



Effectiveness and Efficiency of Strategic and Systemic Therapy in Naturalistic Settings: Preliminary Results from a Systemic Practice Research Network (SYPRENE)

Grégoire Vitry^{a,b,c}, Rytis Pakrosnis,^d Olivier Gabriel Brosseau^e and Nathalie Duriez^c

SYPRENE, an international digital systemic practice research network (PRN), was established in 2014 to fill the gap in practice-based research on the effectiveness and efficiency of systemic therapies, starting with strategic therapies. This article reports initial outcome results. So far, twenty-seven therapists from several countries using the MRI brief therapy strategic approach have collected data for a wide range of problem diagnoses. Data are reported here for 1,150 completed cases. The drop-out rate was 19%. Significant improvement or complete problem resolution was reported by therapists in 80% and by patients in 90% of cases, with an average of 5.4 sessions and 5.3 months of treatment. Patients reported significant improvement on the GHQ-12 with a large effect size, and 76% of cases were evaluated as reaching reliable and clinically significant change.

Practitioner points

1. SYPRENE is a digital practice research network
2. Outcome data have been collected on 1,079 cases.
3. Evidence supports the effectiveness of strategic therapy.

Keywords: brief therapy; MRI; outcome research; research methodology; strategic therapy; systemic consultation.

Thanks to Michael Hoyt, Xavier Briffault, Aurélien Baelde and Scott Miller for consultation.

Strategic therapy was inspired by the work of Milton Erickson (1980; Haley, 1973) and Gregory Bateson (1972, 1979) and flourished in the 1980s,

^a Director LACT, Paris, France

^b University Paris Descartes, CERMES3, Paris, France

^c Laboratoire de Psychopathologie et Neuropsychologie, University Paris 8, Saint-Denis, France

^d Vytautas Magnus University, Kaunas, Lithuania

^e LACT, Paris, France

initially at the Mental Research Institute (MRI) in Palo Alto, California (Fisch *et al.*, 1982; Watzlawick *et al.*, 1974), and then later in Washington, DC (e.g. Haley, 1977, Richeport and Carlson, 2012; Madanes, 1981, 1984) and in Milan (e.g. Boscolo *et al.*, 1987). Working with patients from an interactional-systemic perspective that favours short-term problem-solving strategies, this therapeutic approach has spread around the world and has taken many different forms and variations (Hoyt, 2019; Nichols, 2003). Although MRI strategic-interactional therapy has ‘stood the test of time’ (Hale and Frusha, 2016) and the general conclusion can be made that strategic therapy works (Horigian *et al.*, 2015, 2016; Jackson *et al.*, 2018; Nardone and Portelli, 2005; Pietrabissa *et al.*, 2016; Wittezaele and Nardone, 2016), the limited transferability of laboratory-based efficacy research into everyday clinical practice is often stressed in the literature (Baucom *et al.*, 2018; Lebow, 2018; Tasca *et al.*, 2015; Teachman *et al.*, 2012). Looking for solutions to fill the gap between laboratory and naturalistic psychotherapy outcome studies, practice research networks (PRNs) were introduced in the 1990s in the hope that large numbers of cases coming from real-life clinical practice would allow obtaining specific information valid for everyday practice and greater generalisability of results and conclusions (Johnson *et al.*, 2017; Thurin *et al.*, 2012).

The Systemic Practice Research Network (SYPRENE), created by LACT¹ Research in 2014, is a PRN of professionals, including therapists from several international institutes (Vitry *et al.*, 2020). Following a ‘real-world’ research perspective (Seligman, 1995), SYPRENE’s main goal is to fill the research gap on the effectiveness and efficiency of strategic therapy by developing a new therapy data recording and assessment platform that would create a database and allow communication between and amongst practitioners and researchers. SYPRENE aims to use a minimum number of general measures to reduce costs and time demands, simplify assessment and training and allow comparisons between different cases and techniques. Effectiveness characterises the measure of a treatment performance under natural clinical conditions. Efficacy deals with the measure of a treatment performance under ideal laboratory conditions such as randomised controlled trials (RCTs) used by evidence-based medicine. Efficiency is the ratio between resources and effectiveness.

¹LACT is an acronym signifying ‘Liberating Actions for Transformation’. LACT is a free-standing research, training and psychotherapy institute based in Paris. It is supported by tax credits from the French government. SYPRENE was originally called ‘LACT Research’; we changed the name to SYPRENE in 2018.

SYPRENE's purpose and goals, the process of development and data collecting/recording system were described in detail in a previous article (Vitry *et al.*, 2020). The present article introduces preliminary outcome results of the SYPRENE PRN. It explains the methodology for evaluating the problem-solving outcome from both the therapist's and the patient's points of view. Outcome results, combining patients' and practitioners' perspectives, are presented. Finally, a link between problem-solving outcomes and changes in patients' psychosocial functioning is analysed, informing the discussion about the validity of different methods of outcome evaluation and the convergence between outcomes measured by different means and from different perspectives.

Methods

SYPRENE Digital Data Collection and Recording System

Data in SYPRENE are collected from both practitioners and patients, and then entered digitally into the data-recording system by the practitioners (Vitry *et al.*, 2020). Practitioners all undergo an initial 90-min training in the use of SYPRENE, and then engage in a monthly online conference to discuss cases and assessments.

Standardised information for each patient's case is recorded:

- date, time, duration and setting of a session;
- problem diagnosis, level of potential risk;
- type of session (first appointment, continuing and last appointment);
- patient's evaluation of the therapeutic alliance (Session Rating Scale; SRS, Miller and Duncan, 2004 – discussed below) scores obtained after each session;
- stage of the therapy [e.g. defining the objective of the therapy; unblocking (problem attempted solution); consolidation (when problem is solved); follow-up at 6 months; type of ending (completed or drop-out)];
- patient's evaluation of the outcome: (1) General Health Questionnaire (GHQ-12 – discussed below) scores obtained before the first, fourth and after the last session; (2) Outcome Rating Scale (ORS; Miller and Duncan, 2004) for each session²; (3) patient's evaluation of the problem-solving

²Looking for the best solutions, different outcome measures were added to the SYPRENE data recording system at different times. The ORS and SRS (Miller and Duncan, 2004) as well as the patient's evaluation of the problem-solving outcome and the evaluation of change were added in the last trimester of 2019. Thus, only measures with a substantial number of cases, allowing reliable statistical analysis, were used in the present article (see Outcome measures).

