

Examining concurrent validity and item selection of the Wants and Needs Outcome Measure (SWAN-OM) in a web-based therapy service

Santiago De Ossorno Garcia, Julian Edbrooke-Childs, Louisa Salhi, Florence Ruby, Aaron Sefi, Jenna Jacob

Submitted to: JMIR Formative Research
on: June 08, 2022

Disclaimer: © The authors. All rights reserved. This is a privileged document currently under peer-review/community review. Authors have provided JMIR Publications with an exclusive license to publish this preprint on its website for review purposes only. While the final peer-reviewed paper may be licensed under a CC BY license on publication, at this stage authors and publisher expressly prohibit redistribution of this draft paper other than for review purposes.

Table of Contents

Original Manuscript	5
Supplementary Files	35
0.....	36
Figures	37
Figure 1.....	38
Multimedia Appendixes	39
Multimedia Appendix 1.....	40
Multimedia Appendix 2.....	40

Preprint
JMIR Publications

Examining concurrent validity and item selection of the Wants and Needs Outcome Measure (SWAN-OM) in a web-based therapy service

Santiago De Ossorno Garcia¹ BSc, MSc, PhD; Julian Edbrooke-Childs^{2,3} BSc, MSc, PhD; Louisa Salhi^{1,4} BSc, MSc, AFS, PhD; Florence Ruby² BSc, MSc, PhD; Aaron Sefi¹ BA, MSc, PhD; Jenna Jacob² BSc, MSc, PhD

¹Kooth Plc London GB

²Anna Freud Centre, CORC London GB

³UCL London GB

⁴University of Kent Canterbury, Kent GB

Corresponding Author:

Santiago De Ossorno Garcia BSc, MSc, PhD

Kooth Plc

5 Merchant Square

London

GB

Abstract

Background: Single-session mental health interventions are frequently attended by children and young people in both web-based and face-to-face therapy settings. The Session “Wants” and “Needs” Outcome Measure (SWAN-OM) is an instrument developed in a web-based therapy service to overcome challenges of collecting outcomes and experiences of single-session therapies. It provides pre-defined goals of the session, selected by the young person prior to the intervention, on which progress towards achievement is scored at the end of the session.

Objective: To evaluate the instrument’s psychometric properties, including concurrent validity against other frequently used outcome and experience measures, at a web-based service.

Methods: The SWAN-OM was administered for a period of six months to 1401 children and young people (aged 10 to 32; 79.3% White; 77.59% female) accessing single-session therapy in a web-based service. Item correlations with comparator measures and hierarchical logistic regressions to predict item selection were calculated for concurrent validity and psychometric exploration.

Results: Most frequently selected items are “Feel better” (N= 431; 11.61%) and “Find ways I can help myself” (N= 411; 11.07%); unpopular items are “Feel safe in my relationships” (N= 53; 1.43%) and “Learn the steps to achieve something I want” (N= 58; 1.56%). The SWAN-OM is significantly correlated with the Experience of Service Questionnaire (ESQ), particularly the item: “Feel better”(rs(109) = .48, p < .001), the Youth Counselling Impact Scale (YCIS), particularly the item: “Learn the steps to achieve something I want” (rs(22) = .76, p < .001), and the Positive and Negative Affect Schedule (PANAS), particularly the items: item “Learn how to feel better” (rs(22) = .72, p < .001) and were “Explore how I feel” (rs(70) = -.44, p < .001).

Conclusions: The SWAN-OM demonstrates good concurrent validity with common measures of outcome and experience. Analysis suggests that lesser endorsed items may be removed in future iterations of the measure, to improve functionality. Future research is required to explore SWAN-OM’s potential to measure meaningful change, in a range of therapeutic settings.

(JMIR Preprints 08/06/2022:40122)

DOI: <https://doi.org/10.2196/preprints.40122>

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?

✓ **Please make my preprint PDF available to anyone at any time (recommended).**

Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

No, I do not wish to publish my submitted manuscript as a preprint.

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain visible to the public.

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <http://www.jmir.org>, I will be able to access the full text of my article.

Preprint
JMIR Publications

Original Manuscript



Examining concurrent validity and item selection of the Wants and Needs Outcome Measure (SWAN-OM) in a web-based therapy service

Authors

Santiago De Ossorno Garcia^{1*} [0000-0001-6660-2109](https://orcid.org/0000-0001-6660-2109)

Julian Edbrooke-Childs^{2,3} ORCID: 0000-0003-0401-4058

Louisa Salhi^{1,4} [0000-0001-6458-1391](https://orcid.org/0000-0001-6458-1391)

Florence Ruby²

Aaron Sefi¹

Jenna Jacob² ORCID: 0000-0003-1006-1547

¹Kooth Plc, London, UK

²Anna Freud Centre, CORC, 4-8 Rodney Street, London, N1 9JH, UK

³UCL, Gower St, London WC1E 6BT, UK

⁴University of Kent, Canterbury, Kent

*

Correspondence:

Santiago de Ossorno Garcia
sossorno@Kooth.com

Abstract

Background: Single-session mental health interventions are frequently attended by children and young people in both web-based and face-to-face therapy settings. The Session “Wants” and “Needs” Outcome Measure (SWAN-OM) is an instrument developed in a web-based therapy service to overcome challenges of collecting outcomes and experiences of single-session therapies. It provides pre-defined goals of the session, selected by the young person prior to the intervention, on which progress towards achievement is scored at the end of the session.

Objective: To evaluate the instrument’s psychometric properties, including concurrent validity against other frequently used outcome and experience measures, at a web-based service.

Methods: The SWAN-OM was administered for a period of six months to 1401 children and young people (aged 10 to 32; 79.3% White; 77.59% female) accessing single-session therapy in a web-based service. Item correlations with comparator measures and hierarchical logistic regressions to predict item selection were calculated for concurrent validity and psychometric exploration.

Results: Most frequently selected items are “*Feel better*” (N= 431; 11.61%) and “*Find ways I can help myself*” (N= 411; 11.07%); unpopular items are “*Feel safe in my relationships*” (N= 53; 1.43%) and “*Learn the steps to achieve something I want*” (N= 58; 1.56%). The SWAN-OM is significantly

correlated with the Experience of Service Questionnaire (ESQ), particularly the item: “*Feel better*” ($r(109) = .48, p < .001$), the Youth Counselling Impact Scale (YCIS), particularly the item: “*Learn the steps to achieve something I want*” ($r(22) = .76, p < .001$), and the Positive and Negative Affect Schedule (PANAS), particularly the items: item “*Learn how to feel better*” ($r(22) = .72, p < .001$) and were “*Explore how I feel*” ($r(70) = -.44, p < .001$).

Conclusions: The SWAN-OM demonstrates good concurrent validity with common measures of outcome and experience. Analysis suggests that lesser endorsed items may be removed in future iterations of the measure, to improve functionality. Future research is required to explore SWAN-OM’s potential to measure meaningful change, in a range of therapeutic settings.

Keywords: Single-session therapy; instrument evaluation; digital mental health; SWAN-OM; web-based therapy; patient-reported outcome measure; concurrent validity.

Introduction

In the field of digital mental health and services, where the aim is often to increase access to services and provide choice and flexibility, brief and sometimes only one session is often the common engagement form of professional support. Digital mental health services are well placed to deliver brief and focused interventions, with human-mediated support, as well as through evidence-based, self-guided programs.^{1,2}

Clients as users using web-based delivered services will often mimic face-to-face services with a single-session engagement, with one-off sessions being the most frequent across services.^{3,4} The opportunities to examine the changes and monitoring of interventions in the digital healthcare context are promising, thanks to tracking information technology, ease and quick access to the intervention, and large data volumes that can be collected and analyzed quickly.⁵ A digitally-enabled intervention of web-based, single-session therapy (SST) intervention is a good starting point to understand how change takes place in this intervention and continue to collect evidence about the effectiveness and impact of SST.

What are single-session therapies?

Single-Session Therapies (SST) or One-At-A-Time (OAAT) approaches⁶ are interventions delivered by practitioners across a range of settings and psychological support services. SST and OAAT approaches are broad, and they have been defined as “*A purposeful endeavour where both parties set out with the intention of helping the client in one session, knowing that more help is available if needed.*”⁷ SST is conducted by professionals who seek to use their existing therapeutic skillset to address the presenting concerns or problems within one session and assumes that support will not be ongoing over several weeks or months.⁸ SST uses a strength-based approach to solve problems; the focus lies on client-led “in-session” goals rather than longer-term therapeutic goals, making the most of someone’s circumstances. This type of support lends itself to a person-centered approach in which objectives and outcomes are client led, rather than manualized course of therapy outcomes.⁹ Most services offering single-session therapy, like in traditional “Walk-in therapy”,¹⁰ people access the service at the “point of need”, and no appointment is necessary to receive support. This additional support is offered alongside SSTs and other brief interventions,^{11,12} or services may offer further support that the individual can use after (e.g. signposting advice, further counselling sessions, group sessions, referrals, etc.). Some use the term “one-at-a-time”¹³ to avoid misunderstanding or resistance by the clinical community that single-session therapy means “only one”; in single sessions and other brief interventions, the work can continue beyond the implicit “one time” misconception often attributed to the SST term, so both terms SST and OAAT, despite some differences, often implies

that more help is available and people can access it “at the time of need”.⁶

Single-session therapies and brief interventions are gaining momentum and adoption amongst mental health services, especially for young populations. Brief interventions like SST may help to address the unmet needs of people waiting for services and help to meet the steady increase of demand for mental health support in recent decades. This increased demand calls for a transformation of services through novel models of effective delivery, including SST and brief approaches.^{14,15} For example, metanalytic evidence suggested that single-session targeted interventions can be effective in anxiety reduction, conduct problems, and substance use.^{16–18} However, single-session interventions, or SST, have not always been considered a meaningful and effective type of therapeutic intervention. This is partially due to the assumption that one session may indicate dissatisfaction or service drop-out, as the client has not finished their course of therapy or assessment as design of practice is concerned. This is, however, more difficult to discern for internet-delivered and web-based interventions,¹⁹ especially if the SST model is not explicit. It has taken some time for the field of counselling to recognize SST as a relevant therapeutic intervention.²⁰ Evidence from walk-in therapy clinics across the world²¹ have contributed to the acceptance of SSTs, as well as evidence from some studies reporting that 70% to 95% of people receiving SST were satisfied with their session.^{3,22,23} A further trial reported that one session was perceived as enough at the clinics when offering support.²⁴ Single sessions and brief interventions are important to be examined, as they may not only be a cost-effective way to increase access and provide a scalable solution for the rising demand of mental health support, but also a way to understand needs and access from the population that mental health services are intending to serve.

SST has been previously recommended as a treatment of choice for children and young people presenting with mental health difficulties²³ and has the potential to be one of the drivers for system change and transformation described in “THRIVE” framework to support children and young people’s mental health.²⁵ Moreover, the National Health System (NHS) in the UK has made a move to accept one-session interventions as one of the potential changes that may help to tilt the needle on waiting times for Improving Access to Psychological Therapies²⁶; this supports the wider popularity of SST and its recognition as a therapeutic intervention. This is further supported by evidence by Children and Adolescent Mental Health Services (CAMHS) in the UK, where one session is the most common way of engagement with services.²⁷ In addition, SST may be preferred by certain groups of young people when accessing therapy, namely those who value choice and flexibility when receiving support; these are two clearly defined factors for pluralistic and accessible psychotherapy provision.^{18,23,28} One-session engagement from specialist mental health services at CAMHS are evidenced as being most commonly attended by young people with complex needs, from minoritized ethnic groups, with relational difficulties with peers, and also by those with less frequently occurring problems; this study, however, did not examine SST in their evaluation of engagement.²⁷ It is therefore yet to be known how useful it will be to implement tailored therapies like SST to the most common way of engagement in services and its effectiveness and outcomes to individuals.

Outcome measurement for single-session therapies

A range of outcome measures are used to measure SST effectiveness.^{29–31} These often are targeted to specific mental health difficulties (e.g. anxiety, depression) such as the Revised Children Anxiety and Depression Scale¹, Pediatric Anxiety Rating Scale,³² or Counselling Progress and Depth Rating Instrument,^{33,34} amongst others. However, there is a prevailing challenge related to the short-term nature of the SST work and the disconnect with the longer-term measurement of mental health difficulties, and authors have highlighted the importance of measuring immediate changes after SST.¹⁸ On one hand, SST seems to be influenced by unspecific factors to the overall success of SST and impact on change score and clinical improvement³⁵. On the other hand, SST is not a treatment

modality that easily enables the measurement of change over several time points,³⁶ although there is a growing effort and evidence to demonstrate clinically significant improvement over time with controlled studies^{17,37,38}; most of them showed short-term improvement only after one month and failed to demonstrate improvement in further follow-ups.^{39,40} Other outcome instruments for psychotherapy are designed assuming a monitoring connotation, or a course of therapy that does not match the SST framework, where using a series of scores to monitor change is not expected. (e.g. YP-CORE and CORE⁴¹) This also raises considerable issues when embarking on testing the reliability (in terms of measurement error) of SST outcome measures. Single session therapies often assume that ongoing sessions are not required for improvement.³⁸ This emphasizes why it may not be possible, or appropriate, to examine the test-retest and repeated measurement in SST outcomes, as these should be related directly to the session outcomes and experience of the intervention, rather than something re-measurable at a later measurement point.

