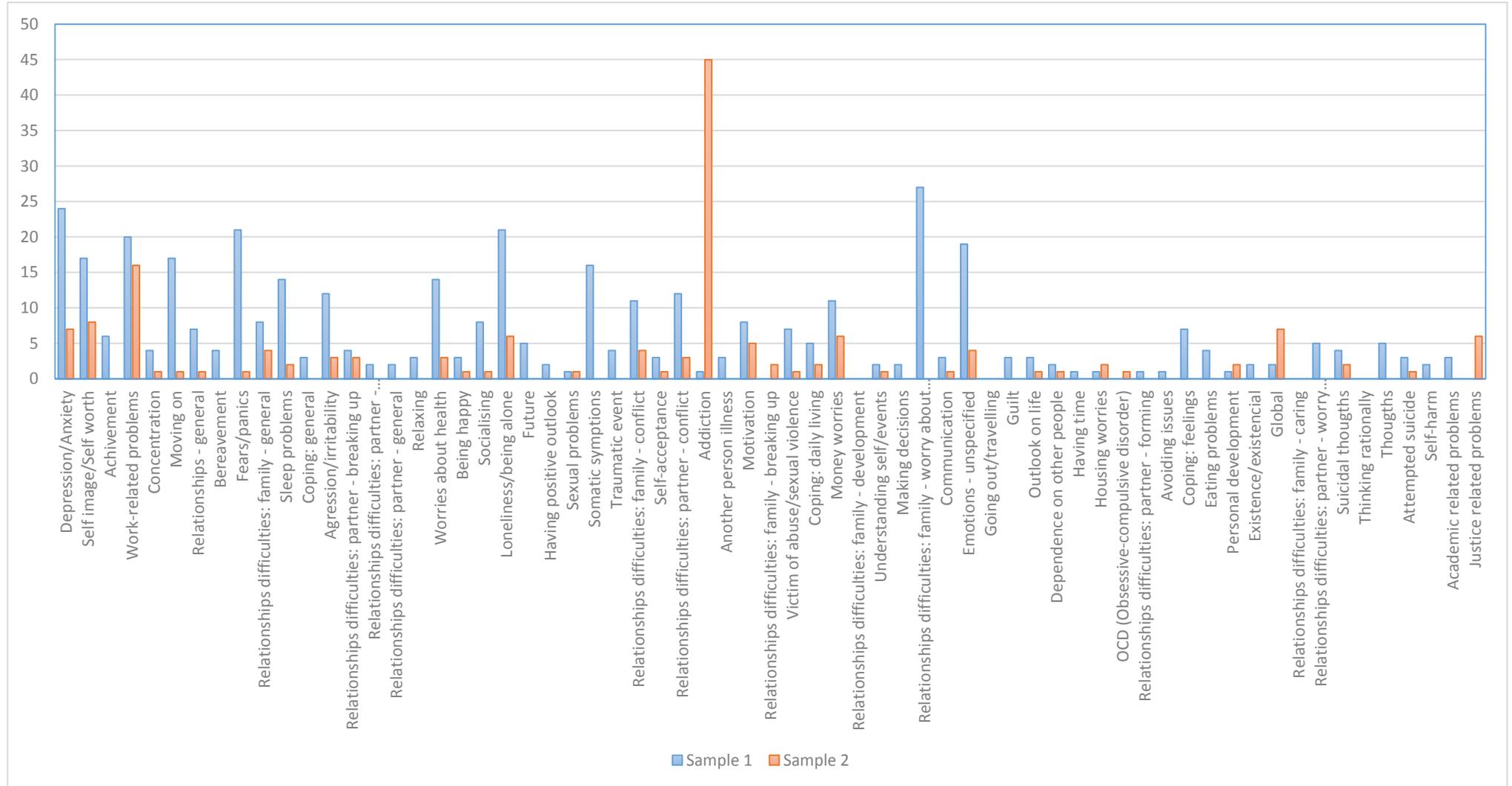


Figure 3. Sub-themes mentioned in the PQ, by each sample



5.3 Mapping of CORE-OM and PHQ-9

Of all the sub-themes indicated by patients on the PQ, 20 (37.78%) were not covered by CORE-OM. A larger number of sub-themes mentioned in PQ - 43 (70.50%) - were not mapped in PHQ-9 (see table 7).