- outcome (on an 11-point scale) at the end of the therapy; (4) patient's evaluation of change (to what extent — not at all, little, mostly or completely) before the fourth and after the last session;
- practitioners' evaluation of the problem-solving outcome (on an 11-point scale) at the end of the therapy).

Outcome measures

General Health Questionnaire (GHQ-12) (Goldberg and Williams, 1991) is a self-administered questionnaire composed of twelve questions, developed as a unidimensional scale for measuring 'mental distress'. Answers are scored on a 5-point scale (0–4) to quantify the level of subjective psychological suffering. Thus, scores can range between 0 and 48. The GHQ-12 provides a general health score and is widely used to measure change in psychotherapy research and everyday practice; it is quick and easy to complete by a patient; it can be used for different situations and problems; and it has been translated into a number of languages and possesses good validity and reliability characteristics, with consistency ranging between .76 and .94 (Lesage *et al.*, 2011; Sánchez-López and Dresch, 2008; Werneke *et al.*, 2000).

Practitioners' and patients' evaluation of the problem-solving outcome

To what extent the patient's presenting problem had been solved during the therapy in each case was assessed by the therapist and patient at the end of the therapy on a 0–10-point scale (from 0 = problem unsolved to 10 = problem solved).³

These assessments are closely related and integrated into the strategic therapy process. At the beginning of the therapy, a practitioner evaluates the contextual problem that brings a patient to therapy, the redundant attempted solution (Vitry *et al.*, in press) that keeps a patient stuck in the situation and the minimum needed to resolve the problem. Next, a practitioner works on the different levels of intervention (strategy, communication and relationship) using direct prescriptions or indirect suggestions. From the second session onwards, the practitioner asks at the beginning of each session if the patient has done the prescriptions (behaviours, thoughts or other homework) they

³The problem-solving scale originated with Molnar and de Shazer (1987) and was originally called the 'inverted scale'.

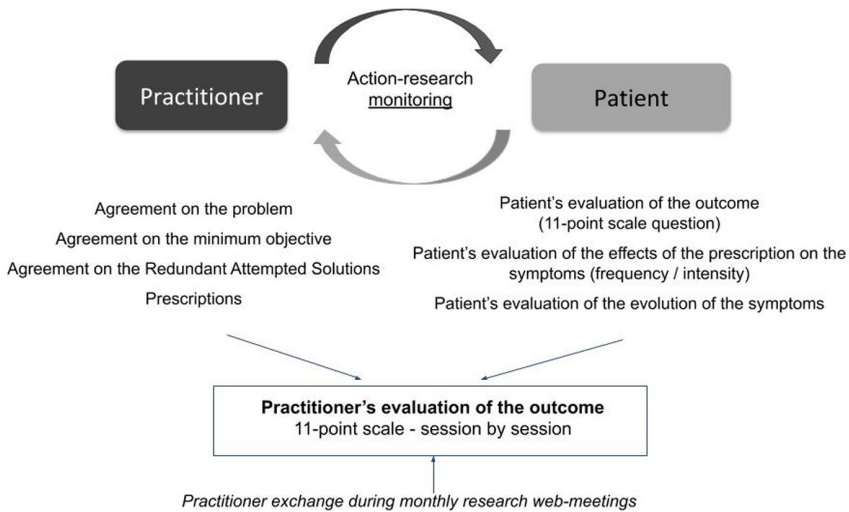


Figure 1. Interactional scheme of a practitioner's and patient's evaluation of the outcome for each session.

have agreed upon. Once it is determined that the prescriptions have been implemented, the practitioner asks the patient if this has had an effect on symptoms in terms of reducing frequency and/or intensity. This iterative process (Figure 1) allows a practitioner to assess the patient's problem-solving outcome from 0 to 10 (from problem unsolved to problem solved) session by session.

When the patient's unsuccessful attempted solutions are stopped and the therapy process comes to an end, the practitioner asks the patient to rate the problem-solving outcome on a 0–10-point scale (from problem unsolved to problem solved). At times, practitioners may have different evaluations from those of patients if a therapist considers, for instance, that a patient underestimates his or her own progress (which may be part of the patient's difficulty, e.g. for a perfectionistic patient) or overestimates his or her progress (e.g. for someone who tends to be overly optimistic or exaggerates their accomplishments).

For the outcome analysis, practitioners' and patients' assessments were *prima facie* and somewhat arbitrarily categorised as

1. Problem solved (successful intervention or improved substantially) if a practitioner/patient evaluated the extent to which a problem had been

- solved as 7 or more (indicating substantial or complete resolution of the problem at the end of therapy);
2. Problem only partially solved (improved somewhat or intervention with moderate improvement) if a practitioner/patient evaluated the extent of problem resolution as 4–6 (indicating partial remission of symptoms at the end of therapy);
 3. Problem not solved (failed intervention or little improved) if a practitioner/patient evaluated the resolution of the problem as 3 or less (indicating that the treatment was not effective).

Participants

Practitioners. Analysis and results are based on data from the twenty-seven active members of SYPRENE⁴ (those who have provided ratings for at least one case), of whom fifteen were female and twelve were male. Their average age was 49 years ($SD = 5.4$), ranging between 37 and 60 years. The nationalities of the SYPRENE practitioners are as follows: fourteen French, six Italian, two Mexican, one Spanish, one Belgian, one Swiss, one Canadian and one South African. Practitioners reported having an average of 16 years of experience as psychologists/psychotherapists ($n = 11$), therapists/counsellors ($n = 16$). All practitioners identified themselves as ‘strategic therapists’, and 18% ($n = 4$) also specified Ericksonian hypnotherapy as a secondary specialisation.