In pluralistic services, in particular, the therapeutic background and practitioner perceptions influence the course of therapy, allowing different therapeutic approaches to be used in the SST; thus it is difficult to systematize or explain the components that lead to effectiveness.⁴² Overall, tracking progress from SSTs can be deemed a challenge, and further follow-up with the young person is not obtained or sought by providers regularly, thus making it difficult to obtain longitudinal data outside of controlled studies. There is also a further challenge of capturing personalized outcomes and goals, which complement the pluralistic nature of SST work, due to the brevity of these interactions. Balancing the need for a short, tailored measure, that can serve as an outcome within the brief nature of the single-session, but also as a measurement instrument that helps to focus the brief encounter and maximize time working with the person-chosen goals.

The suggested solution, which addresses the challenges highlighted, is a patient-reported outcome measure that captures individual “Wants” and “Needs” of the single-session. It highlights the importance of delivering a pluralistic and person-centered intervention in SST, and it was designed with this in mind. The instrument also assumes that to be able to obtain single-session outcomes (a meaningful measurement of change), where the session goal expectations should be led and personalized by the client, too.⁴³ Alignment between practitioner and patient-therapeutic outcomes expectations are critical when providing SST, especially when the session aims and focus has been identified by both parties. Practitioners often need to assess if indeed these expectations brought to SST are realistic for this type of presenting concerns or problems alongside monitoring any disclosure or indication for risk of harm and safeguarding. An instrument that set a limited number of “in-session” goals can help to develop this alignment, as well as to enhance the delivery of SST in web-based services. The SST measure provides a solution to collect, in a systematic way, aggregated SST outcomes for services delivering SST.⁴⁴ As a patient-reported instrument, it also provides the client with choice, by giving control of what they expect to cover in SST and introducing the ability to personalize these “Wants” and “Needs” if preferred by the user.

What is the SWAN-OM?

To address the need for a tailored measure to track single-session therapeutic work, a new instrument contextualized for SST was required; the “*Session Wants And Needs Outcome Measure*” (SWAN-OM⁴⁵) was developed in a digital web-based mental health service delivering SST and OAAT approaches via text-based synchronous messages. The measure was originally developed in a four-phase design to examine the content and face validity of the measure aimed at children and young people (aged 10 to 25), including a pilot of the measure and usability testing with relevant stakeholders, including practitioners and diverse groups of young people.⁴⁵

The instrument provides service users with a list of “Wants” and “Needs” to choose from, alongside a personalized option (“free-text”). The SWAN-OM has a novel format with a two-step filtering logic, where young people can select from six higher-level themes, and within these themes, specific items. Once the number of selected themes is explored, young people can select up to three items in total to cover in their SST goals. Once these “in-session” goals are selected, the practitioner can look at what the person has chosen to focus their SST. This gives information to the practitioner on how their intervention can be tailored to each individual, especially in the context of a digital mental health service; anonymous, free, and accessible where users may present with a wide range of concerns.

Outcomes and experiences after the session are measured on a Likert scale indicating how much they achieved what they initially wanted. This instrument can determine if “Wants” and “Needs” were met throughout the SST encounter. At the end of their session, they are only asked about how much they achieved the items they selected, rather than how much they achieved across all instrument items. This provides a novel way to measure what young people “Want” or “Need” from an SST in a web-based therapy service. It also facilitates the formulation of “in-session” goals as items for the practitioner to structure their session. The two-stage logic measure structure, going from a group of themes for selection to item display and selection of “Needs”, provides a manageable “in-session” goal setting activity for young people; this logic structure was suggested by young people in a stakeholder workshop during its development in order to present the information. The SWAN-OM structure means that traditional psychometric testing may not be appropriate; however, there is an opportunity to examine this measure at an item level, which we go on to explore in this paper.

Aims

This study aims to explore the concurrent validity of the SWAN-OM using other standardized instruments as comparator measures, chosen due to relevance, similarity of items, and immediacy. The study further explores construct validity through prediction of item selection based on the population characteristics. We also discuss limitations and considerations when examining psychometric properties of instruments which have novel structural designs, such as the SWAN-OM.

The present study examines the data collected from an evaluation conducted in a web-based counselling service, where SWAN-OM was administered alongside the other measures. We hypothesized that a therapeutic encounter of SST should have a positive association with positive emotional changes. Therefore, positive SWAN-OM scores would correlate positively with a positive affect scale; we also expect to see positive changes in the affect scale before and after the SST. We expect most items from SWAN-OM to correlate positively with a session progress rating scale. Finally, we expect positive SWAN-OM scores to correlate positively with an experience or satisfaction of service measure.

Methods

Participants

Young people who participated in this study were users of a digital mental health service (Kooth.com). To be eligible to take part, young people had to have no previous engagement with counselling sessions within the service. Young people are commonly aged between 10 to 25 years old and anonymously register on the digital health platform; there are some exceptions with a small number of services seeing older young adults in specific locations. All participants were required to have requested access to a synchronous text-based chat session with a practitioner in the digital service online. Data from young people was collected on the service between January 2021 and June

2021. Only data from young people who had explicitly provided research consent when using the service was available for this study. Gender, Age and Ethnicity were self-reported variables collected directly from the young people as part of the service sign-up process for the digital service.

Over this evaluation period, 1401 young people accessed 1901 chats within the service. On average, a young person accessed the chat 3.2 times during the study period (with a minimum of 1 and a maximum of 26 chats), and each chat lasted on average 52 minutes ($SD=21.6$; extreme outliers removed). Young people who took part were aged between 10 and 32, with an average Age of 15.9 years ($SD=2.9$). Most young people accessing the service were Female ($N=1087$; 77.59%) from a White ethnic background ($N=1111$; 79.3%). In total, 1435 (75.13%) chats included information about the participants' presenting concerns, as reported by practitioners. The majority indicated experiencing difficulties with anxiety/stress, suicidal thoughts, self-harm, and family relationships (Table 1).

Table 1. Demographic characteristics. ^a

Demographic variables	n	%
Gender		
Agender	47	3.35%
Female	1087	77.59%
Gender fluid	57	4.07%
Male	210	14.99%
Age		
10 to 14	492	35.12%
15 to 19	788	56.25%
20 and above	121	8.64%
Ethnicity		
Any other ethnic group	10	0.71%
Asian / Asian British	106	7.57%
Black African / Caribbean / Black British	49	3.5%
Mixed multiple ethnic group	78	5.57%
White	1111	79.3%
Not Specified	47	3.35%
Presenting Concerns		
Mental Health	1105	57.85%
External issues	734	38.43%
Suicidal thoughts / Self-harm	601	31.47%
Risk	163	8.53%
Physical / Other	83	4.35%
No information provided	475	24.87%

^a Data collected from 1401 young people attending 1910 chats. Percentages are based on the total number of young people for the categories Gender, Age and Ethnicity; percentages are based on the total number of chats for the categories related to Presenting Concerns. Percentages reported for Presenting Concerns do not add up to 100% as young people can be assigned more than one presenting concerns per chat.

Instruments and variables

Young People's Characteristics

The service collects four different categories for Gender (Male, Female, Agender, and Genderfluid). Age was collected as a continuous variable and divided in 3 age groups for analysis purposes (10 to 14 years, 15 to 19 years, 20 years and above). Ethnicity was also grouped into five categories as

recommended by the ONS with the available service data.⁴⁶

Following each chat that takes place in the digital service, practitioners record the “presenting concerns” the young person experience during the session (the service routinely reports and records presenting concerns after one chat); this information was collected from the service for each participant of the study. The variable “presenting concerns” was grouped into five higher-order categories to enable analysis (Mental Health [MH]; External [E]; Suicidal thoughts/self-harm [SS]; Risk [R]; Physical or other problems [PO]). These categories were composed by more specific concerns or problems. For example, risk was composed of presenting problems such as psychotic episodes, trauma, abuse, and parental mental health, whereas physical or other problems will group financial hardship, physical illness, or pregnancy. These are practitioner reported presetting concerns that were discussed in the SST and captured by the practitioner or clinician.

Session Wants and Needs Outcome Measure

The Session Wants And Needs Outcome Measure (SWAN-OM)⁴⁵ is a 21-item single-session outcome measure; face and construct validity of the instrument has been previously examined as part of its development and design within the digital service.⁴⁵ Young people are presented with the SWAN-OM prior to the intervention (Pre-chat item selection) and immediately after the intervention (post-chat item scoring).

The instrument follows a two-step logic:

- First, ask the young person to select any of the following: “*In my chat I would like to...*” (A: “Understand what help I can get”; B: “Share my story with someone”; C: “Set and achieve my goals”; D: “Explore my emotions”; E: “Improve my relationships”; F: “Learn ways to cope”).
- Second, depending on the theme selection, the 21-items from SWAN-OM will be displayed after a “*select up to 3 things in to focus on in your chat today*” that reflect their aims for the chat session. These are seen as “in-session” pre-defined goals tailored to the “Wants” and “Needs” reported by the young person at the point of access to the session (Figure 1).
- After the SST intervention (post-chat item scoring), young people are again presented with the instrument and asked “*Did your chat support you in the way you hoped?*” to indicate how much progress they had made on each, using a 5-point Likert ([-2]: “strongly disagree” to [2]: “strongly agree”) with follow-up statements that match what was selected prior to the session (see statements in Appendix).
- One of the 21 items is a free-text option (write your own) for personalization. For this personalized item, young people are presented with the following text at the post-chat item-scoring stage: “*I chose to write my own focus before the chat and I was supported the way I hoped.*”

[insert figure 1 here]