Table 7. *Sub-themes not covered by CORE-OM and PHQ-9*

Sub-themes	CORE-OM	PHQ-9
Work-related problems	x	x
Moving on		x
Relationships – general		x
Fears/panics		x
Relationship difficulties: family – general		x
Coping: general		x
Aggression/irritability		x
Relationship difficulties: partner – breaking up		x
Relationship difficulties: partner – development		x
Relationship difficulties: partner – general		x
Worries about health;		x
Being happy		x
Socializing		x
Loneliness/being alone		x
Sexual problems	x	x
Traumatic event		x
Relationship difficulties: family – conflict		x
Relationship difficulties: partner – conflict		x
Addiction		x
Another person's illness	x	x
Relationship difficulties: family – breaking up		x
Victim of abuse/sexual violence	x	x
Coping: daily living		x
Money worries	x	x
Relationship difficulties: family – development		x
Understanding self/events	x	x
Making decisions	x	x
Relationship difficulties: family – worry about another	x	x
Going out/traveling	x	x

Guilt		X
Dependence on other people		X
Having time	X	X
Housing worries	X	X
OCD		X
Relationship difficulties: partner – forming		X
Avoiding issues	X	X
Coping: feelings		X
Eating problems	X	
Existence/existential	X	X
Global	X	X
Relationship difficulties: family – caring		X
Relationship difficulties: partner – worry about another	X	X
Academic-related problems	X	X
Justice-related problems	X	X
Concentration	X	
Bereavement	X	

5.4 Sub-themes not covered by CORE-OM and PHQ-9

A large part of sample population (n=72, 68.57%) reported at least one item that is not mapped by CORE-OM. The most indicated sub-themes not mapped in CORE-OM were: “Addiction” (n=46), “Work-related problems” (n=36) and “Relationships difficulties: family – worry about another” (n=25) (table 8).

Regarding PHQ-9, almost all sample (n=103, 97.14%) indicated at least one item that is not mapped on the instrument. “Addiction” (n=46), “Work-related problems” (n=36), “Loneliness/Being alone” (n=27) and “Relationships difficulties: family – worry about another” (n=25) were the sub-themes most reported that were not mapped in PHQ-9 (table 8).

Table 8. Number of responses and number of patients per sub-theme and its matching with the CORE-OM and PHQ-9

PQ sub-themes	Total			
	Total number of responses on PQ (n=563)	number of patients making each PQ response (n=105)	Matching with CORE-OM	Matching with PHQ-9
Addiction	46	30	Possible no	No
Work-related problems	36	28	No	No
Depression/Anxiety	31	23	Yes	Yes
Loneliness/being alone	27	22	Yes	No
Relationships difficulties: family - worry about another	27	17	No	No
Self image/Self worth	25	19	Yes	Yes
Emotions - unspecified	23	16	Possible yes	Possible yes
Fears/panics	22	14	Yes	No
Moving on	18	13	Possible yes	No
Worries about health	17	14	Yes	No
Sleep problems	17	14	Yes	Yes
Somatic symptoms	16	10	Yes	Yes
Agression/irritability	15	9	Yes	No
Relationships difficulties: family - conflict	15	9	Possible yes	No
Relationships difficulties: partner - conflict	15	14	Possible yes	No
Money worries	14	14	No	No
Motivation	13	10	Possible yes	Yes
Relationships difficulties: family - general	12	11	Possible yes	No
Socialising	9	8	Yes	No
Global	9	8	No	No
Relationships - general	8	8	Yes	No
Victim of abuse/sexual violence	8	5	No	No
Relationships difficulties: partner - breaking up	7	6	Possible yes	No
Coping: daily living	7	6	Yes	No
Coping: feelings	7	7	Possible yes	No
Achivement	6	6	Yes	Possible yes