Consistent with the theory of redundant attempted solutions (RAS; Vitry *et al.*, in press), from which the MRI strategic approach was derived, when people try to solve problems, they often apply the very strategy that perpetuates the problems they seek to solve. Disrupting RAS can resolve ‘stuckness’ and thus allow change to occur. In MRI-style strategic-systemic therapy (Watzlawick *et al.*, 1974; Fisch *et al.*, 1982; Nardone and Portelli, 2005; Nardone and Watzlawick, 1993, 2005), this is accomplished as the practitioner and client continuously provide one another with feedback and make adjustments.

Patients. Since the beginning of SYPRENE data collection, data for 2,027 patients seen individually have been recorded from the twenty-seven active therapists (data for the current article were retrieved from the system on 3 April 2020), including the following (Figure 2): (1) 1,150 patients (57%) with a final outcome evaluation on at least one

⁴Ninety therapists opened an account to learn more about SYPRENE but did not contribute data.

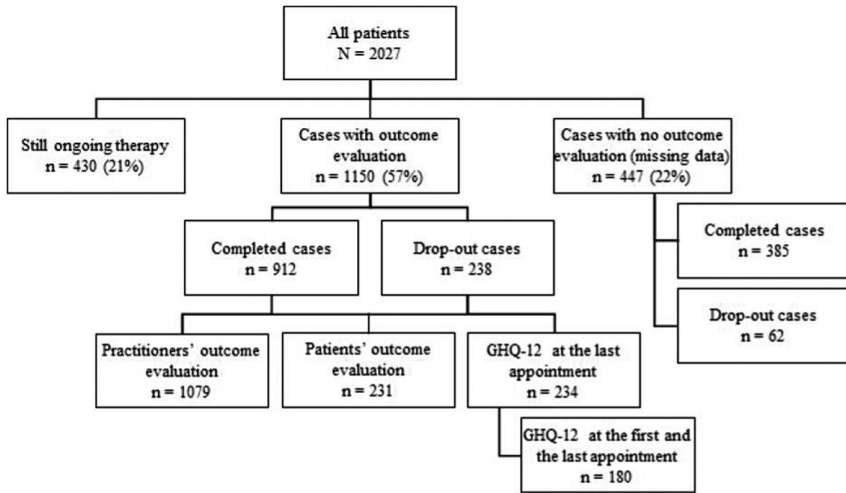


Figure 2. Flow of patient enrolment and evaluation.

measure (912 of them completed therapy with a closing meeting and 238 terminated therapy early, thus were considered as drop-outs); (2) 447 patients (22%) without an outcome evaluation (385 of whom were 'visitors' merely seeking information or for whom therapists did not collect data necessary for the outcome evaluation, and 62 who dropped out early and thus were not assessed at the last meeting); and (3) 430 patients (21%) who were still undergoing therapy at the moment of data retrieval from the system. Therefore, out of 1,597 patients for whom the therapy was finished at the moment of data retrieval, 300 (19%) were considered drop-outs (terminated early or did not come to an assigned appointment).

Thus, the general sample for the outcome analysis was 1,150 patients with outcome data on at least one of three measures. Of those, 1,079 patients had a practitioner's problem-solving outcome evaluation; in 231 cases problem-solving outcome was evaluated by the patients and in 224 cases⁵ the GHQ-12 was filled in at the last session

⁵Note: the sum of provided numbers does not correspond to the general number of patients with outcome evaluation ($n = 1,150$) because for some patients the outcome was assessed on all measures, whereas for some, on two or only one measure, which again could vary in different cases. Thus, depending on the measure, different sample sizes will be used when analysing outcome data.

(180 of these cases had the GHQ-12 at the first and the last sessions, allowing comparison of these scores to assess the magnitude of change on this measure).

Out of 1,150 patients with outcome evaluation, 64% were self-identified female and 36% were self-identified male. Marital status (reported by 65% of patients) was 48% living alone (single, separated, widowed) and 52% living as a couple (married or in a union). The average age (indicated by 70% of patients) was 39.7 ($SD = 13.6$) years; 14% were under 25 years old, 51% between 25 and 44 years old, 31% between 45 and 64 years old and 4% were over 65 years old. Occupationally, 67% were employed and 33% were inactive (e.g. without professions, job seeking or retired).

Missing data analysis. As seen in Figure 2, outcome data for 447 patients (28% of 1,597 cases for whom therapy was finished in one or another way) were not collected. Evaluation by the therapist was introduced as soon as the database was created (2014), GHQ-12s were fully implemented in 2018 and evaluation by the patient fully implemented in 2019. It can be observed that 2018 was the pivotal year from which automated procedures were introduced that encourage practitioners to fill in the necessary data at the time of file closure. These pop-up procedures remind the practitioner that it is necessary to enter certain data. Practitioners are gradually reducing their rate of missing data.

The assumption that missing data occurred due to organisational and accidental reasons and that there is no systematic bias for missing data can be partially supported by the fact that no differences between patient groups with and without outcome data were found in regard to gender, age, relationship status or initial practitioners' and patients' problem severity evaluations as well as initial GHQ-12 scores.

Results

Statistical analysis was conducted using SPSS-21 software. For the 1,150 cases completed with outcome assessment, the average number of sessions was 5.4 ($SD = 4.4$), ranging from 1 to 35 sessions. In 10% of cases, intervention was completed after a single session; in 29% of cases, after two to three sessions; in 23% of cases, after four to five sessions; in 15% of cases, after six to seven sessions; and in 8% of cases, after eight to nine sessions. Only 11% of patients needed between ten and fifteen sessions, and another 4% needed more than fifteen sessions. The average

(mean) duration of treatment for all 1,150 patients with outcome data was 5.3 months ($SD = 5.8$).