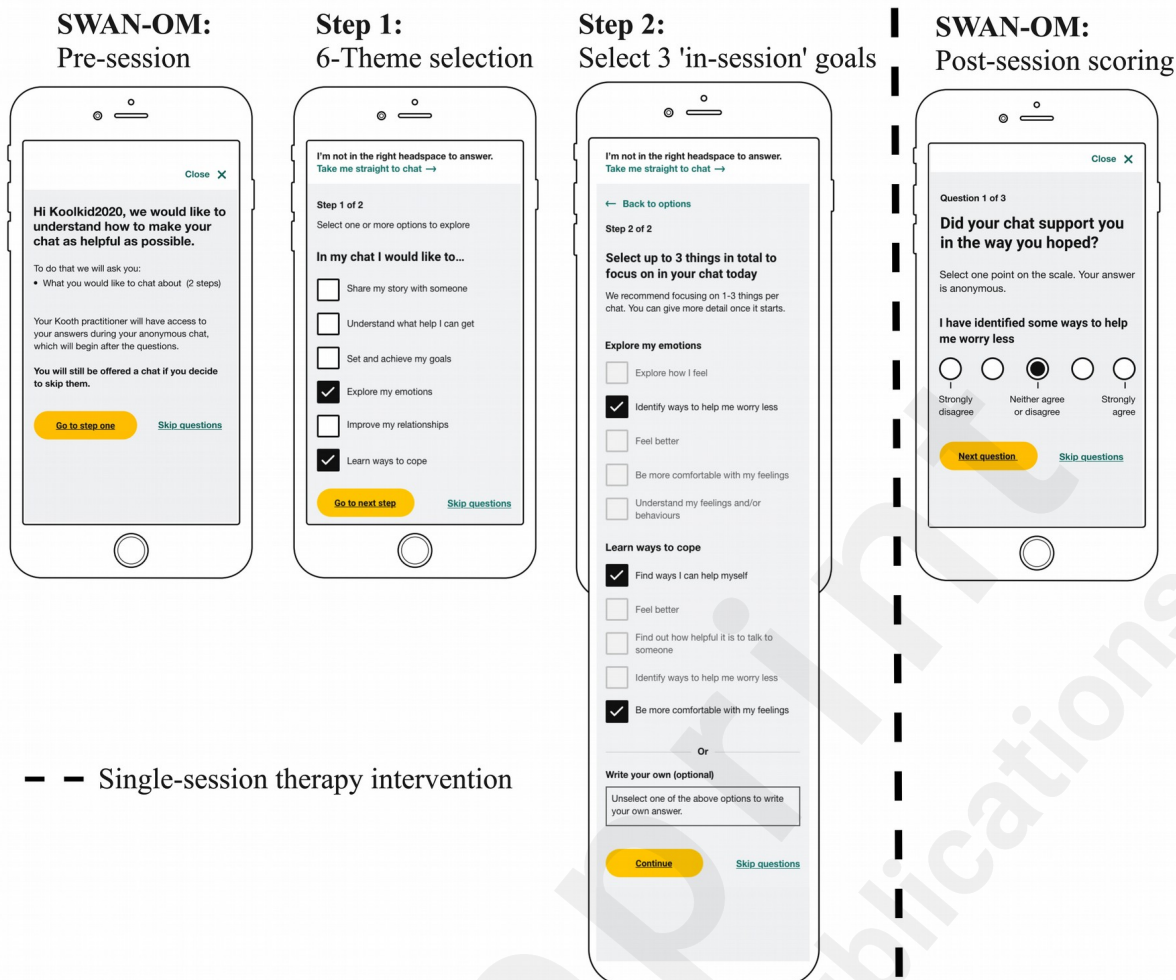


Figure 1- Smartphone wireframes of SWAN-OM at Kooth.com

Positive and Negative Affect Schedule

The Positive and Negative Affect Schedule (PANAS)⁴⁷ was selected as an instrument due to its wide use in the research literature and the immediacy of measuring emotional states, as a proxy for quality of the SST therapeutic encounter measured before and after the session. The PANAS is a measure of affect that can be used with children aged 6-18⁴⁸. It has been extensively validated in different languages showing excellent psychometric properties, and it is the most widely used instrument to measure affect.⁴⁹ It includes 10-items assessing affect in the present moment and is divided across two sub-scales: Positive Affect (PANAS-PA) and Negative Affect (PANAS-NA). The following 10 feelings are used: Sad, Happy, Scared, Miserable, Cheerful, Proud, Afraid, Joyful, Mad (Angry), and Lively. PANAS assesses a person's positive and negative trait affect using a 5-point scale (1= "very slightly or not at all"; 5= "extremely"). The schedule has been validated in general and clinical populations,⁴⁷ which makes it a suitable instrument to use in a digital service where clinical and non-clinical populations are accessing the service. In the current study, PANAS was administered at Time 1 (pre-chat) and Time 2 (post-chat) before and after the single-session intervention. In our sample, both Negative and Positive sub-scales showed good internal consistency at Time 1 and Time 2: Negative subscale, $\alpha = .75$ (Time 1) and $\alpha = .84$ (Time 2); Positive subscale, $\alpha = .82$ (Time 1) and $\alpha = .90$ (Time 2).

Youth Counselling Impact Scale

The insights subscale of the Youth Counselling Impact Scale (YCIS)^{50,51} aligns to measure the impact of the session, or perceptions of having made progress within a session, which are also associated

with clinical treatment outcomes.⁵²⁻⁵⁴ The YCIS is a 6-item scale assessing young people's perceptions of the impact of individual mental health counselling sessions have on their thoughts, feelings, and behaviors. The YCIS was designed to be used with people aged 11 to 18 and showed good psychometric properties.⁵⁵ The original instrument is divided into two sub-scales: Insights, assessing impact immediately after the session, and Change, measuring the impact of the two weeks following the session.⁵⁶ The 3-item Insight sub-scale was selected for the study; in this sub-scale, young people indicate how well each item reflected the outcome of the session using a 5-point Likert scale (1= "Not at all"; 5= "Totally"). The Insight subscale showed good levels of internal consistency Cronbach's alpha $\alpha = .86$ for the study.

Experience of Service Questionnaire

The Experience of Service Questionnaire (ESQ, formerly CHI-ESQ)⁵⁷ was selected to compare with SWAN-OM after the single-session took place, to understand the satisfaction that the user had with the service and with the care provided; the instrument is used across mental health CAMHS services for quality and experience assessments and evaluations. The ESQ is a 12-item questionnaire that can be used with young people aged 12 to 18 to measure feedback about a mental health service and, specifically satisfaction with care and with the environment; it is commonly used for CAMHS services to measure the subjective experience of satisfaction.⁵⁸ The 9-item "Satisfaction with Care" subscale was used in the study. Some items were adapted and rephrased to match the context; for example, "*Overall, the help I have received on Kooth (the service) is good*". Young people rated each item on a 4-point Likert scale (1= "Certainly true"; 3= "Not true"). Items scored as 4= "Don't know" were treated as missing data. Responses were reversed before the total score was calculated so that a higher ESQ score indicated higher service satisfaction. The internal consistency of the ESQ Satisfaction with Care subscale showed good levels, with a Cronbach's alpha $\alpha = .87$.

SWAN-OM evaluation procedure

The instrument was implemented and evaluated for a period of six months (Jan 2021 - June 2021) at Kooth.com (<https://kooth.com>), a web-based therapy service based in the UK, via synchronous text messaging. The service is anonymous at the point of entry and provides person-centered text-based SST and brief interventions, free and accessible with an internet connection to children and young people who wish to register and access the service.

During the evaluation period, the SWAN-OM was implemented for SSTs. A total of 120 practitioners from the web-based therapy service were recruited and trained to administer SWAN-OM at Kooth.com. Each practitioner attended a training session of 60 minutes and was provided with a manual containing guidance on how to use the instrument in the platform, internal clinical governance procedures, SST relevant literature, and frequently asked questions about the instrument and the research study. Ad hoc support was provided through instant messaging software by the research group to each practitioner. All practitioners involved in the study were part of the service workforce. Therefore, practitioners were in training or obtained their counselling or clinical qualifications as mental health practitioners. The service holds a pluralistic view on their training and therapeutic background, but all within encompassing a person-centered framework to deliver care.

The SST intervention was delivered over a 40- to 60-minute text-based chat in the online synchronous messaging system at the web-based service. The SST broad aims were to engage, conduct a brief assessment, and meet the needs of young people where possible. Brief risk assessments and safeguarding protocols were prioritized above those aims as part of service provision. The approach of SST delivery by practitioners within the service was pluralistic with a

broad range of therapeutic orientations. The SST interventions delivered during the evaluation considered the brief-intervention mindset and its blend with traditional approaches to counselling,⁵⁹ in addition to the already established evidence-base on SST.^{13,18,60}

The SWAN-OM was administered when practitioners clicked a button in the platform to launch the questionnaires in the front-end view of the user. The battery of instruments was administered at the same point in time, before the chat (Time 1: pre-SST; PANAS and SWAN-OM) and after the chat (Time 2: post-SST; PANAS, YCIS, ESQ, and SWAN-OM).

Young people could skip the measures if they wished to at the time of accessing the service. The practitioner was able to access the item selection of SWAN-OM at Time 1 and then start the SST when ready. Following the end of their SST chat with a practitioner, young people were asked to complete the post-session measures. Individuals who skipped the administration of the questionnaires at Time 1 were not presented with the other measures at Time 2. Unique identifiers in the platform were used to collect demographic and characteristic variables and relevant information for the study and construct the dataset containing the SSTs for analysis.

Statistical Analysis

The dataset was cleaned and analyzed using R open-source package and language for statistical computing.⁶¹ Completion rates for the SWAN-OM and the other comparator measures alongside descriptive statistics were calculated for the study. To investigate the psychometric properties of SWAN-OM, different analytical strategies were followed, taking into consideration the structure of the instrument and the sample size availability and suitability of the analyses conducted during the evaluation. Pair-wise comparisons across the administered measures were performed to investigate the concurrent validity of the instrument using correlations, and regression models were computed to understand participant characteristics and item-selection responses for SWAN-OM.

Analysis considerations

Considerations were made regarding the analysis given the nature and structure of SST and the instrument. The SST outcomes based on the patient expectations will not be repeated in a second session (even if the patient did return), as the “in-session” goals for the SST can only take place in the present moment and may not be continued in the future. Therefore, while high face validity may be achieved, it may be a challenge to obtain traditional forms of internal and structural psychometric validity, as re-testing the measure will not measure the same “Wants” or “Needs” again and therefore would be measuring a different SST outcome.

It was important to understand how the personal characteristics of users affected their selection of “Wants” and “Needs” for the session. This is especially important for the SWAN-OM, as item selection is shown through two-stage logic filtering of items (to be shown) based on the young person’s theme selection. The items initially selected in SWAN-OM should help to identify if the instrument works in similar ways across a diverse cohort of young people, as well as to inform the most frequent “Wants” and “Needs” for SST requested by children and young people at the service, and if their characteristics predict the type of items people will select for their SST.

Predicting item selection

Two approaches were used to assess the association between young people’s demographic characteristics and item selection on the SWAN-OM. The first approach used multilevel logistic regressions, predicting theme or item selected by young people’s Gender, Age, Ethnicity, and Presenting concerns. Multilevel regression accounts for the nested nature of the data whereby young people can access several chats and may complete SWAN-OM more than once. Log-likelihood ratio

tests were used to determine whether adding demographic variables available to the model explained a significantly larger amount of variance. We used the larger category as a reference (for example, females were used as a reference when exploring the impact of Gender in item selection).

The second approach used Chi-square tests to examine demographic characteristics associated with item selection to complement the multilevel regressions, especially when sample sizes were too small for specific themes or items. This occurred when small cell sizes were found in the different demographic characteristic variables and in the item selected vs. item not selected (which would be anticipated given young people were invited to select up to three items per session). A sub-sample of 930 young people attending to 1131 was used to analyze the data; cases were discarded if young people had missing data for their demographic variables.

Concurrent validity

Concurrent validity was explored by examining correlations between scores on individual SWAN-OM items and total scores on the comparator measures (YCIS, ESQ, and PANAS), as well as individual items from YCIS and ESQ. Overall, the correlation analysis amongst measurements will provide evidence for concurrent validity of the SWAN-OM covering the domains for measuring quality of care⁶²; further comparison between PANAS at Time 1 and 2 will provide further construct validity on affect changes after SST.

The Shapiro-Wilk test of normality confirmed non-normal distribution (significant at $p < .05$); therefore, Spearman-ranked correlations were used for the analysis. Each pairwise correlation was calculated using data available for each SWAN-OM item (Supplementary Table 7) and each comparator measure. In instances where individuals had completed SWAN-OM and comparator measures more than once, within-individual average scores were computed prior to calculating the correlations. Pairwise correlations with less than 20 cases were not reported. Only cases where SWAN-OM was completed for Time 1 and Time 2 were included, comprising a sub-sample of 577 young people attending 696 chats that were used for these analyses.

Results

Completion rates and descriptive statistics

Most young people accessing the service for a single-session selected SWAN-OM items before their session (SWAN-OM pre-chat item selection ($n=1503$, 78.69%). After their chat session, 696 young people completed the measure scoring (46.31%), while 112 (7.45%) users skipped the measure scoring and 695 (46.24%) left the chat before the measures scoring could be presented during the evaluation. When considering only individuals who saw the SWAN-OM and were able to complete their chat intervention, the vast majority completed the measure scoring after their chat session ($N=969$, 86.14%), providing a good indicator of acceptability and completion rates for the instrument.

Completion rates from the other comparative instruments were high in the pre-chat item selection, with more than 73% of young people selecting items on the measure. After the session, measure engagement rates decreased overall, with less than 24% of the sample completing the PANAS, YCIS, and ESQ (Supplementary Table 1).

The SWAN-OM allows young people to select up to three items from a list of 21-items by previously selecting in between six-theme categories to display a group of items (between two to five), there is no limitation on the number of themes that young people can select, allowing in that scenario to display up to the 21-items for choice. Young people selected on average and 2.47 items ($SD = .76$; median = 3).

The frequency of item selection within the overall sample ranged selection rates from 0.51% to 11.61% (Table 2). The frequency analysis identified popular items, such as “*Feel better*” ($N= 431$; 11.61%), “*Find ways I can help myself*” ($N= 411$; 11.07%), and unpopular amongst the sample such as “*Feel safe in my relationships*” ($N= 53$; 1.43%) or “*Learn the steps to achieve something I want*” ($N= 58$; 1.56%).