Suicidal thoughts	6	4	Yes	Yes
Justice related problems	6	6	No	No
Concentration	5	4	No	Yes
Future	5	4	Yes	Possible yes
Relationships difficulties: partner - worry about another	5	4	No	No
Thoughts	5	3	Yes	Possible yes
Bereavement	4	3	No	No
Being happy	4	4	Yes	Possible yes
Traumatic event	4	3	Possible yes	No
Self-acceptance	4	4	Yes	Yes
Communication	4	4	Yes	Possible yes
Outlook on life	4	3	Yes	Possible yes
Eating problems	4	1	No	Yes
Attempted suicide	4	3	Yes	Possible yes
Coping: general	3	3	Yes	No
Relaxing	3	3	Yes	Yes
Another person illness	3	3	No	No
Understanding self/events	3	3	No	No
Guilt	3	3	Yes	No
Dependence on other people	3	3	Possible yes	No
Housing worries	3	2	No	No
Personal development	3	3	Possible no	Possible yes
Academic related problems	3	3	No	No
Relationships difficulties: partner - development	2	2	Possible no	No
Relationships difficulties: partner - general	2	1	Possible yes	No
Having positive outlook	2	2	Yes	Possible yes
Sexual problems	2	2	No	No
Relationships difficulties: family - breaking up	2	1	Possible yes	No
Making decisions	2	1	No	No
Existence/existencial	2	2	No	No
Self-harm	2	1	Yes	Possible yes
Having time	1	1	No	No
OCD (Obsessive-compulsive disorder)	1	1	Possible yes	No
Relationships difficulties: partner - forming	1	1	Possible no	No

Avoiding issues	1	1	No	No
Relationships difficulties: family - development	0	0	Possible no	No
Going out/travelling	0	0	No	No
Relationships difficulties: family - caring	0	0	Possible no	No
Thinking rationally	0	0	Possible yes	Possible yes

6. Discussion

This study compares PQ with CORE-OM and PHQ-9 in order to investigate which items are/aren't covered by the standardized instruments. Results show that there is a large number of patients that indicate sub-themes in PQ not covered by standardized instruments. This work comes to show its contribution to the further development of the characteristics of PQ. In the study of Elliott and colleagues (2016), psychometric characteristics were explored; now we come to investigate the characteristics and results of the instrument's application, since there are few studies that explore its use in clinical contexts.

The results show that 68.6% of the patients indicated at least one sub-theme in PQ that was not mapped in CORE-OM and 97.1% indicated at least one sub-theme that was not mapped in PHQ-9. This difference may be due to the fact that CORE-OM is an instrument consisting of 34 items and PHQ-9 only contemplates 9. Because PHQ-9 is a specific instrument for depression, items related to this theme are the only ones covered by the instrument. CORE-OM being an instrument that covers areas such as patient well-being, problems and symptoms, functioning and risk becomes more comprehensive, so it was expected that there would be a higher percentage of items covered.

The most mentioned theme in PQ, not mapped in the two standardized instruments, is "work-related problems". However, in the drug abuse sample the most indicated sub-theme is "Addiction", while in the clinical population sample this sub-theme is mentioned only once. Sample 1 had several highly mentioned sub-themes, representing the diversity of the population.

Regarding the quality of the items, the two samples show different quality levels: while in sample 1 most items have good quality, in sample 2 items are mostly vague. Sample 1 is comprised by clinical population and so it becomes a more diverse and representative sample of the general population; sample 2 being constituted by drug abuse population is just representative of that type of population, which due to their addiction problems may be affected cognitively (Rigoni, Oliveira & Andretta, 2006), thus showing the results of low-quality items.

Through the obtained results in this study, we can see that the use of this individual instrument brings benefits when compared to standardized instruments: the nature of the information enables great proximity to clinical reality, since it is able to capture the patients's individual problems to the fullest (Sales & Alves, in press). This characteristic

can not be demonstrated by CORE-OM and PHQ-9 since their answers are predetermined and do not meet the exact problems of the patients. In this study, this is demonstrated by the number of PQ items that these instruments do not cover. An assessment made only with CORE-OM or PHQ-9 would not allow the understanding of what really worries the patients because some of their concerns would never be mentioned in its items.

“Addiction”, “work-related problems”, “relationships difficulties: family – worry about another” and “loneliness” were the most indicated sub-themes that were not mentioned in CORE-OM and PHQ-9 items – here we can move towards another advantage of PGOMs seen through PQ: it allow the patient to indicate the problems that led to therapy and that can be changed with the course of treatment (Sales e Alves, in press). Selecting predetermined items on CORE-OM and PHQ-9, allows patients to indicate some of the problems that are affecting their clinical condition; however, these may not be problems that they really want to work in therapy. Thus, PQ becomes an instrument that will better meet the patients’s needs, taking into account their real problems and feelings.