Table 1 presents a summary of the types of problem diagnoses for which the clients came to therapy. For the sake of brevity, only problem categories included in ten or more cases are presented. The items in the table include interactional diagnoses (de Scorraille *et al.*, 2017; Vitry *et al.*, 2019; Watzlawick, 2009) such as relational problems at place of work, relational problems with mistrust of others or relational problems within the couple. Data in Table 1 indicate that interpersonal conflict, problems of self-esteem, mistrust and anxiety-related issues were the problems most frequently reported by patients. It is also notable that, for most problem categories, the success rate ranged between 70% and 95%. The highest success rates (over 90%) were for perfectionism, phobias and problems related to change in the workplace, while the lowest success rates (59%) were for issues related to borderline personality.

Problem-solving outcome results based on practitioners' and patients' evaluations

The SYPRENE data recording system requires that, in each case after the last appointment, the degree of problem improvement is evaluated on an 11-point scale (0–10). The average evaluation by practitioners at the last appointment was 6.0 ($SD = 2.8$, $n = 1079$) and ranged from 0 to 10. Patients' evaluations on the same scale also ranged from 0 to 10, but the average score was somewhat higher: 7.1 ($SD = 2.6$, $n = 231$). Having in mind sample size differences, we calculated averages only for those cases where problem-solving outcome was evaluated by both practitioners and patients ($n = 219$). In these instances, the practitioners' average evaluation was 7.3 ($SD = 2.3$), and the patients' was 7.1 ($SD = 2.6$).

Furthermore, practitioners' and patients' evaluations were separated into three categories, reflecting the success rates of intervention (see Outcome Measures section in Methods). Thus, out of the total 1,079 patients who had practitioners' problem-solving outcome evaluations, in 53% of cases problems were evaluated by practitioners as solved or improved substantially (rated 7–10), 27% as only partially solved or improved somewhat (rated 4–6) and 20% as unsolved or little improved (rated 0–3). Out of 231 patients who evaluated their problem-solving outcome on a 0–10-point scale at the last session, in 70% of cases problems were evaluated as solved or improved substantially, 20% as only partially solved or improved somewhat and 10% as unsolved or little improved [among these, 4% ($n = 9$) of cases were evaluated by patients as 0].

TABLE 1 *Frequencies of patients' problem diagnoses and practitioners' evaluation of success rates for each problem diagnosis category (n = 1,150)*

| Category of patients' problem diagnoses | Number of cases | Failed intervention | Intervention with some improvement | Successful intervention | Success rate % (complete and partial) |
|---|-----------------|---------------------|------------------------------------|-------------------------|---------------------------------------|
| Problems of self-esteem | 117 | 20 | 34 | 63 | 97 (83%) |
| Relational problem with mistrust paranoia | 86 | 13 | 27 | 46 | 73 (85%) |
| Pathological doubt | 82 | 9 | 22 | 51 | 73 (89%) |
| Relational problems at place of work | 72 | 5 | 28 | 39 | 67 (93%) |
| Relation problems with conflict | 63 | 15 | 16 | 32 | 48 (76%) |
| Relational problems within the family | 50 | 6 | 18 | 26 | 44 (88%) |
| Obsessive disorder | 47 | 8 | 9 | 30 | 39 (83%) |
| Relational problems within couple | 46 | 10 | 12 | 24 | 36 (78%) |
| Panic disorder | 41 | 6 | 5 | 30 | 35 (85%) |
| Burn-out | 40 | 6 | 17 | 17 | 34 (85%) |
| Chronic fatigue syndrome | 39 | 6 | 11 | 22 | 33 (85%) |
| Generalised anxiety disorder | 39 | 8 | 10 | 21 | 31 (79%) |
| Depressive disorder | 37 | 11 | 11 | 15 | 26 (70%) |
| Post-traumatic stress disorder | 29 | 4 | 5 | 20 | 25 (86%) |
| Procrastination | 27 | 3 | 10 | 14 | 24 (89%) |
| Phobic behaviour | 25 | 2 | 11 | 12 | 23 (92%) |

(Continues)

TABLE 1 (CONTINUED)

| Category of patients' problem diagnoses | Number of cases | Failed intervention | Intervention with some improvement | Successful intervention | Success rate % (complete and partial) |
|---|-----------------|---------------------|------------------------------------|-------------------------|---------------------------------------|
| Perfectionism disorder | 23 | 2 | 7 | 14 | 21 (91%) |
| Borderline personality disorder | 19 | 4 | 6 | 9 | 15 (79%) |
| Avoidance disorder | 19 | 2 | 8 | 9 | 17 (89%) |
| Pathological grief | 18 | 2 | 3 | 13 | 16 (89%) |
| Performance anxiety | 17 | 2 | 5 | 10 | 15 (88%) |
| Substance abuse (alcohol and other drugs) | 16 | 5 | 6 | 5 | 11 (69%) |
| OCD (obsessive compulsive disorder) | 16 | 3 | 2 | 11 | 13 (73%) |
| Harassment | 15 | 3 | 6 | 6 | 12 (80%) |
| Problems related to change at work | 14 | 2 | 3 | 9 | 12 (86%) |
| Acute stress disorder | 13 | 0 | 4 | 9 | 13 (100%) |
| Persistent depressive disorder | 11 | 2 | 5 | 4 | 9 (82%) |
| Specific phobia/social anxiety disorder | 10 | 1 | 5 | 4 | 9 (90%) |

Thus, practitioners evaluated 80% of cases and patients 90% of cases as successful (problem solved to improved somewhat). Again, when success rates were calculated for those cases which have both practitioners' and patients' evaluations ($n = 219$), results revealed that practitioners evaluated 93% and patients 89% of the cases as successful.