Table 2. SWAN-OM item selection, completion, and average scores. ^{a,b}

Theme	N	Item	Pre-chat item selection		Post-chat item scoring			
			n	%	n	%	Mean	SD
A		Total	470	12.66%				
	1	Be comfortable asking for help outside Kooth	276	7.43%	134	7.89%	0.54	1.00
	2	Find information about how to keep myself safe	194	5.22%	95	5.59%	1.18	0.87
B		Total	991	26.69%				
	3	Feel listened to	333	8.97%	138	8.13%	1.58	0.71
	4	Talk about something I haven't told anyone before	187	5.04%	71	4.18%	1.32	1.19
	5	Identify a solution to a problem in my life	179	4.82%	75	4.42%	0.96	0.85
	6	Be able to open up to people in my life	169	4.55%	67	3.95%	0.58	0.99
	7	Find out how useful it is to talk to someone	70	1.89%	27	1.59%	1.44	0.85
	8	Feel safe in my relationships	53	1.43%	20	1.18%	0.5	0.95
C		Total	150	4.04%				
	9	Learn how to feel better	92	2.48%	44	2.59%	1.00	1.08
	10	Learn the steps to achieve something I want	58	1.56%	35	2.06%	1.14	1.03
D		Total	967	26.04%				
	11	Understand my feelings and/or behaviours	373	10.05%	159	9.36%	0.72	0.87
	12	Identify ways to	233	6.28%	112	6.60%	0.96	0.92

		help me worry less						
	13	Explore how I feel	225	6.06%	114	6.71%	1.16	1.02
	14	Be more comfortable with my feelings	136	3.66%	64	3.77%	0.98	0.88
E		Total	215	5.79%				
	15	Explore difficulties in my relationships	78	2.10%	37	2.18%	0.78	1.23
	16	Identify solutions to improve my relationships	75	2.02%	38	2.24%	0.76	1.02
	17	Learn how to manage conflict with others	43	1.16%	22	1.30%	1.14	0.77
	18	Learn how to relate to other people	19	0.51%	12	0.71%	0.33	0.89
F		Total	842	22.68%				
	19	Feel better	431	11.61%	212	12.49%	0.90	0.94
	20	Find ways I can help myself	411	11.07%	186	10.95%	1.06	0.96
NA	21	Free text	78	2.10%	36	2.12%	0.94	1.17

^a Frequency of SWAN-OM item selection is provided for pre-chat item selection, percentages at Time 1 are calculated based on total ($n = 3713$) items selected during 1503 chats; frequency, mean, and standard deviation (SD) are provided for SWAN-OM items at Time 2, percentages at post-chat item scoring, are calculated based on total ($n = 1698$) items rated during 696 chats.

^b Themes: A: 'Understand what help I can get'; B: 'Share my story with someone'; C: 'Set and achieve my goals'; D: 'Explore my emotions'; E: 'Improve my relationships'; F: 'Learn ways to cope'. N: item number.

Predicting SWAN-OM item selection

There was an acceptable amount of data to compute hierarchical logistic regressions on five selected items. For the items "Feel better" and "Identify a solution to a problem in my life", no significant differences between demographic variables were found predicting the selection of these two items ($p > .05$).

Hierarchical logistic regressions for the rest of the items accounted for some differences in items selection prediction across demographic characteristics. The item "Find ways I can help myself" showed that demographic characteristics of young people significantly predict the selection of this item (Table 3). Compared to females, males were more likely to select this item (OR = 0.51, CI = [0.3-0.87]; overall model significance: $\chi^2(5) = 8.95, p = .03$). In regards to Ethnicity, Black African, Black Caribbean and Black British individuals were significantly less likely to select the item when compared to White young people (OR = .27; CI = [0.08-0.86]; $\chi^2(11) = 22.81, p < .001$); this difference on item selection also was found for people from any other ethnic group and those who did not provide Ethnicity data (OR = 0.12, CI = [0.03-0.53]). Asian and Asian British were also less likely to select "Find ways I can help myself" albeit significant at a trend level (OR = .56; CI = [0.28-1.11]), as well as those young people experiencing presenting concerns around risk within the

service (OR = 0.56, CI = [0.32-0.98]; $\chi^2(16) = 9.26, p = .10$).

Table 3. Results of the hierarchical multi-level logistic regressions for SWAN-OM Item 20: “*Find ways I can help myself*”.^{ab}

SWAN-OM Item 20	OR	p-value	95% CI
Gender *			
Male vs. Female	0.51*	0.013	0.30-0.87
Agender vs. Female	0.59	0.229	0.25-1.4
Gender Fluid vs. Female	1.14	0.752	0.50-2.64
Age			
10 to 14 vs. 15 to 19	0.79	0.181	0.55-1.12
20 and above vs. 15 to 19	0.88	0.667	0.50-1.57
Ethnicity *			
Asian/Asian British vs White	0.56 ⁺	0.094	0.28-1.11
Black African/Caribbean/Black British vs. White	0.27*	0.027	0.08-0.86
Mixed multiple ethnic group vs. White	0.69	0.256	0.36-1.31
Any other/Unknown vs. White	0.12*	0.005	0.03-0.53
Presenting concerns^b			
Mental health	0.97	0.889	0.65-1.45
External	0.90	0.542	0.66-1.25
Suicidal thoughts / Self-harm	1.23	0.211	0.89-1.72
Risk	0.56 ⁺	0.042	0.32-0.98
Physical / Other	0.60	0.153	0.30-1.21

^a Data collected from 910 young people attending 1131 chats. OR = odds ratio. CI = confidence interval. For variables to be significant, the model in which they were entered and the effect of the variable within the model both had to meet the $p < .05$ threshold. Model significance is indicated in parentheses next to the predicting variable. (*) indicate a significant effect at $p < .05$. (+) indicate an effect at trend level ($p < .1$). ^b Adding presenting concerns did not reach $p < .05$ significant level for model fit; therefore, although the effect of risk was significant, we have interpreted this as trend level.

For the item “*To understand my feelings and/or behaviours*” we found differences in Age and Gender in item selection prediction (Table 4). Results shown that young people aged between 10 to 14 were significantly less likely to select this item when compared to 15 to 19 years old (OR = .58, CI = [0.38-0.88]; $\chi^2(7) = 7.88, p = .02$). Males were also less likely to select this item compared to females, albeit at trend level (OR = .54, CI = [0.30-0.96]; $\chi^2(5) = 5.76, p = .12$). Other characteristics did not show any significant power to predict the selection of this item.

Table 4. Results of the hierarchical multi-level logistic regressions for SWAN-OM Item 13: “*Understand my feelings and/or behaviours*.”^a

SWAN-OM Item 13	OR	p-value	95% CI
Gender⁺			
Male vs. Female	0.54 ⁺	0.035	0.30-0.96
Agender vs. Female	0.59	0.293	0.22-1.58
Gender Fluid vs. Female	0.76	0.591	0.29-2.04
Age *			
10 to 14 vs. 15 to 19	0.58*	0.010	0.38-0.88
20 and above vs. 15 to 19	1.09	0.782	0.58-2.05
Ethnicity			

Asian/Asian British vs White	1.00	1.000	0.48-2.07
Black African/Caribbean/Black British vs. White	0.94	0.916	0.33-2.73
Mixed multiple ethnic group vs. White	0.97	0.933	0.48-1.98
Any other/Unknown vs. White	1.93	0.157	0.78-4.82
Presenting concerns			
Mental health	1.17	0.496	0.74-1.84
External	0.90	0.556	0.63-1.29
Suicidal thoughts / Self-harm	1.00	0.997	0.69-1.45
Risk	0.63	0.140	0.34-1.16
Physical / Other	1.13	0.724	0.57-2.27

^a Data collected from 910 young people attending 1131 chats. OR = odds ratio. CI = confidence interval. For variables to be significant, the model in which they were entered and the effect of the variable within the model both had to meet the $p < .05$ threshold. Model significance is indicated in parentheses next to the predicting variable. (*) indicate a significant effect at $p < .05$. (+) indicate an effect at trend level ($p < .1$).

Finally, for the item “*Feel listened to*” hierarchical regression results (Table 5) indicated that young people experiencing risk as a presenting concern within the service were significantly more likely to select this item (OR = 2.51, [1.28-4.9]; $\chi^2(16) = 11.83$, $p = .04$) compared to those without risk issues. A trend was found to those between 10 to 14 more likely to select this item compared to those aged 15 to 19 years old (OR = 1.59, CI = [1.03-2.45]; $\chi^2(7) = 4.95$, $p = .08$).

Table 5. Results of the hierarchical multi-level logistic regressions for SWAN-OM Item 6: “*Feel listened to*”. ^a

SWAN-OM Item 6	OR	p-value	95% CI
Gender			
Male vs. Female	0.71	0.267	0.38-1.31
Agender vs. Female	1.43	0.466	0.55-3.7
Gender Fluid vs. Female	1.62	0.342	0.6-4.42
Age ⁺			
10 to 14 vs. 15 to 19	1.59 ⁺	0.037	1.03-2.45
20 and above vs. 15 to 19	0.86	0.690	0.41-1.79
Ethnicity			
Asian/Asian British vs White	1.57	0.251	0.73-3.40
Black African/Caribbean/Black British vs. White	2.00	0.228	0.65-6.16
Mixed multiple ethnic group vs. White	0.95	0.888	0.43-2.06
Any other/Unknown vs. White	0.90	0.840	0.33-2.49
Presenting concerns [*]			
Mental health	1.01	0.958	0.63-1.64
External	1.26	0.261	0.84-1.89
Suicidal thoughts / Self-harm	0.93	0.709	0.61-1.39
Risk	2.51 [*]	0.007	1.28-4.90
Physical / Other	1.64	0.199	0.77-3.48

^a Data collected from 910 young people attending 1131 chats. OR = odds ratio. CI = confidence interval. For variables to be significant, the model in which they were entered and the effect of the variable within the model both had to meet the $p < .05$ threshold. Model significance is indicated in parentheses next to the predicting variable. (*) indicate a significant effect at $p < .05$. (+) indicate an effect at trend level ($p < .1$).

For the rest of items of SWAN-OM, Chi-square test were computed for each item and a single demographic variable. These analyses did not consider the repeated measures within the sample of young people and co-variance between demographic characteristics, for presenting concerns

comparisons between groups were calculated between those with the presenting concerns against those who did not present that problem in the sample.

The Chi-square comparisons on presenting concerns (Supplementary Table 2) found that young people who experience presenting concerns around risk (e.g. victim of crime, trauma) were more likely to select the item “*Feel safe in my relationships*” ($\chi^2(1) = 11.33, p < .001$). Young people experiencing suicidal thoughts and/or self-harm were less likely to select the item “*Identify solutions to improve my relationships*” ($\chi^2(1) = 11.47, p < .001$) and more likely to select “*Find information about how to keep myself safe*” ($\chi^2(1) = 18.62, p < .001$).

Regarding the Chi-square comparisons computed for the rest of demographic variables, it was found that young people aged 10 to 14 were significantly more likely to select the items “*Be able to open up to people in my life*” ($\chi^2(2) = 17.98, p < .001$) and “*Talk about something I haven’t told anyone before*” ($\chi^2(2) = 14.67, p < .001$). The older age group (20 and above) were more likely to select the item “*Identify solutions to improve my relationships*” ($\chi^2(2) = 22.77, p < .001$).

In terms of Ethnicity, while there appeared to be an effect of young people from Black African/Caribbean and Black British backgrounds being significantly more likely to select the item “*Learn how to relate to other people*” (very small; $\chi^2(4) = 22.30, p < .001$) the cell sizes were very small meaning this finding is unreliable. Young people who did not provide Ethnicity information and those who identified as part of any other ethnic background were less likely to select the item “*Find ways I can help myself*” ($\chi^2(4) = 21.18, p < .001$). No significant differences in item selection for Gender were found (Supplementary Table 3).

Variability of SWAN-OM item scores

The variability of item scoring was explored by looking at the population average of the selected items at post-chat item scoring stage when the SWAN-OM was considered completed after a single-session intervention.

Average scores of SWAN-OM items at the post-chat item scoring stage ranged between 0.33 (SD=.89) for the item “*Learn how to relate to other people*” and 1.58 (SD= .71) for the item “*Feel listened to*” (Table 2). Some items had larger positive scores on average, perhaps because they captured aims that could be more easily implemented during a single-chat intervention as part of the “in-session” goals within the online service. For instance, “*Feel listened to*” and “*Find out how useful it is to talk to someone*” were amongst the best scored items during the study. Other items, such as “*Be able to open up to people in my life*”, “*Feel better*” or “*Explore difficulties in my relationships*”, had on average lower scores across the sample. Interpretation of these scores need to account for the variation in the number of cases used to calculate the average scores, which varied from 12 to 212 chats for the available 20 items.

Concurrent validity

Pairwise Spearman-ranked correlations were calculated for individual items and total scores for the YCIS, ESQ, and PANAS to explore the hypothesis how these instruments will correlate with SWAN-OM. Pairwise correlations with less than 20 cases were not reported. Some of the results shown (Table 6) should take into consideration the large number of significance test calculated and the relatively small number of paired cases available for some of the correlation tests (Supplementary Table 4).