When comparing the results of this study to the work that it was based on (Ashworth, Robinson, Shepherd, Conolly, & Rowlands, 2007) we can see some similarities. Ashworth’s and colleagues study (2007) showed that CORE-OM was unable to map 60% of the items listed in PSYCHLOPS (individualized measured in the study) and that the topics most indicated in PSYCHLOPS not covered by the CORE-OM were work-related problems and relationships issues. These data are in line with the results we’ve obtained, showing relatively similar values and both coinciding in the not mapped themes.

In a similar work from Neves, Sales and Ashworth (2015), our results become even closer. In a comparison between PSYCHLOPS and the two standardized measures used in this study, 73.8% of individualized measure items were not mapped by CORE-OM and 96.2% of the items were not covered by PHQ-9. The most indicated subthemes that were not represented in the standardized instruments were work-related problems, relational issues, money worries and addiction. The similarity between the two studies is due to the particularity of both using the same sample. However, the application of both individualized instruments is different, since PQ is made in an interview and PSYCHLOPS dares the patients to complete a questionnaire on paper by themselves. Thus, the variations found between the two studies may be related to differences in administration measures, which can influence the quality and

accuracy of data that is retrieved. Whatever the method of application may be, there will always be influences in the answers, even if for minor differences in the choice of words, order of questions or the answers format (Bowling, Bond, Jenkinson, & Lamping, 1999; Bowling, 2005).

Thereby, the results found in our study are compatible with others who also explored the application characteristics of individualized measures. Although the PGOMs are different, the conclusions that can be drawn from this study confirm that the use of these measures becomes beneficial in relation to nomothetic methods. This paper adds that the particular use of PQ proves essential in reaching the patients's real problems and their application in clinical practice becomes quite useful.

Limitation and future studies

The limitation of this study is the use of two completely different samples. There are disparities between the number of patients, number of items indicated in PQ and quality of the items. These differences do not allow a balanced sample population, nor a fair comparison for the two samples.

In future studies it might be worthwhile to continue the use of PQ applying it during and at the end of treatment in order to validate its contribution to the patients' change during the therapy process.

7. References

- Alves, P., Sales, C. M. D., & Ashworth, M. (2013). Enhancing the patient involvement in outcomes: a study protocol of personalised outcome measurement in the treatment of substance misuse. *BMC Psychiatry*, 13, 337-349. doi: 10.1186/1471-244X-13-337.
- Alves, P., Sales, C. M. D., & Santos, T. (2014). *Involving patients in health care provision*. Paper presented in SPR 2014 Conference. Copenhagen.
- Ashworth, M., Evans, C., & Clement, S. (2009). Measuring psychological outcomes after cognitive behaviour therapy in primary care: A comparison between a new patient-generated measure "PSYCHLOPS" (Psychological Outcome Profiles) and "HADS" (Hospital Anxiety and Depression Scale). *Journal of Mental Health*, 18, 169-177. doi: 10.1080/09638230701879144.
- Ashworth, M., Robinson, S., Evans, C., Shepherd, M., Conolly, A., & Rowlands, G. (2007). What does an idiographic measure (PSYCHLOPS) tell us about the spectrum of psychological issues and scores on a nomothetic measure (CORE-OM)? *Primary Care & Community*, 12, 7-16. doi: 10.1080/17468840701560805.
- Ashworth, M., Shepherd, M., Christey, J., Matthews, V., Wright, K., Parmentier, H., Robinson, S., & Godfrey, E. (2004). A client-centred psychometric instrument: the development of PSYCHLOPS. *Counselling and Psychotherapy Research*, 4, 27-33. doi: 10.1080/14733140412331383913.
- Barkham, M., Margison, F., Leach, C., Lucock, M., Mellor-Clark, J., Evans, C., Benson, L., Connell, J., Audin, K., & McGrath, G. (2001) Service profiling and outcomes benchmarking using the CORE-OM: toward practice-based evidence in the psychological therapies. *Journal of Consulting and Clinical Psychology*, 69, 184-196.