Outcome results based on GHQ-12

Since 2018, when the GHQ-12 was introduced to the SYPRENE data recording system, we have collected data from 180 patients who completed the questionnaire at the beginning of their first and at the end of their last appointment (Figure 1). The within-group comparison of the GHQ-12 mean scores at the first ($M = 30.83$, $SD = 8.52$) and the last ($M = 10.97$, $SD = 5.54$) sessions revealed a statistically significant decrease of subjective suffering ($t = 26.58$, $p < .001$) with large effect size (Cohen $d = 1.68^6$, confidence interval for $d = 1.44$ – 1.92).

The difference of mean scores and effect size demonstrates statistical significance of change on a group level, but due to its naturalistic nature, the data obtained from SYPRENE do not provide the opportunity for comparison with any kind of control group. Thus, analysis of the reliable and clinically significant change (Jacobson and Truax, 1991), which allows estimation of the benefits of an intervention on the individual level compared with normative population data (how close to the normative or 'healthy' population a client has come during the therapy), was performed for the GHQ-12 data.

The concept of reliable change reflects the extent to which individual change on a measure exceeds the interval which represents the measurement variability of the instrument employed (Evans *et al.*, 1998). It is assessed using the standard error (SE) of measurement considering that two repeated measurements were made: $SE_{diff} = SD_I \sqrt{2} \sqrt{1-r}$ (where SD_I is the standard deviation of baseline observations; r is the reliability of the measure⁷). The change is considered as reliable with probability of 95% if it is larger than $SE_{diff} \times 1.96$ (Evans *et al.*, 1998).

⁶An online effect size calculator (Lenhard and Lenhard, 2016) for single-group repeated measures design (as suggested by Morris, 2008) was employed.

⁷To simplify and shorten the encoding process of SYPRENE practitioners, enter the total GHQ-12 scores, which does not allow calculation of internal reliability of the instrument based on collected data. Thus, the GHQ-12 internal reliability coefficient, provided in the literature for French and Spanish non-clinical populations (Lesage *et al.*, 2011; Sánchez-López and Dresch, 2008) was used in this analysis since most data in this study come from these national populations.

According to this formula, patients whose GHQ-12 score difference between the first and the last appointments exceeded 11.0 points were considered as reaching reliable change.

The change of a measure can be considered as clinically significant when there is a greater likelihood of the patient's post-treatment score falling closer to the average score of the normative (non-clinical) population than to the average score of the treated (clinical) population. Statistically, this requires the determination of the cut-off point where the probability of coming from each of the distributions is equal (Evans *et al.*, 1998): $[(mean_{clin} \times SD_{norm}) + (mean_{norm} \times SD_{clin})] / (SD_{norm} + SD_{clin})$ ⁸. In our data, the estimated cut-off point was 22.8, so the patients whose GHQ-12 scores were above the cut-off at the first appointment and below the cut-off at the last appointment were considered as having reached clinically significant change. The difference of the GHQ-12 score between the first and the last appointments was categorised as achieving or not achieving the reliable change, clinically significant change or both (Table 2).

Results suggest that change, as measured by GHQ-12, can be considered as reliable in 80% of cases and as clinically significant in 81% of cases. Accordingly, 76% ($n = 137$) of the patients could be considered as reaching both reliable and clinically significant change. These results are encouraging, especially considering that as many as 16% of the patients started better than the cut-off for clinically significant change (their GHQ-12 scores were in the normal population range at the first appointment); thus, eliminating them from the calculations would dramatically increase (up to 91%) the number of patients reaching reliable and clinically significant change. Finally, there were only 2% ($n = 3$) of patients who could (their initial GHQ-12 scores were in the range of the clinical group) but failed to reach clinically significant change.

Relationship between outcome results obtained by different measures

Since three different outcome measures representing practitioners' and patients' perspectives were used, we further address the question of the relationship between outcome results obtained using these measures. Since the problem-solving outcome was evaluated by practitioners and

⁸For normative data, we used GHQ-12 mean scores and SD from the articles, presenting GHQ-12 data for non-clinical French and Spanish populations (Lesage *et al.*, 2011; Sánchez-López and Dresch, 2008). Due to the coding differences mentioned in the articles presented GHQ-12 data with the 0-1-2-3 scoring system, mean GHQ-12 scores for normative population were recalculated to match the 0-1-2-3-4 coding system used in this article.

TABLE 2 *The distribution of participants in different groups of reliable and clinically significant change*

| | | Reliable change reached? | | Total |
|--|---|--------------------------|------------|-----------|
| | | No | Yes | |
| Clinically significant change reached? | Started better than cut-off (no need for clinically significant change) | 24 | 5 | 29 (16%) |
| | No | 3 | 2 | 5 (3%) |
| | Yes | 9 | 137 | 146 (81%) |
| Total: | 36 (20%) | 144 (80%) | 180 (100%) | |

patients on a similar 0–10-point scale, analysis of these evaluations was conducted and revealed strong correlation ($r = .77$, $p < .001$, $n = 219$). In addition, moderate agreement between practitioners' and patients' categorised evaluations of problem-solving outcomes (solved, partially solved and not solved) was revealed ($\kappa = .550$ (95% CI, .44 to .66), $p < .001$).