Analysis reveal that nine SWAN-OM items were positively correlated with the ESQ total scores. The items “Be comfortable asking help outside Kooth” ($rs(64) = .41, p < .001$), “Understand my feeling and/or behaviours” ($rs(86) = .39, p < .001$), “Feel better” ($rs(109) = .48, p < .001$), and “Find ways I can help myself” ($rs(98) = .42, p < .001$) showed a significant positive association with ESQ total scores at $p < .001$. In addition, five items from SWAN-OM showed a significant positive association with a $p < .05$ significance level on ESQ total scores.

Thirteen items from SWAN-OM showed a significant positive correlation with total scores of the YCIS insight subscale. Items such as “Learn the steps to achieve something I want” ($rs(22) = .76, p < .001$), the personalized “free-text” option ($rs(22) = .7, p < .001$), and “Explore how I feel” ($rs(70) = .69, p < .001$) were amongst the ones with highest Spearman-rank coefficients showing large associations. Furthermore, additional four items from SWAN-OM correlated positively with YCIS total scores at $p < .05$ significance level.

Pairwise correlations were also computed between individual SWAN-OM items and individual items of the YCIS and ESQ. Similar trends were observed between correlations of total scores and individual items from the ESQ and YCIS (Supplementary Table 5-6).

Six items from SWAN-OM showed a significant negative correlation with PANAS Negative Affect subscale (PANAS-NA) total scores. The items were “Explore how I feel” ($rs(70) = -.44, p < .001$), “Understand my feelings and/or behaviours” ($rs(87) = -.39, p < .001$), “Feel better” ($rs(134) = -.30, p < .001$), “Find ways I can help myself” ($rs(114) = -.30, p < .001$), “Talk about something I haven't told anyone before” ($rs(43) = -.34, p < .05$), and “Feel listened to” ($rs(80) = -.24, p < .05$). Fourteen items from SWAN-OM showed statistically significant correlations with PANAS Positive Affect (PANAS-PA) subscale total scores. The item “Learn how to feel better” ($rs(22) = .72, p < .001$) showed the highest significant association and “Find information about how to keep myself safe” ($rs(66) = .24, p < .05$) the lowest coefficient amongst the items that showed an association.

Young people experienced a significant improvement in positive affect following their SST, represented by changes in the scores before and after the session in PANAS, with an average change of $M=3.08$ ($SD = 3.95$) between Time 1 and Time 2, $t(452) = 16.58, (p < .001)$. They also experienced a significant reduction in negative affect following their chat, with an average change of $M=-4.03$ ($SD = 3.61$), $t(452) = -23.71, (p < .001)$. Positive affect significantly changed between pre and post SST on average total scores. Negative affect significantly reduced after SST when comparing pre and post score of PANAS.

Table 6. Pairwise Spearman-ranked correlations between individual SWAN-OM items and YCIS, ESQ, and PANAS total scores. ^{a-c}

N	Item	ESQ	YCIS	PANAS-NA	PANAS-PA
1	Be comfortable asking for help outside Kooth	.413 $p = .001$.495 $p < .001$	-.186 $p = .113$.356 $p = .002$
2	Find information about how to keep myself safe	.306 $p = .017$	0.452 $p < .001$	-.082 $p = .504$.239 $p = .049$
3	Feel safe in my relationships	--	--	--	--
4	Be able to open	.219	.643	-.199	.469

	up to people in my life	p = .199	p < .001	p = .200	p = .002
5	Talk about something I haven't told anyone before	.153 p = .379	.339 p = .023	-.338 p = .023	.35 p = .018
6	Feel listened to	.329 p = .005	.473 p < .001	-.238 p = .031	.309 p = .005
7	Find out how useful it is to talk to someone	--	--	--	--
8	Identify a solution to a problem in my life	.349 p = .025	.572 p < .001	-.151 p = .310	.412 p = .004
9	Learn how to feel better	.284 p = .212	.595 p = .002	-.367 p = .078	.717 p < .001
10	Learn the steps to achieve something I want	--	.764 p < .001	-.215 p = .313	.273 p = .197
11	Explore how I feel	.255 p = .060	.686 p < .001	-.445 p < .001	.372 p = .001
12	Be more comfortable with my feelings	.116 p = .55	.634 p < .001	-.105 p = .562	.372 p = .033
13	Understand my feelings and/or behaviours	.388 p < .001	.614 p < .001	-.389 p < .001	.35 p = .001
14	Identify ways to help me worry less	.266 p = .045	.681 p < .001	-.239 p = .053	.393 p = .001
15	Explore difficulties in my relationships	--	.496 p = .014	-.180 p = .400	.305 p = .147
16	Learn how to relate to other people	--	--	--	--
17	Learn how to manage conflict with others	--	--	--	--
18	Identify solutions to improve my relationships	--	.507 p = .023	.004 p = .987	.324 p = .164
19	Feel better	.483 p < .001	.684 p < .001	-.301 p < .001	.441 p < .001
20	Find ways I can help myself	.423 p < .001	.599 p < .001	-.301 p = .001	.423 p < .001
21	Free text	.483	0.700	-.380	0.609

$p = .023$ $p < .001$ $p = .067$ $p = .002$

^a Items with less than 20 cases were not reported.. Pairwise correlations are based on varying sub-samples depending on the data available.

^b N: SWAN-OM item number.

^c More information is available in Supplementary Table 4.

Discussion

In the context of brief psychotherapy and solution-focus interventions, there is a measurement deficit in demonstrating outcomes within the single-session therapeutic intervention. Evidence suggests that this type of intervention and service delivery can be effective in reducing waiting times and increasing access to psychotherapy and mental health support,³⁸ but literature examining what contributes to its clinical effectiveness are inconclusive³⁶. Further studies stress the importance of how a single point of engagement with mental health services appears as a frequent option when monitoring service engagement,⁶³ including 46% of children and young people mental health engaging services in the UK with only one appointment²⁷; note that this percentage will include treatment drop-outs or people who booked further appointments and were not planned as an SST.

Currently, despite work on outcome measures and examining the effectiveness of single-session intervention continues, even at a more rapid pace in the context of the COVID-19 pandemic and the new mental health needs emerging in the population.² Most of these efforts scope for targeting to specific mental health difficulties,²⁹⁻³¹ and none of the previously used instruments have been designed with an SST and pluralistic view of service delivery in mind. There is a fundamental challenge related to the short-term nature of the SST work, and the disconnect with the measurement of mental health difficulties, which are often measured over several time points, or across longer time periods, or where SST outcome measures have been developed; they tend to be solely symptom-based.^{32,64}

This study explored the validity of a new outcome measure for SST (SWAN-OM), which aims to provide a patient-reported outcome measure that captures the “Wants” and “Needs” of the single-session itself, and the associated achievement of these wants. This paper focused on examining concurrent and convergent validity, as well as importantly exploring how the measure is used by different individuals, by exploring demographic characteristics and item selection. By providing early evidence for instrument validation of the SWAN-OM, we hope that this measurement may also contribute to demonstrating the effectiveness of SST across services.

The novel design provides young people with the most approachable version of the instrument, given that the structural design was driven by young people’s participatory design ideas.⁴⁵ This is novel and exciting, as it puts the user at the forefront of the instrument’s design. Whilst the structure of the SWAN-OM measures means that traditional psychometric testing may not be appropriate, this offers the opportunity to examine this measure at an item level. Additionally, some considerations should be taken for psychometric evaluation and exploration of properties, particularly when attempting to provide test-retest reliability. Test-retest reliability is not applicable to the SWAN-OM instrument, no two SST are the same experience, as it is not reasonable to examine the repeated measurement of SST outcomes, given that the main SST principle is rooted in the possibility that the intervention will be the “one and only” encounter between practitioner and client.³ An advantage of this novel instrument use is that the SWAN-OM changes to fit the needs of the young person for every SST they attend. The SST outcomes are expected to change from one session to the next one; hence the SWAN-OM changes, too, with each administration, so each SWAN-OM is personalized to each SST,

focusing the measurement on the intervention impact rather on the information of the underlying factor structure that contributes to that impact. Despite this, there have been efforts to conceptualize these non-specific factors on SST,^{42,65} future advances in SST and psychotherapy should continue to investigate and unveil those factors so the effectiveness of SST can continue to improve alongside its popularity and impact to provide well-being to society.

Principal results

Completion Rates

Completion rates are of high importance when examining a new outcome measure. SWAN-OM at pre-chat item selection and item selection was high, with over three quarters (78.69%) of the young people who saw the measure completing it. Of the young people who viewed the post-session SWAN-OM, there was a high completion rate showing good acceptability of the measure (86.14%). This suggests young people found the measure reasonable to complete after an SST and is in line with previous research, which showed SWAN-OM had good levels of acceptability and face validity.⁴⁵

This is encouraging, as the two-stage structural design of the SWAN-OM was created as a way to organize information^{66,67} and reduce cognitive load⁶⁸ on the young people completing the instrument. Additionally, from this study, it appears that three items are enough to take forward to the session, as most young people selected between two and three items during the study. This is in line with what practitioners said during the instrument development⁴⁵; three goals were a manageable number of “in-session” objectives and set out clear expectations about what can be achieved in a single-session therapy encounter.

Item selection and variability

On an item level, Age, Ethnicity, and Gender significantly predict differences in some of the item-level selections by young people. For instance, young people who had *risk* related presenting concerns were more likely to select the item “*To feel listen to*”, highlighting the importance of distress disclosure,⁶⁹ and “*Feel safe in my relationships*”; this is in line with the presenting concerns risk grouping often covering abuse, trauma, and bullying.

Overall, despite differences seen at item level through exploring demographic characteristics, the SWAN-OM was designed for and was administered to a wide range of ages and backgrounds, with choice and a person-centered approach at the heart of the measure design. Therefore, the standardization in selection of all items was unlikely to be achieved across items; yet, we recommend further investigation to determine if the instrument should be adapted based on specific characteristics and cultural factors, especially for future cross-cultural validation studies.⁷⁰

In regard to the variability of scores for items, some SWAN-OM items, when selected, were scored lower on average than others; however, the overall average of “in-session” pre-defined goal scores was positive across all items in the study. This highlights the importance of specificity on measurement when dealing with change in an SST intervention, so change can be directly reported after the intervention. SWAN-OM scores the perception of change in individual “in-session” goals for the single-session, despite non-specific factor effects on practitioner skills, therapeutic alliance, and contextual factors of the session.⁷¹ The assumption from these results is that single-session outcomes included in SWAN-OM can be achieved in the context of SST. However, differences in average scores were found between items, leading to the question of whether certain items are realistically achievable in one session, e.g., “*Feel better*”, which suggests a change and maintenance of a positive emotional state of the individual, may be unrealistic from the input of one session, or

may not change when dealing with negative affect responses in psychotherapy.⁷² In addition, “*Explore difficulties in my relationship*” can be particularly difficult to address in one session alone, as interpersonal functioning can be difficult to describe or understand, with evidence that long periods of therapeutic input is required for clients to report changes in interpersonal functioning.^{73–75} These examples provide some indication for item modification or deletion and suggest further research to collect feedback from practitioners and children and young people in regard to the use of SWAN-OM, it also provides an initial indication how personal characteristics may influence the difficulty score in each item of the measure.