- Beck, A.T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961) Na inventory for measuring depression. *Archives of General Psychiatry* 4, 561-571.
- Bergner, M., & Rothman, M. L. (1987). Health status measures: na overview and guide for selection. *Annual review of public health*, 8, 191-210. doi: 10.1146/annurev.pu.08.050187.001203.
- Bowling, A. (2005). Mode of questionnaire administration can have serious effects on data quality. *Journal of public health*, 27, 281-291. doi:10.1093/pubmed/fdi031.
- Bowling, A., Bond, M., Jenkinson, C., & Lamping, D. L. (1999). Short Form 36 (SF-36) Health Survey questionnaire: which normative data should be used? Comparisons between the norms provided by the Omnibus Survey in Britain, the Health Survey for England and the Oxford Healthy Life Survey. *Journal of Public Health*, 21, 255-270. doi: 10.1093/pubmed/21.3.255.
- Clark, A. , Hook, J. , & Stein, K. (1997). Counselors in primary care in Southampton: A questionnaire survey of their qualifications, working arrangement, and casemix. *British Journal of General Practice*, 47, 613–617.
- Crawford, M. J., Rutter, D., Manley, C., Weaver, T., Bhui, K., Fulop, N., & Tyrer, P. (2002). Systematic review of involving patients in the planning and development of health care. *British Medical Journal*, 325, 1-5. doi: 10.1136/bmj.325.7375.1263.
- Elliott, R. (2010). Psychotherapy change process research: Realizing the promise. *Psychotherapy Research*, 20, 123-135. doi: 10.1080/10503300903470743.
- Elliott, R. (2012). *Quality rating system for PQ items*. Unpublished manuscript. Counselling Unit, University of Strathclyde.

- Elliott, R., Mack, C., & Shapiro, D. (1999). *Simplified Personal Questionnaire Procedure* [Unpublished manuscript]. University of Toledo: Toledo, Ohio. Retrieved from <http://www.experiential-researchers.org/instruments/elliott/pqprocedure.html>.
- Elliott, R., Wagner, J., Sales, C. M. D., Rodgers, B., Alves, P., & Na, M. J. (2016). Psychometrics of the personal questionnaire: a client-generated outcome measure. *Psychological Assessment, 28*, 263-278. doi:10.1037/pas0000174.
- Evans, C., Margison, F., & Barkham, M. (1998). The contribution of reliable and clinically significant change methods to evidence-based mental health. *Evidence Based Mental Health, 70-72*. doi:10.1136/ebmh.1.3.70.
- Evans, C., Mellor-Clark, J., Margison, F., Barkham, M., Audin, K., Connell, J., & McGrath, G. (2000). CORE: Clinical Outcomes in Routine Evaluation. *Journal of Mental Health, 9*, 247–255. doi:10.1080/jmh.9.3.247.255.
- Fitzpatrick, R., Davey, C., Buxton, M. J., Jones, D. R. (1998). Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment, 2*, 1-86. doi: 10.3310/hta2140.
- Hedinsson, H., Kristjansdottir, H., Olason, D., Sigurdsson, J. F. (2013). A validation and replication study of the patient-generated measure 'PSYCHLOPS' (Psychological Outcome Profiles) on an Icelandic clinical population. *European Journal of Psychological Assessment, 29*, 89-95. doi: 10.1027/1015-5759/a000136.
- Hunsley, J., & Mash, E. (2007). Evidence-based Assessment. *Annual Review of Clinical Psychology, 329-351*.

- Kazis, L. E., Anderson, J. J., & Meenan, R. F. (1989). Effect sizes for interpreting changes in health status. *Medical care*, 27, 178-189.
- Kiresuk, T., & Sherman, R. (1968). Goal attainment scaling: a general method of evaluating comprehensive mental health programmes. *Community Mental Health*, 4, 443-453. doi: 10.1007/BF01530764.
- Kirshner, B., & Guyatt, G. (1985). A methodological framework for assessing health indices. *Journal of chronic diseases*, 38, 27-36. doi:10.1016/0021-9681(85)90005-0.
- Kroenke, K., & Spitzer, R. L. (2002). The PHQ-9: a new depression diagnostic and severity measure. *Psychiatric Annals*, 32, 1-7.
- Kroenke, K., Spitzer, R. L., & Williams, J. B. (2001). The Phq-9. *Journal of general internal medicine*, 16, 606-613. doi: 10.1046/j.1525-1497.2001.016009606.x.
- Lambert, M., & McRoberts, C. (1993). *Outcome measurement in JCCP:1986-1991*. Poster presented in Annual Meetings of the Western Psychological Association. Phoenix, Arizona.
- McPherson, F.M. & LeGassicke, J. (1965). A single-patient, self-controlled and self-recorded trial for Wy 3498. *British Journal of Psychiatry*, 111, 149-154.
- Monteiro, S., Torres, A., Pereira, A., Albuquerque, E., & Morgadinho, R. (2013). 2077– Preliminary validation study of a portuguese version of the patient health questionnaire (PHQ-9). *European Psychiatry*, 28, 1. doi: 10.1016/S0924-9338(13)76982-7.
- Nacional Institute for Health and Care Excellence. (2016). *Patient experience in adult NHS services*. Retrived from <https://www.nice.org.uk/>.