The problem-solving outcome measures indicate the state or severity of a problem at the end of the therapy. GHQ-12 score at the last appointment also reflects the level of a patient's mental distress after the therapy. Thus, we analysed the relationship between problem-solving outcomes and GHQ-12 scores at the last appointment. To account for possible differences in practitioners' impact on the delivery of therapy, a linear mixed model (Hox *et al.*, 2017) was used to model the relationship between the GHQ-12 at the last appointment (as the outcome variable) with practitioners' or patients' problem-solving outcome evaluations and GHQ-12 at the first appointment as fixed effects, as well as practitioners' identity as the random-effect variable on problem-solving outcome evaluation. The model included random slopes and random intercepts for problem-solving outcome evaluation. The software 'R' was used, along with the 'lmer' library. Model was fitted with restricted maximum likelihood estimation (REML). As presented in Table 3, the fixed-effect slope coefficient of practitioners' problem-solving outcome evaluation was -1.42 ($SD = 0.42$, $t = -3.32$). Similar results (Table 4) were obtained for patient's problem-solving outcome evaluation with a linear mixed model (fixed-effect patients' problem-solving slope coefficient was -1.24 ($SD = 0.60$, $t = -2.08$). Practitioners' and patients' random intercepts' standard deviation were equal to 33.5 and 61, respectively, on the GHQ-12 scale ranging from 0 to 48 (patients' standard deviation is higher than the GHQ-12 scale range due to the fact that random intercept is estimated as a Gaussian law which does not take the scale's limited range into account).

Furthermore, practitioners' problem-solving outcome evaluation random slope's standard deviation was 0.75, while the one for patients was 0.8. Compared with fixed effects slope coefficients of practitioners (-1.42) and patients (-1.24), these standard deviations suggest that for some the effect may be negligible and for others it may be much larger. However, as random effects are computed according to a normal law, 95% of practitioners and patients have a coefficient between two standard deviations. Thus, except for extreme cases, the relationship between problem solving and GHQ-12 is negative, with some being close to 0 and others quite close to -3 . This

TABLE 3 *Linear mixed model for predicting GHQ-12 score at the last appointment from practitioner problem-solving outcome evaluation and GHQ-12 score at the first appointment with practitioner identity as random effect*

| <i>Random effects</i> | | | | |
|--|--|----------|------|-----------------|
| Group | Name | Variance | SD | <i>r</i> |
| Therapist identity | (Intercept) | 33.51 | 5.78 | - |
| | Practitioner problem solving outcome evaluation last appointment | 0.57 | 0.75 | -0.99 |
| Residuals | | 25.92 | 5.09 | - |
| Number of observations: 174, Groups: Therapist identity, 8 | | | | |
| <i>Fixed effect</i> | | | | |
| | Name | Estimate | SE | <i>t</i> -value |
| | (Intercept) | 22.85 | 3.81 | 6.00 |
| | Practitioner problem-solving outcome evaluation last appointment | -1.42 | 0.43 | -3.33 |
| | GHQ-12 score at the first appointment | -0.00077 | 0.05 | -0.017 |

Note: Model formula: GHQ-12 score at the last appointment ~ Practitioner problem solving outcome evaluation + GHQ-12 score at the first appointment + (1 + Practitioner problem solving outcome evaluation | Therapist identity).

TABLE 4 *Linear mixed model for predicting GHQ-12 score at the last appointment from patient problem solving outcome evaluation and GHQ-12 score at the first appointment with practitioner identity as random effect*

| <i>Random effects</i> | | | | |
|---|---|----------|------|-----------------|
| Group | Name | Variance | SD | <i>r</i> |
| Therapist identity | (Intercept) | 61.23 | 7.82 | - |
| | Patient problem-solving outcome evaluation last appointment | 0.64 | 0.80 | -1 |
| Residuals | | 31.94 | 5.65 | - |
| Number of observations: 61, Groups: Therapist identity, 7 | | | | |
| <i>Fixed effect</i> | | | | |
| | Name | Estimate | SE | <i>t</i> -value |
| | (Intercept) | 19.92 | 5.96 | 3.34 |
| | Patient problem solving outcome evaluation last appointment | -1.24 | 0.60 | -2.08 |
| | GHQ-12 score at the first appointment | 0.08 | 0.07 | 1.03 |

Note: Model formula: GHQ-12 score at the last appointment ~ Patient problem solving outcome evaluation + GHQ-12 score at the first appointment + (1 + Patient problem solving outcome evaluation | Therapist identity).

suggests that improving the problem-solving score generally leads to a better mental state, reflected by a decrease of the GHQ-12 score at the last appointment compared with the first appointment. To conclude, results suggest that solving people's specific problems during therapy is correlated with a better mental state at therapy's end.

Discussion

Results presented in this article suggest the following:

1. Initial outcome results from SYPRENE are promising. They suggest that strategic therapy helped patients achieve considerable improvement. Outcome results indicate a problem-solving improvement rate of 80% as evaluated by practitioners, and 90% as evaluated by patients. In addition, patients reported significant improvement in their psychosocial functioning with a large effect size. In 76% of cases, this improvement was evaluated as reaching reliable and clinically significant change. These results

reflect the considerable improvement of patients' problem situations and decrease in subjective suffering for most of the patients and are comparable with those obtained by other authors who report that systemic and strategic therapies are effective, with success rates ranging between 60% and 87% (e.g. Jackson *et al.*, 2018; Nardone and Watzlawick, 2005; Pietrabissa *et al.*, 2016).

2. SYPRENE data suggest good efficiency for strategic therapy. Encouraging outcome results were achieved with an average of 5.4 sessions, within an average of 5.3 months. Length of the therapy in this study is comparable with results obtained by the Centro di Terapia Strategica d'Arezzo in their research over the past three decades (Nardone and Balbi, 2012). In addition, the drop-out rate was only 19%, which compares favourably with other large-scale studies (see Barrett *et al.*, 2008; Swift and Greenberg, 2012; Wells *et al.*, 2013).
3. There is a significant positive relationship between outcomes evaluated by different measures and from different perspectives. While all outcome measures in SYPRENE rely on self-report evaluations by patients and their therapists, credibility of outcome data comes from the fairly good correspondence between the various evaluations with the following implications: (1) large improvement was obtained from patients' and practitioners' perspectives, indicating that strategic therapy helps to improve the problem situation and decrease patients' feelings of subjective distress; (2) practitioners' and patients' evaluations of problem-solving outcome were strongly correlated with moderate agreement, suggesting that problem improvement was perceived quite similarly from both perspectives; (3) practitioners' and patients' evaluations of problem improvement were significantly linked to the patients' GHQ-12 scores at the end of therapy, even when controlling for GHQ-12 baseline scores and differences between therapists. At this point, we can conclude that outcome evaluations on all three measures point to the same direction and are related but not identical. Such results might suggest that measures from different sources and of different nature (such as subjective scales and standardised questionnaires) could reflect different perspectives of therapeutic change. Hence, practice research networks, such as SYPRENE, should include both patients' and practitioners' perspectives when assessing outcomes in naturalistic settings to provide a greater overall understanding of therapeutic change.