Concurrent Validity

Three different instruments and subscales were used to explore how well the measure interacts with other well-established standardized measures. Overall, we observed good concurrent validity between individual SWAN-OM items and the comparator measures (PANAS, YCIS, and ESQ), meaning that some, but not all, items were associated with negative and positive affect, experience or satisfaction, and impact of the session. The exploration between PANAS at Time 1 and 2 showed a significant change on positive affect and reduction of negative affect immediately after the single-session; this may be linked with the often reported high rates of satisfaction levels in SST and walk-in clinic studies,^{3,24,76} and highlights that affect changes are likely to occur as a result of an SST intervention.⁴

Limitations and wider Considerations

There are wider considerations when examining the findings regarding the psychometric properties of SWAN-OM. Some of these considerations are commonly encountered in psychometric testing research, others are related to the idiographic nature of one of the items, as well as the dynamic nature of the assessment and time of observation for an intervention, and other common challenges like sample size and population diversity were also found in the applied research context.⁷⁷ The nature of the measure focuses on measuring the SST intervention; therefore, the changes in scores will not be expected to be maintained across time when administered twice. Therefore, the psychometric consistency of SWAN-OM scoring could not be calculated through common test-retest approaches. Further, the item selection process in the instrument defines the construct of the SST, making internal consistency values and interpretation have less relevance.⁷⁸ Item selection was purposefully limited to young people being able to select between one to three items, impacting the volume of data per item. This limitation by design was determined by clinical judgement based on the expected number of objectives that can be covered in SST.⁴⁵

Moreover, the pairwise comparisons and correlations often had low sample sizes for certain items that were selected less commonly during the study. Yet, inter-item correlations were explored for those items with enough number of responses, showing overall significant correlations between items in line with the reliability scores of themes. Regarding items to focus the single-session, some preliminary exploration of the reliability of items for SWAN-OM has been examined by experts through content validity indexes during the instrument development process.⁴⁵

Taking this into consideration, the SWAN-OM is recommended to be used at an item level where each item is treated individually as part of SST, rather than at the theme level. This provides the granularity for the practitioners and young people to match up the “Wants” and “Needs” with session progress, experience, and structure. Items can be aggregated across a young person’s session, or across specific items between young people, to monitor SST appropriateness in supporting that specific “Want” or “Need”. Further investigation of practitioner perceptions of total scores, multiple administration scores, or even service-level data usability for each item and the total score of the instrument is required. SWAN-OM does not rate the perception of the clinician on what outcomes

were achieved, which is suggested to be the best stakeholder to discern the effectiveness and impact of SST to date. On the other hand,

SWAN-OM is the first ever patient-reported outcome measure designed for SST; further studies should look at how the clinician-reported and patient-reported measures interact when measuring SST. Finally, this study was conducted at one digital web-based mental health service, therefore restricting the generalizability of the results to other digital contexts or non-text-based SSTs. This is important, as the instrument was also developed through a service-specific program theory, making it difficult to interpret and define the individual constructs being measured⁷⁹ and ensure the items relate to wider services and represent the same common “Wants” and “Needs” from SST across children and young people. The results show some indication of questionnaire fatigue in line with previous studies,⁸⁰ and considerations should be made on balancing the number of instruments that can be used for comparisons between instruments when testing validity, especially in the digital context. Future research will explore response rates and item selection preferences in non-digital and non-text-based services.

Future Directions

Future research on the instrument’s ongoing validity should try to answer what works best for whom and under what conditions as an attempt to generalize its use and standardization. This is especially important due to the wide variability of practices that someone may observe in single-session therapeutic encounters.

We also find some items performed poorly or were rarely selected, suggesting some future item adaptations for those items. Therefore, we recommend making some items less absolute, as these items were selected less frequently and appear to be harder to achieve in a single-session and could be why they are being selected less frequently. An example of this would be changing the item “*Feel better*” to a less absolute version, perhaps more grounded on temporality and immediacy: “*To start to feel better.*” This is more in line with the outcomes that can be achieved in a single-session, and we expect this to result in a more equal spread of item selection. Therefore, future research needs to examine the inclusion of certain items or adaptations to the instrument. We recommend input from stakeholder groups, especially to examine the validity of the instrument in other walk-in services provided in the community.

Finally, with the increased need and popularity of SSTs, we recommend exploring the perception of meaningful change after a single-session and how this aligns with SWAN-OM outcome scores. A longitudinal-controlled study could then examine the maintenance and magnitude of change in SST for those who were administered with SWAN-OM in their sessions. This will also provide evidence towards how these Walk-in and SST service delivery can impact, at a large scale, waiting times and access across the wider mental health welfare system.³⁸

Conclusion

In this paper, we present evidence on the validity of a new outcome measure for SST (SWAN-OM), which is a patient-reported outcome measure that captures the “Wants” and “Needs” of the single-session therapy itself. The outcome and experience aspects measure the associated achievement of these “Wants” or “Needs” and the experiences of the session, from the perspective of the client (children and young people: 11 to 25 years).

This study showed positive results on some of the psychometric properties assessed when the data and the study design allowed it. Wider considerations have been discussed on the novel structure and two-stage logic of the measure. Nevertheless, there are indications of good construct validity by

using and comparing concurrent instruments that showed good reliability for the study, and further verification that changes in affect scores take place after the SST with differences found between the pre-chat and post-chat measurement during the evaluation period.

Our study also provides a way to compare and validate measures that may be required of a specific design and structure and acknowledges the challenges ahead to align new technological advancements in questionnaire development and personalization with psychometric properties and validation. This opens a door to the proliferation of measures which consist of a combination of nomothetic and idiographic items.⁴⁴ Outcome and experience measurement will continue to proliferate in digital contexts; it is important advances are made to assess the properties of these instruments and how they may be used outside the digital context and in face-to-face mental health services, as well as contribute to the knowledge and evidence base for SST.

Acknowledgements

Charlie Patterson, Georgia Sugarman, Felicity Moon, members of the product team at Kooth plc, and contributors to the development of the instrument. The Kooth Clinical Team, led by Dr. Hannah Wilson, who guided the development and ensured that clinical expertise was embedded in the instrument and research. The whole team of senior practitioners and counsellors at Kooth plc, led by Michelle Bracken and Victoria Lewis, who contributed to administering the tool and collecting feedback from its use. Thanks to Laura Lynch Becherer to support proofreading. Finally, Kate Dalzell and Luís Costa da Silva, for ongoing consultation and expertise relating to outcome and experience measures and psychometric testing.

Data Availability Statement

The data sets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

Santiago de Ossorno Garcia, Louisa Salhi, and Aaron Sefi are researchers employed and receive honorarium by Kooth plc. Julian Edbrooke-Childs, Jenna Jacob, and Florence Ruby work (or worked) on the CORC project, which encourages the use of idiographic and nomothetic outcomes and experience measures. Julian Edbrooke-Childs reports funding from NHS England and NHS Improvement outside of the submitted work; Julian Edbrooke-Childs and Jenna Jacob report funding from the Wellcome Trust outside of the submitted work. All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1964 and its later amendments. Informed consent was obtained from all participants for being included in the study when registered at Kooth.com.

Abbreviations

UK – United Kingdom

SWAN-OM – Session Wants And Needs Outcome Measure

SST – Single-Session Therapy

OAAT- One-At-A-Time

CAMHS – Children and Adolescent Mental Health Services

NHS – National Health System (UK)

PANAS – Positive Affect and Negative Affect Scale

PANAS-NA – PANAS negative affect subscale

PANAS-PA – PANAS positive affect subscale

YCIS – Youth Counselling Impact Scale
ESQ – Experience of Service Questionnaire
OR – Odd Ratio
SD – Standard Deviation
CI – Confidence Interval
MH – Mental Health
E - External
SS – Suicidal thoughts and Self-harm
R - Risk
PO – Physical or other problems
ONS – Office for National Statistics (UK)

References

1. Schleider J, Weisz J. A single-session growth mindset intervention for adolescent anxiety and depression: 9-month outcomes of a randomized trial. *J Child Psychol Psychiatry Allied Discip.* 2018;59(2):160-170. doi:10.1111/jcpp.12811
2. Schleider JL, Mullarkey MC, Fox KR, et al. A randomized trial of online single-session interventions for adolescent depression during COVID-19. *Nat Hum Behav.* 2022;6(2):258-268. doi:10.1038/s41562-021-01235-0
3. Talmon M. *Single-Session Therapy: Maximizing the Effect of the First (and Often Only) Therapeutic Encounter.* Jossey-Bass; 1990.
4. Hymmen P, Stalker CA, Cait CA. The case for single-session therapy: Does the empirical evidence support the increased prevalence of this service delivery model? *J Ment Heal.* 2013;22(1):60-71. doi:10.3109/09638237.2012.670880
5. Bauer S, Moessner M. Technology-enhanced monitoring in psychotherapy and e-mental health. *J Ment Heal.* 2012;21(4):355-363. doi:10.3109/09638237.2012.667886
6. Dryden W. Single-Session One-At-A-Time Therapy: A Personal Approach. 2020;41(3):283-301. doi:10.1002/anzf.1424
7. Dryden W. Providing help at the point of need: insights from single-session therapy. University and College Counselling, May 2020. Published July 2020. Accessed February 14, 2022. <https://www.bacp.co.uk/bacp-journals/university-and-college-counselling/may-2020/providing-help-at-the-point-of-need/>
8. Paul KE, Van Ommeren M. A primer on single session therapy and its potential application in humanitarian situations. *Intervention.* 2013;11(1):8-23. Accessed January 11, 2022. https://www.interventionjournal.com/sites/default/files/A_primer_on_single_session_therapy_and_its.3_0.pdf
9. McLeod J. Developing Pluralistic Practice in Counselling and Psychotherapy: Using What the Client Knows. *Eur J Couns Psychol.* 2013;2(1):51-64. doi:10.5964/EJCOP.V2I1.5
10. Hoyt MF, Bobele M, Slive A, Young J, Talmon M. Single-session therapy by walk-in or appointment: Administrative, clinical, and supervisory aspects of one-at-a-time services. In: Hoyt MF, Bobele M, Slive A, Young J, Talmon M, eds. *Single-Session Therapy by Walk-In or Appointment.* Routledge/Taylor & Francis Group; 2018:3-24. doi:10.4324/9781351112437-1
11. Young K. Change in the Winds: The Growth of Walk-In Therapy Clinics in Ontario, Canada. In: *Single-Session Therapy by Walk-In or Appointment.* 1st ed. Routledge; 2018:59-71. doi:10.4324/9781351112437-4
12. Weeks M, Zook-Stanley L. Walk-In Counseling Center of Minneapolis/St. Paul: The Magic

- of Our Model for Clients and Volunteers. In: *Single-Session Therapy by Walk-In or Appointment*. 1st ed. Routledge; 2018:135-146. doi:10.4324/9781351112437-9
13. Talmon M. When Less is More: Lessons from 25 Years of Attempting to Maximize the Effect of Each (and Often Only) Therapeutic Encounter. *Aust New Zeal J Fam Ther*. 2012;33(1):6-14. doi:10.1017/AFT.2012.2
 14. Chorpita BF. Commentary: Metaknowledge is power: envisioning models to address unmet mental health needs: reflections on Kazdin (2019). *J Child Psychol Psychiatry Allied Discip*. 2019;60(4):473-476. doi:10.1111/jcpp.13034
 15. Kazdin AE. Annual Research Review: Expanding mental health services through novel models of intervention delivery. *J Child Psychol Psychiatry*. 2019;60(4):455-472. doi:10.1111/jcpp.12937
 16. Bertuzzi V, Fratini G, Tarquinio C, et al. Single-Session Therapy by Appointment for the Treatment of Anxiety Disorders in Youth and Adults: A Systematic Review of the Literature. *Front Psychol*. 2021;12:3659. doi:10.3389/FPSYG.2021.721382/BIBTEX
 17. Schleider JL, Weisz JR. Little Treatments, Promising Effects? Meta-Analysis of Single-Session Interventions for Youth Psychiatric Problems. *J Am Acad Child Adolesc Psychiatry*. 2017;56(2):107-115. doi:10.1016/j.jaac.2016.11.007
 18. Schleider JL, Dobias ML, Sung JY, Mullarkey MC. Future Directions in Single-Session Youth Mental Health Interventions. *J Clin Child Adolesc Psychol*. 2020;49(2):264-278. doi:10.1080/15374416.2019.1683852
 19. Firth N, Barkham M, Delgado J, Allery K, Woodward J, O'cathain A. Socioeconomic Deprivation and Dropout from Contemporary Psychological Intervention for Common Mental Disorders: A Systematic Review. *Adm Policy Ment Heal Ment Heal Serv Res*. 2021;49:490-505. doi:10.1007/s10488-021-01178-8
 20. Young K, Jebreen J. Recognizing Single-Session Therapy as Psychotherapy. *J Syst Ther*. 2020;38(4):31-44. doi:10.1521/JSYT.2019.38.4.31
 21. Miller JK, Platt JJ, Conroy KM. Single-Session Therapy in the Majority World: Addressing the Challenge of Service Delivery in Cambodia and the Implications for Other Global Contexts. In: *Single-Session Therapy by Walk-In or Appointment*. Routledge; 2018:116-134. doi:10.4324/9781351112437-8
 22. Miller JK. Walk-In Single Session Team Therapy: A Study of Client Satisfaction. *J Syst Ther*. 2008;27(3):78-94. doi:10.1521/JSYT.2008.27.3.78
 23. Perkins R. The effectiveness of one session of therapy using a single-session therapy approach for children and adolescents with mental health problems. *Psychol Psychother Theory, Res Pract*. 2006;79(2):215-227. doi:10.1348/147608305X60523
 24. Harper-Jaques S, Foucault D. Walk-In Single-Session Therapy: Client Satisfaction and Clinical Outcomes. *J Syst Ther*. 2014;33(3):29-49. doi:10.1521/JSYT.2014.33.3.29
 25. Wolpert M, Harris R, Hodges S, et al. *THRIVE Framework for System Change*. (Payne C, Perry E, eds.). CAMHS Press; 2019. Accessed February 15, 2022. <https://www.annafreud.org/media/9254/thrive-framework-for-system-change-2019.pdf>
 26. National Collaborating Centre for Mental Health. The Improving Access to Psychological Therapies Manual. Published online 2018:1-15. <https://www.england.nhs.uk/wp-content/uploads/2018/06/the-iapt-manual.pdf>
 27. Edbrooke-childs J, Hayes D, Lane R, Liverpool S, Deighton J. Association between single session service attendance and clinical characteristics in administrative data. *Clin Child Psychol Psychiatry*. Published online 2021. doi:10.1177/13591045211002609
 28. Schleider JL, Dobias M, Sung J, Mumper E, Mullarkey MC. Acceptability and utility of an open-access, online single-session intervention platform for adolescent mental health.