- Neves, I., Sales, C. M. D., & Ashworth, M. (2015). *What do we gain when we listen to the patients? Comparison of PSYCHLOPS with two nomothetic measures*. Masters dissertation. Universidade de Évora, Évora.
- Ogles, B. M. (2013). Measuring change in psychotherapy research. In M.J. Lambert (Ed.), *Bergin & Garfield's Handbook of psychotherapy research and behavior change* (6th ed.). New York: Wiley.
- Overington, L. & Ionita, G. (2012). Progress monitoring measures: A brief guide. *Canadian Psychology*, 53, 82–92.
- Philips, J. (1986). Shapiro personal questionnaire and generalized personal questionnaire technique: A repeated measures individualized outcome measurement. In L. Greenberg, & W. Pinsof, *The psychotherapeutic process: A research handbook* (557-589). New York: Guildford.
- Rigoni, M. D. S., da Silva Oliveira, M., & Andretta, I. (2006). Conseqüências neuropsicológicas do uso da maconha em adolescentes e adultos jovens. *Ciências & Cognição*, 8, 118-126.
- Robinson, S., Ashworth, M., Shepherd, M., & Evans, C. (2006). In their own words: a narrative-based classification of clients' problems on an idiographic outcome measure for talking therapy in primary care. *Primary Care and Mental Health*, 4, 165-173. doi:10.1080/09638230701879144.
- Sales, C. M. D., & Alves, P. C. G. (2012). Individualised Patient-Progress Systems: Why We Need to Move Towards a Personalized Evaluation of Psychological Treatments. *Canadian Psychology*, 53, 115-121. doi: 10.1037/a0028053.

- Sales, C. M. D., & Alves, P. C. G. (in press). Patient centred assessment in Psychotherapy: A review of individualized assessment tools. *Clinical Psychology: Science and Practice*.
- Sales, C., Gonçalves, S., Fragoeiro, A., Noronha, S., & Elliott, R. (2007). Psychotherapists openness to routine naturalistic idiographic research. *Mental Health and Learning Disabilities Research and Practice*, 4, 145–161. 10.5920/mhldrp.2007.42145.
- Sales, C. M. D., Moleiro, C., Evans, C., & Alves, P. C. G. (2012). Versão Portuguesa do CORE-OM: Tradução, adaptação e estudo preliminar das suas propriedades psicométricas. *Revista de Psiquiatria Clínica*, 39, 54-59.
- Shapiro, M.B. (1961). A method of measuring changes specific to the individual psychiatric patient. *British Journal of Medical Psychology*, 34, 151-155. doi: 10.1111/j.2044-8341.1961.tb00940.x.
- Shapiro, M. B. (1969). Short-term Improvements in the Symptoms of Affective Disorder. *British Journal of Social and Clinical Psychology*, 8, 187-188. doi: 10.1111/j.2044-8260.1969.tb00606.x.
- Spitzer, R. L., Kroenke, K., Williams, J. B., & Patient Health Questionnaire Primary Care Study Group. (1999). Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. *Jama*, 282, 1737-1744. doi: 10.1001/jama.
- Streiner, D. L., & Norman, G. R. (2003). *Health Measurement Scales: A Practical Guide to Their Development and Use*. Oxford: Oxford University Press.
- Turner-Stokes, L. (2011). Goal attainment scaling its relationship with standardized outcome measures: A commentary. *Journal of Rehabilitation Medicine*, 70-72. doi:10.2340/16501977-0656.

Wagner, J., & Elliott, R. (2001). The simplified personal questionnaire. Unpublished manuscript, University of Toledo, Department of Psychology.

Ware, J. (1984). Methodological considerations in the selection of health status assessment procedures. In N. Wenger, C. Mattson, C. Furberg, & J. Elinson, *Assessment of quality of life in clinical trials of cardio-vascular therapies* (87-111). New York: LeJacq Publishing Inc.

Zigmond, A. S., & Snaith, R. P. (1983). The Hospital Anxiety and Depression Scale. *Acta Psychiatrica Scandinavica*, 67, 361-370. doi: 10.1111/j.1600-0447.1983.tb09716.x.