Limitations, challenges and opportunities

Although the presented preliminary results are promising, they are somewhat limited by the size of our sample as well as a lack of broad national heterogeneity. In terms of effectiveness, the present results mainly reflect the perspective of the therapist and the patient regarding overall outcome ratings. The recent introduction (in 2018) of the GHQ-12 patient-rated questionnaire at the beginning and end of the therapy process ($n = 180$) has shown encouraging results. We are also considering the use of additional questionnaires about the patient's experience of the change process. However, the benefits of more measures need to be weighed against the possibility of asking too many questions. In terms of *efficiency*, we will further explore possible factors, including for those cases when problems remain unresolved. In addition to measuring the effectiveness and efficiency of interventions, SYPRENE data will make it possible to evaluate more qualitative dimensions and change processes.

Since SYPRENE is a living PRN, we have naturalistic patterns of data rather than having identical cell sizes throughout. In the future, more data will allow more robust statistical analyses, such as evaluation of the multiple factors (e.g. gender, age, relationship status and number of sessions) possibly influencing the outcome results. One of the important issues for future research is control of missing data. In the present study, even though this group did not differ from the participants with complete outcome data in respect to sociodemographic characteristics, we do not know for sure whether there were no differences in life circumstances or other variables possibly accounting for outcome results. To reduce the proportion of missing data, we are considering the possibility of an additional evaluation several months after the therapy. The evaluation of the role of practitioner differences is also an important question which we hope to explore more in the near future. SYPRENE data allow raising questions about the processes of therapeutic change, and future observations will help us to understand better which prescriptions will be most useful.

Funding information

No conflicts of interest.

References

- Barrett, M. S., Chua, W.-J., Crits-Christoph, P., Gibbons, M. B. and Thompson, D. (2008). Early withdrawal from mental health treatment: implications for psychotherapy practice. *Psychotherapy: Theory, Research, Practice, Training*, **45**(2): 247–267.
- Bateson, G. (1972) *Steps to an ecology of mind*. San Francisco, CA: Chandler Publishing.
- Bateson, G. (1979) *Mind and nature: A necessary unity*. New York: Dutton.
- Baucom, D. H., Fischer, M. S., Worrell, M., Corrie, S., Belus, J. M., Molyva, E. and Boeding, S. E. (2018) Couple-based intervention for depression: an effectiveness study in the National Health Service in England. *Family Process*, **57**(2): 275–292.
- Boscolo, L., Cecchin, G., Hoffman, K. and Penn, P. (1987) *Milan systemic family therapy: conversations in theory and practice*. New York: Basic Books.
- de Scorraille, C., Brosseau, O. and Vitry, G. (2017) *Quand le travail fait mal: une clinique de la relation pour soigner les maux au travail*. InterÉditions.
- del Sánchez-López, M. P. and Dresch, V. (2008) The 12-Item General Health Questionnaire (GHQ-12): reliability, external validity and factor structure in the Spanish population. *Psicothema*, **20**(4): 839–843.
- Erickson, M. H. (1980) *Collected papers* (Vols. 1-4; E. L. Rossi, Ed.). New York: Irvington.
- Evans, C., Margison, F. and Barkham, M. (1998) The contribution of reliable and clinically significant change methods to evidence-based mental health. *Evidence-Based Mental Health*, **1**(3): 70–72.
- Fisch, R., Weakland, J. H. and Segal, L. (1982) *The tactics of change: doing therapy briefly*. San Francisco, CA: Jossey-Bass.
- Goldberg, D. and Williams, P. (1991) *A user's guide to the General Health Questionnaire*. London: NFER-Nelson 1988.
- Hale, D. and Frusha, C. (2016) MRI Brief Therapy: A tried and true systemic approach. *Journal of Systemic Therapies*, **35**(2): 14–24.
- Haley, J. (1973) *Uncommon therapy: the psychiatric techniques of Milton H. Erickson, M.D.* New York: Norton.
- Haley, J. (1977) *Problem-solving therapy: new strategies for effective family therapy*. San Francisco, CA: Jossey-Bass.
- Horigian, V. E., Anderson, A. R. and Szapocznik, J. (2016) Taking brief strategic family therapy from bench to trench: evidence generation across translational phases. *Family Process*, **55**(3): 529–542.
- Horigian, V. E., Feaster, D. J., Robbins, M. S., Brincks, A. M., Ucha, J., Rohrbaugh, M. J., Shoham, V., Bachrach, K., Miller, M., Burlew, A. K., Hodgkins, C. C., Carrion, I. S., Silverstein, M., Werstlein, R. and Szapocznik, J. (2015) A cross-sectional assessment of the long term effects of brief strategic family therapy for adolescent substance use. *American Journal on Addictions*, **24**(7): 637–645.
- Hox, J. J., Moerbeek, M. and van de Schoot, R. (2017) *Multilevel analysis: techniques and applications*. New York: Routledge.
- Hoyt, M. F. (2019) Strategic therapies: roots and branches. *Journal of Systemic Therapies*, **38**(1): 30–43.
- Jackson, J. B., Pietrabissa, G., Rossi, A., Manzoni, G. M. and Castelnovo, G. (2018) Brief strategic therapy and cognitive behavioral therapy for women with binge eating disorder and comorbid obesity: a randomized clinical trial one-year follow-up. *Journal of Consulting and Clinical Psychology*, **86**(8): 688–701.