- JMIR Ment Heal.* 2020;7(6). doi:10.2196/20513
29. Coverley CT, Garralda ME, Bowman F. Psychiatric intervention in primary care for mothers whose schoolchildren have psychiatric disorder. *Br J Gen Pract.* 1995;45(394):235-237. Accessed March 29, 2021. /pmc/articles/PMC1239227/?report=abstract
 30. McCambridge J, Strang J. The efficacy of single-session motivational interviewing in reducing drug consumption and perceptions of drug-related risk and harm among young people: results from a multi-site cluster randomized trial. *Addiction.* 2004;99(1):39-52. doi:10.1111/j.1360-0443.2004.00564.x
 31. Weersing VR, Jeffreys M, Do MCT, Schwartz KTG, Bolano C. Evidence Base Update of Psychosocial Treatments for Child and Adolescent Depression. *J Clin Child Adolesc Psychol.* 2017;46(1):11-43. doi:10.1080/15374416.2016.1220310
 32. Riddle. The Pediatric Anxiety Rating Scale (PARS): Development and Psychometric Properties. *J Am Acad Child Adolesc Psychiatry.* 2002;41(9):1061-1069. doi:10.1097/00004583-200209000-00006
 33. Chardon L, Bagraith KS, Robert JK, King RJ. Counseling activity in single-session online counseling with adolescents: An adherence study. *Psychother Res.* 2011;21(5):583-592. doi:10.1080/10503307.2011.592550
 34. Bagraith K, Chardon L, King RJ. Rating counselor–client behavior in online counseling: Development and preliminary psychometric properties of the Counseling Progress and Depth Rating Instrument. *Psychother Res.* 2010;20(6):722-730. doi:10.1080/10503307.2010.519739
 35. Campbell A, Blais M, Blagys MD, et al. Single-session approaches to therapy: Time to review. *Clin Psychol Psychother.* 2015;22(1):15-26. doi:10.2196/jmir.7725
 36. Courtnage A. Hoping for Change: The Role of Hope in Single-Session Therapy. *J Syst Ther.* 2020;39(1):49-63. doi:10.1521/JSYT.2020.39.1.49
 37. Stalker CA, Riemer M, Cait CA, et al. A comparison of walk-in counselling and the wait list model for delivering counselling services. *J Ment Heal.* 2016;25(5):403-409. doi:10.3109/09638237.2015.1101417
 38. Ewen V, Mushquash AR, Mushquash CJ, Bailey SK, Haggarty JM, Stones MJ. Single-session therapy in outpatient mental health services: Examining the effect on mental health symptoms and functioning. *Soc Work Ment Health.* 2018;16(5):573-589. doi:10.1080/15332985.2018.1456503
 39. Calvete E, Fernández-Gonzalez L, Orue I, et al. The Effect of an Intervention Teaching Adolescents that People can Change on Depressive Symptoms, Cognitive Schemas, and Hypothalamic-Pituitary-Adrenal Axis Hormones. *J Abnorm Child Psychol.* 2019;47(9):1533. doi:10.1007/S10802-019-00538-1
 40. White C, Shanley DC, Zimmer-Gembeck MJ, Walsh K, Hawkins R, Lines K. Outcomes of In Situ Training for Disclosure as a Standalone and a Booster to a Child Protective Behaviors Education Program. *Child Maltreat.* 2019;24(2):193-202. doi:10.1177/1077559518816877
 41. Twigg E, Barkham M, Bewick BM, Mulhern B, Connell J, Cooper M. The young person's CORE: Development of a brief outcome measure for young people. *Couns Psychother Res.* 2009;9(3):160-168. doi:10.1080/14733140902979722
 42. Bloom BL. Focused Single-Session Psychotherapy: A Review of the Clinical and Research Literature. *Br Treat Cris Interv.* 2001;1(1):75-86. doi:10.1093/BRIEF-TREATMENT/1.1.75
 43. Dryden W. *Single-Session Therapy (SST): 100 Key Points and Techniques.* 1st ed. (Dryden W, ed.). Routledge; 2018. doi:10.4324/9780429489648

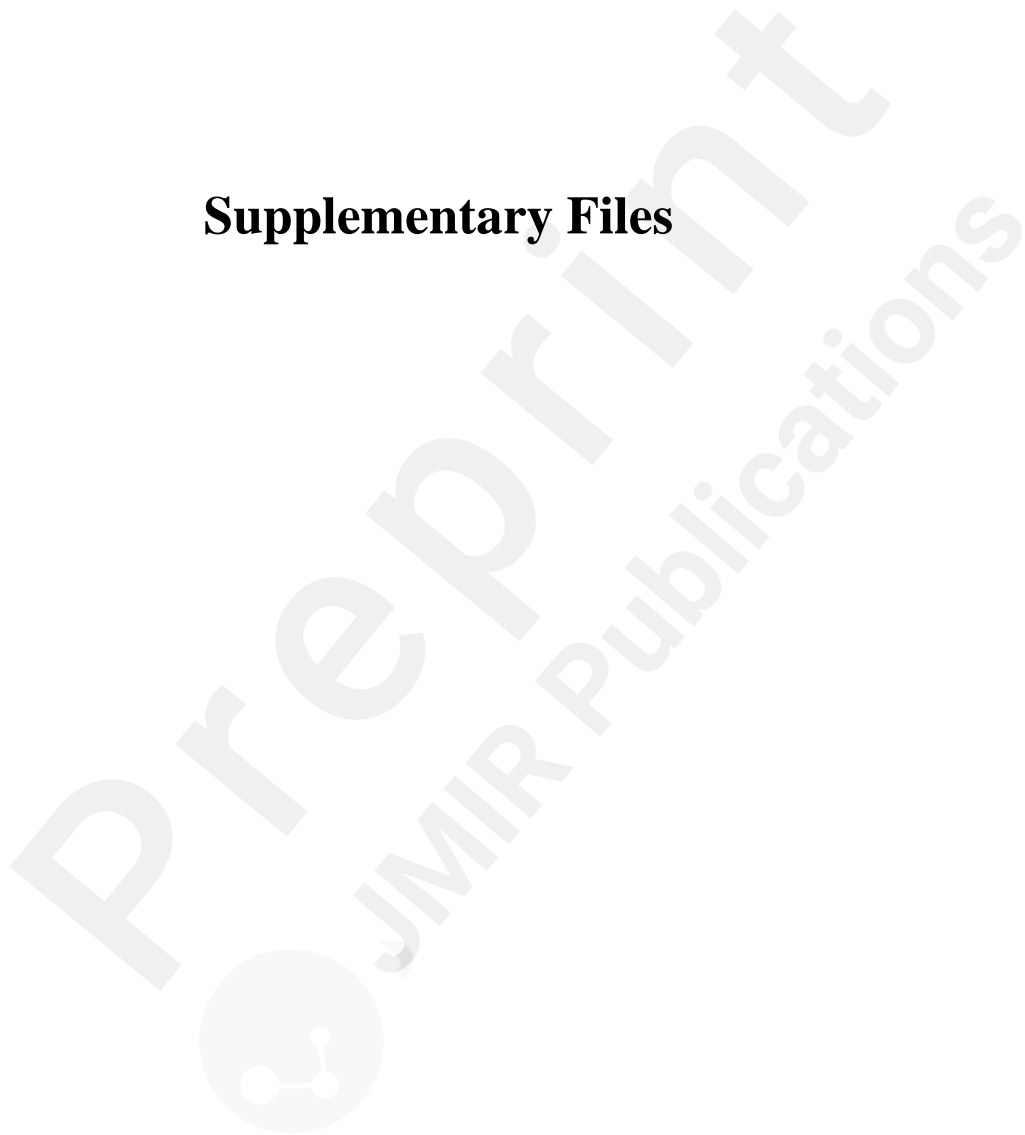
44. Jacob J, Edbrooke-Childs J, Silva LC da, Law D. Notes from the youth mental health field: Using movement towards goals as a potential indicator of service change and quality improvement. *J Clin Psychol*. Published online 2021. doi:10.1002/JCLP.23195
45. de Ossorno Garcia S, Salhi L, Sefi A, Hanley T. The Session Wants and Need Outcome Measure: The Development of a Brief Outcome Measure for Single-Sessions of Web-Based Support. *Front Psychol*. 2021;12:4900. doi:10.3389/FPSYG.2021.748145/BIBTEX
46. Office for National Statistics. Ethnic group, national identity and religion. 2016. Accessed February 15, 2022. <https://www.ons.gov.uk/methodology/classificationsandstandards/measuringequality/ethnicgroupnationalidentityandreligion>
47. Watson D, Clark LA, Tellegen A. Development and validation of brief measures of positive and negative affect: The PANAS scales. *J Pers Soc Psychol*. 1988;54(6):1063-1070. doi:10.1037/0022-3514.54.6.1063
48. Ebesutani C, Regan J, Smith A, Reise S, Higa-McMillan C, Chorpita BF. The 10-item Positive and Negative Affect Schedule for Children, Child and parent shortened versions: Application of item response theory for more efficient assessment. *J Psychopathol Behav Assess*. 2012;34(2):191-203. doi:10.1007/S10862-011-9273-2/TABLES/5
49. Díaz-García A, González-Robles A, Mor S, et al. Positive and Negative Affect Schedule (PANAS): Psychometric properties of the online Spanish version in a clinical sample with emotional disorders. *BMC Psychiatry*. 2020;20(1):1-13. doi:10.1186/S12888-020-2472-1/TABLES/8
50. Bickman L, Athay MM, Riemer M, et al. *Manual of the Peabody Treatment Progress Battery PTPB*. (Bickman L, Athay MM, Riemer M, et al., eds.). Vanderbilt University; 2010. Accessed February 14, 2022. https://peabody.vanderbilt.edu/docs/pdf/cepi/ptpb_2nd_ed/PTPB_2010_Entire_Manual_UPDATE_31212.pdf
51. Riemer M, Athay MM, Bickman L, Breda C, Kelley SD, Vides De Andrade AR. The Peabody Treatment Progress Battery: History and Methods for Developing a Comprehensive Measurement Battery for Youth Mental Health. *Adm Policy Ment Health*. 2012;39(0):3. doi:10.1007/S10488-012-0404-1
52. Kolden GG, Chisholm-Stockard SM, Strauman TJ, Tierney SC, Mullen EA, Schneider KL. Universal session-level change processes in an early session of psychotherapy: path models. *J Consult Clin Psychol*. 2006;74(2):327-336. doi:10.1037/0022-006X.74.2.327
53. Kolden GG. Change in early sessions of dynamic therapy: Universal processes and the generic model of psychotherapy. *J Consult Clin Psychol*. 1996;64(3):489-496. doi:10.1037/0022-006X.64.3.489
54. Kolden GG, Howard KI. An empirical test of the generic model of psychotherapy. *J Psychother Pract Res*. 1992;1(3):225-236. Accessed February 14, 2022. <http://www.ncbi.nlm.nih.gov/pubmed/22700099>
55. Riemer M, Kearns MA. Description and psychometric evaluation of the youth counseling impact scale. *Psychol Assess*. 2010;22(2):259-268. doi:10.1037/a0018507
56. Kearns MA, Athay MM, Riemer M. Measuring Youths' Perceptions of Counseling Impact: Description, Psychometric Evaluation, and Longitudinal Examination of the Youth Counseling Impact Scale v.2. *Adm Policy Ment Health*. 2012;39(0):104. doi:10.1007/S10488-012-0414-Z
57. Attride-Stirling J. *Development of Methods to Capture Users' Views of Child and Adolescent Mental Health Services in Clinical Governance Reviews*.; 2003. Accessed February 14, 2022. https://www.corc.uk.net/media/1215/chi_projectevaluationreport.pdf
58. Brown A, Ford T, Deighton J, Wolpert M. Satisfaction in child and adolescent mental health services: Translating users' feedback into measurement. *Adm Policy Ment Heal*