- Jacobson, N. S. and Truax, P. (1991) Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, **59**(1): 12–19.
- Johnson, L. N., Miller, R. B., Bradford, A. B. and Anderson, S. R. (2017) The Marriage and Family Therapy Practice Research Network (MFT-PRN): Creating a more perfect union between practice and research. *Journal of Marital and Family Therapy*, **43**(4): 561–572.
- Lebow, J. L. (2018) Editorial: Effectiveness research in couple and family therapy. *Family Process*, **57**(2): 271–274.
- Lenhard, W. and Lenhard, A. (2016) Calculation of effect sizes. *Psychometrica*. https://www.psychometrica.de/effect_size.html
- Lesage, F.-X., Martens-Resende, S., Deschamps, F. and Berjot, S. (2011) Validation of the General Health Questionnaire (GHQ-12) adapted to a work-related context. *Open Journal of Preventive Medicine*, **1**(02): 44.
- Madanes, C. (1981) *Strategic family therapy*. San Francisco, CA: Jossey-Bass.
- Madanes, C. (1984) *Behind the one-way mirror: advances in the practice of strategic therapy*. San Francisco, CA: Jossey-Bass.
- Miller, S. D. and Duncan, B. L. (2004) *The outcome and session rating scales: administration and scoring manual*. Chicago, IL: Institute for the Study of Therapeutic Change.
- Molnar, A. and Shazer, S. (1987) Solution-focused therapy: toward the identification of therapeutic tasks. *Journal of Marital and Family Therapy*, **13**(4): 349–358.
- Morris, S. B. (2008) Estimating effect sizes from pretest-posttest-control group designs. *Organizational Research Methods*, **11**(2): 364–386. <https://doi.org/10.1177/1094428106291059>
- Nardone, G. and Balbi, E. (2012) *Sillonner la mer à l'insu du ciel*. Bruxelles, Belgium: Editions Satas.
- Nardone, G. and Portelli, C. (2005) *Knowing through changing: the evolution of brief strategic therapy*. Norwalk, CT: Crown House Publishing.
- Nardone, G. and Watzlawick, P. (1993) *The art of change: strategic therapy and hypnotherapy without trance*. San Francisco: Jossey-Bass.
- Nardone, G. and Watzlawick, P. (2005) *Brief strategic therapy: philosophy, techniques, and research*. Lanham, MD: Jason Aronson.
- Pietrabissa, G., Manzoni, G. M., Gibson, P., Boardman, D., Gori, A. and Castelnovo, G. (2016) Brief strategic therapy for obsessive-compulsive disorder: a clinical and research protocol of a one-group observational study. *British Medical Journal Open*, **6**(3): 1–9.
- Reicheport-Haley, M. and Carlson, J. (Eds.) (2012) *Jay Haley revisited*. New York: Routledge.
- Seligman, M. E. (1995) The effectiveness of psychotherapy. The Consumer Reports study. *American Psychologist*, **50**(12): 965–974.
- Swift, J. K. and Greenberg, R. P. (2012) Premature discontinuation in adult psychotherapy: a meta-analysis. *Journal of Consulting and Clinical Psychology*, **80**(4): 547–559.
- Tasca, G. A., Sylvestre, J., Balfour, L., Chyurlia, L., ... Wilson, B. (2015) What clinicians want: findings from a psychotherapy practice research network survey. *Psychotherapy*, **52**(1): 1–11.

- Teachman, B. A., Drabick, D. A. G., Hershenberg, R., Vivian, D., Wolfe, B. E. and Goldfried, M. R. (2012) Bridging the gap between clinical research and clinical practice: introduction to the special section. *Psychotherapy*, **49**(2): 97–100.
- Thurin, J.-M., Thurin, M. and Midgley, N. (2012) Does participation in research lead to changes in attitudes among clinicians? Report on a survey of those involved in a French practice research network. *Counselling and Psychotherapy Research*, **12**(3): 187–193.
- Vitry, G., de Scorraille, C. and Hoyt, M. F. (in press). Redundant attempted solutions: 50 years of theory, evolution, and new supporting data. *Australian and New Zealand Journal of Family Therapy*, forthcoming.
- Vitry, G., Duriez, N., Suberville, S., Pakrosnis, R., Beau, A., Garcia-Rivera, T., Brosseau, O., Vargas Avalos, P., Bardot, E. and Ray, W. A. (2020) Introducing SYPRENE: an international practice research network for strategic and systemic therapists and researchers. *Family Process*, **59**(4): 1946–1957. <https://doi.org/10.1111/famp.12520>
- Vitry, G., Garcia, T., de Scorraille, C., Paoli, B. and Brosseau, O. (2019). *Stratégies de changement. 16 prescriptions thérapeutiques*. Toulouse, France: Érès.
- Watzlawick, P. (2009). *Insight may cause blindness and other essays* (W. A. Ray & G. Nardone, Eds.). Phoenix, AZ: Zeig, Tucker & Theisen.
- Watzlawick, P., Weakland, J. H. and Fisch, R. (1974) *Change: principles of problem formation and problem resolution*. New York: Norton.
- Wells, J. E., Browne, M. O., et al. (2013). Drop out from outpatient mental health-care in the World Health Organization's World Mental Health Survey initiative. *British Journal of Psychiatry*, **202**(1): 42–49. <https://doi.org/10.1192/bjp.bp.112.11313>
- Werneke, U., Goldberg, D. P., Yalcin, I. and Ustün, B. T. (2000) The stability of the factor structure of the General Health Questionnaire. *Psychological Medicine*, **30**(4): 823–829.
- Wittezeale, J.-J. and Nardone, G. (2016) *Une logique des troubles mentaux*. Le Seuil: Le diagnostic opératoire systémique et stratégique.