- Ment Heal Serv Res.* 2014;41(4):434-446. doi:10.1007/s10488-012-0433-9
59. Barrett-Kruse C. Brief counselling: A user's guide for traditionally trained counsellors. *Int J Adv Couns* 1994 172. 1994;17(2):109-115. doi:10.1007/BF01407967
 60. Young J, Dryden W. Single-session therapy – past and future: an interview. *Br J Guid Couns* . 2019;47(5):645-654. doi:10.1080/03069885.2019.1581129
 61. R Core Team. R. Published online 2020. <http://www.r-project.org/index.html>
 62. Donabedian A. The Quality of Care: How Can It Be Assessed? *JAMA*. 1988;260(12):1743-1748. doi:10.1001/JAMA.1988.03410120089033
 63. Young J, Weir S, Rycroft P. Implementing single session therapy. *Aust New Zeal J Fam Ther.* 2012;33(1):84-97. doi:10.1017/AFT.2012.8
 64. Schleider J, Weisz J. A single-session growth mindset intervention for adolescent anxiety and depression: 9-month outcomes of a randomized trial. *J Child Psychol Psychiatry.* 2018;59(2):160-170. doi:10.1111/JCPP.12811
 65. Campbell A. Single-Session Approaches to Therapy: Time to Review. *Aust New Zeal J Fam Ther.* 2012;33(1):15-26. doi:10.1017/AFT.2012.3
 66. Petrie H, Power C, Cairns P, Seneler C. Using Card Sorts for Understanding Website Information Architectures: Technological, Methodological and Cultural Issues. In: Campos P, Graham N, Jorge J, Nunes N, Palanque P, Winckler M, eds. *Human-Computer Interaction – INTERACT 2011. Lecture Notes in Computer Science, Vol 6949*. Vol 6949 LNCS. Springer; 2011:309-322. doi:10.1007/978-3-642-23768-3_26
 67. Fincher S, Tenenberg J. Making sense of card sorting data. *Expert Syst.* 2005;22(3):89-93. doi:10.1111/j.1468-0394.2005.00299.x
 68. Sweller J. *Cognitive Load Theory*. Vol 55. Academic Press; 2011. doi:10.1016/B978-0-12-387691-1.00002-8
 69. Coates D, Winston T. The Dilemma of Distress Disclosure. In: Derlega VJ, Berg JH, eds. *Self-Disclosure. Perspectives in Social Psychology*. Springer; 1987:229-255. doi:10.1007/978-1-4899-3523-6_11
 70. Hanna L, Hunt S, Bhopal R. Insights from Research on Cross-Cultural Validation of Health-Related Questionnaires: The Role of Bilingual Project Workers and Lay Participants. *Curr Sociol.* 2008;56(1):115-131. doi:10.1177/0011392107084382
 71. Campbell A. Single-session approaches to therapy: Time to review. *Aust New Zeal J Fam Ther.* 2012;33(1):15-26. doi:10.1017/aft.2012.3
 72. Markowitz JC, Milrod BL. The Importance of Responding to Negative Affect in Psychotherapies. *Am J Psychiatry.* 2011;168(2):124. doi:10.1176/APPI.AJP.2010.10040636
 73. Mohr DC. Negative Outcome in Psychotherapy: A Critical Review. *Clin Psychol Sci Pract.* 1995;2(1):1-27. doi:10.1111/J.1468-2850.1995.TB00022.X
 74. Howard KI, Lueger RJ, Maling MS, Martinovich Z. A phase model of psychotherapy outcome: causal mediation of change. *J Consult Clin Psychol.* 1993;61(4):678-685. doi:10.1037//0022-006X.61.4.678
 75. Spence SH, O'Shea G, Donovan CL. Improvements in Interpersonal Functioning Following Interpersonal Psychotherapy (IPT) with Adolescents and their Association with Change in Depression. *Behav Cogn Psychother.* 2016;44(3):257-272. doi:10.1017/S1352465815000442
 76. Perkins R, Scarlett G. The effectiveness of single session therapy in child and adolescent mental health. Part 2: An 18-month follow-up study. *Psychol Psychother Theory, Res Pract.* 2008;81(2):143-156. doi:10.1348/147608308X280995
 77. Morgado FFR, Meireles JFF, Neves CM, Amaral ACS, Ferreira MEC. Scale development: ten main limitations and recommendations to improve future research practices. *Psicol*

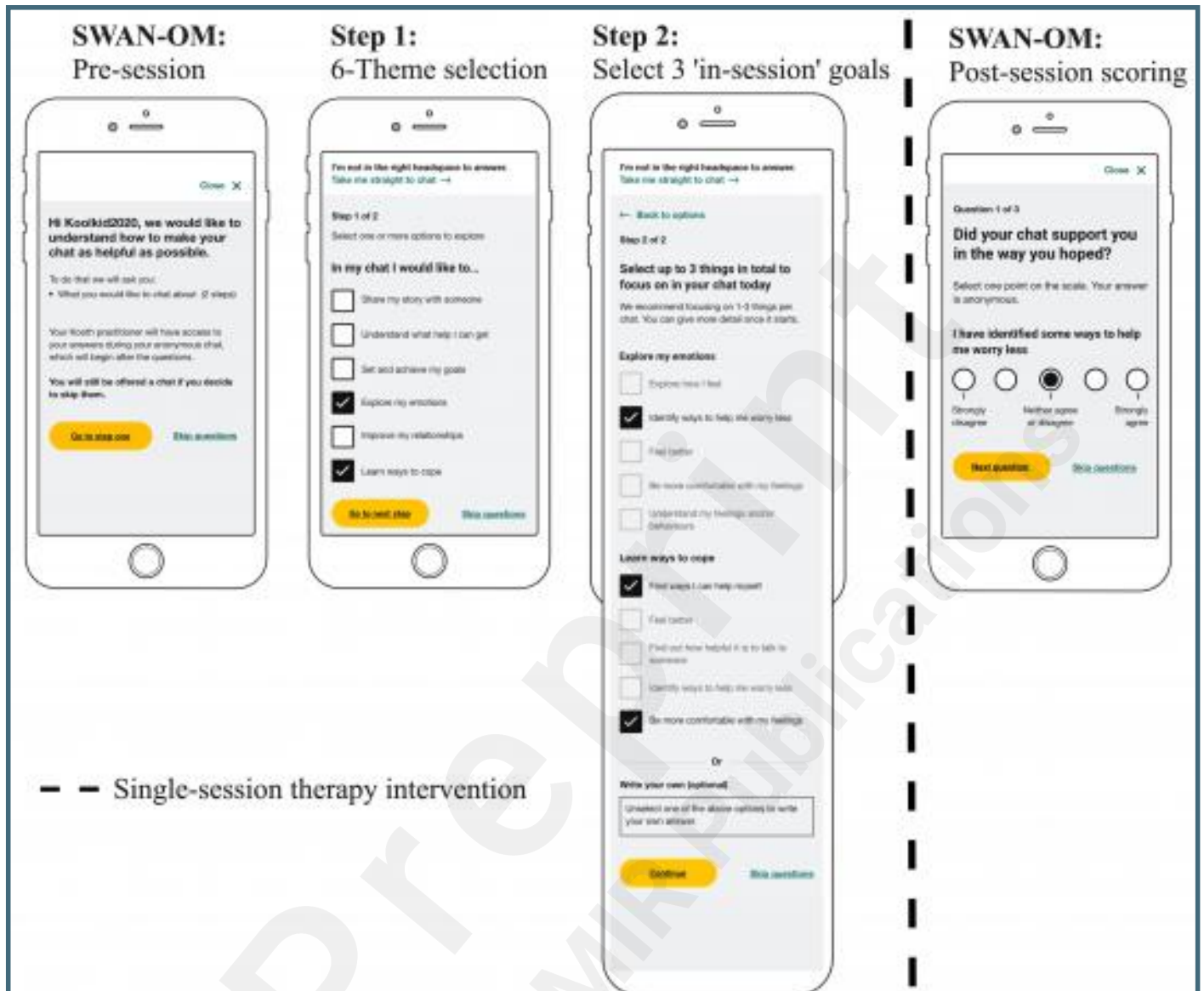
- Reflexão e Crítica* 2017 301. 2017;30(1):1-20. doi:10.1186/S41155-016-0057-1
78. Streiner DL, Kottner J. Recommendations for reporting the results of studies of instrument and scale development and testing. *J Adv Nurs*. 2014;70(9):1970-1979. doi:10.1111/JAN.12402
 79. Hanley T, Sefi A, Grauberg J, Prescott J, Etchebarne A. A theory of change for web-based therapy and support services for children and young people: Collaborative qualitative exploration. *JMIR Pediatr Parent*. 2021;4(1):e23193. doi:10.2196/23193
 80. O'Reilly-Shah VN. Factors influencing healthcare provider respondent fatigue answering a globally administered in-app survey. *PeerJ*. 2017;5(9). doi:10.7717/PEERJ.3785



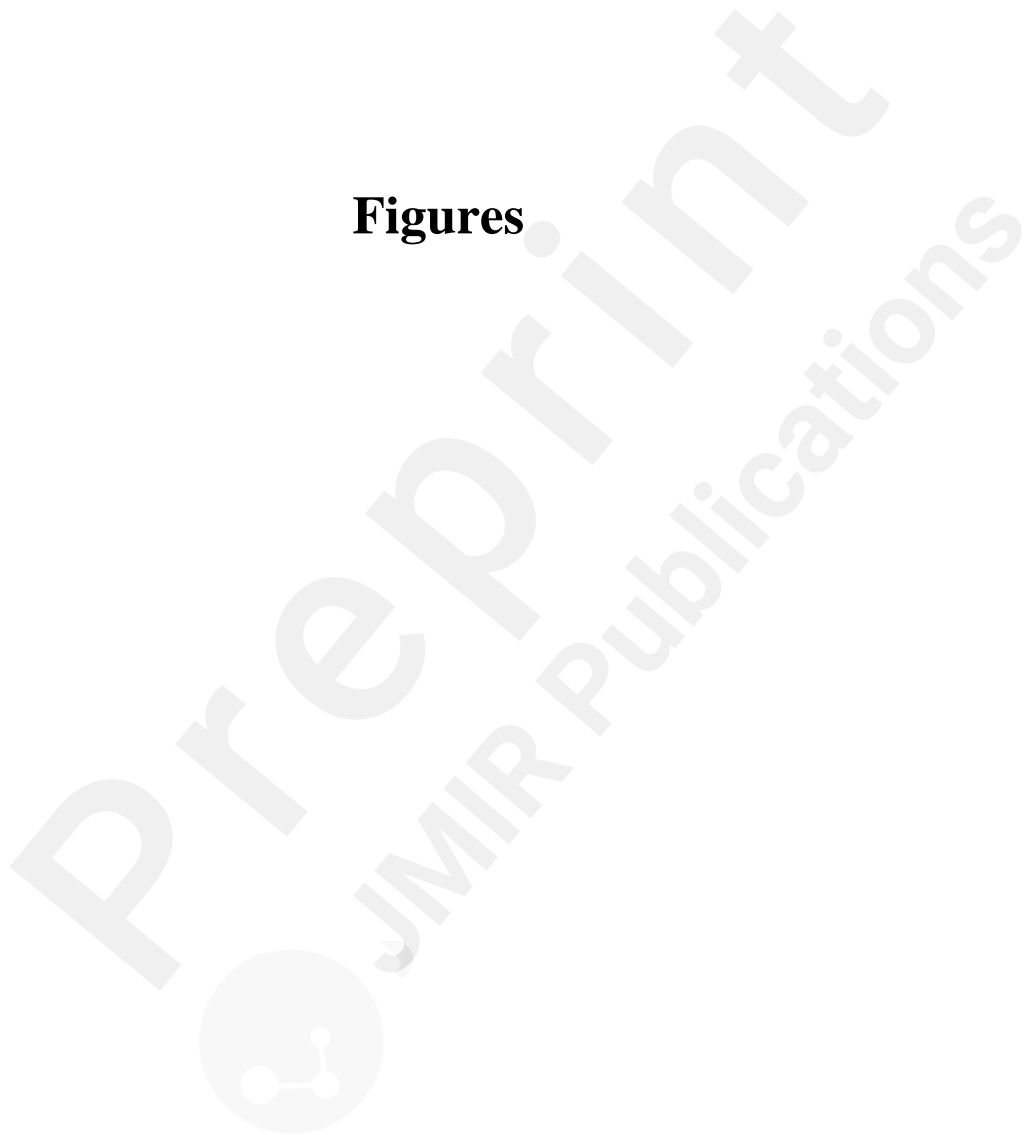
Supplementary Files



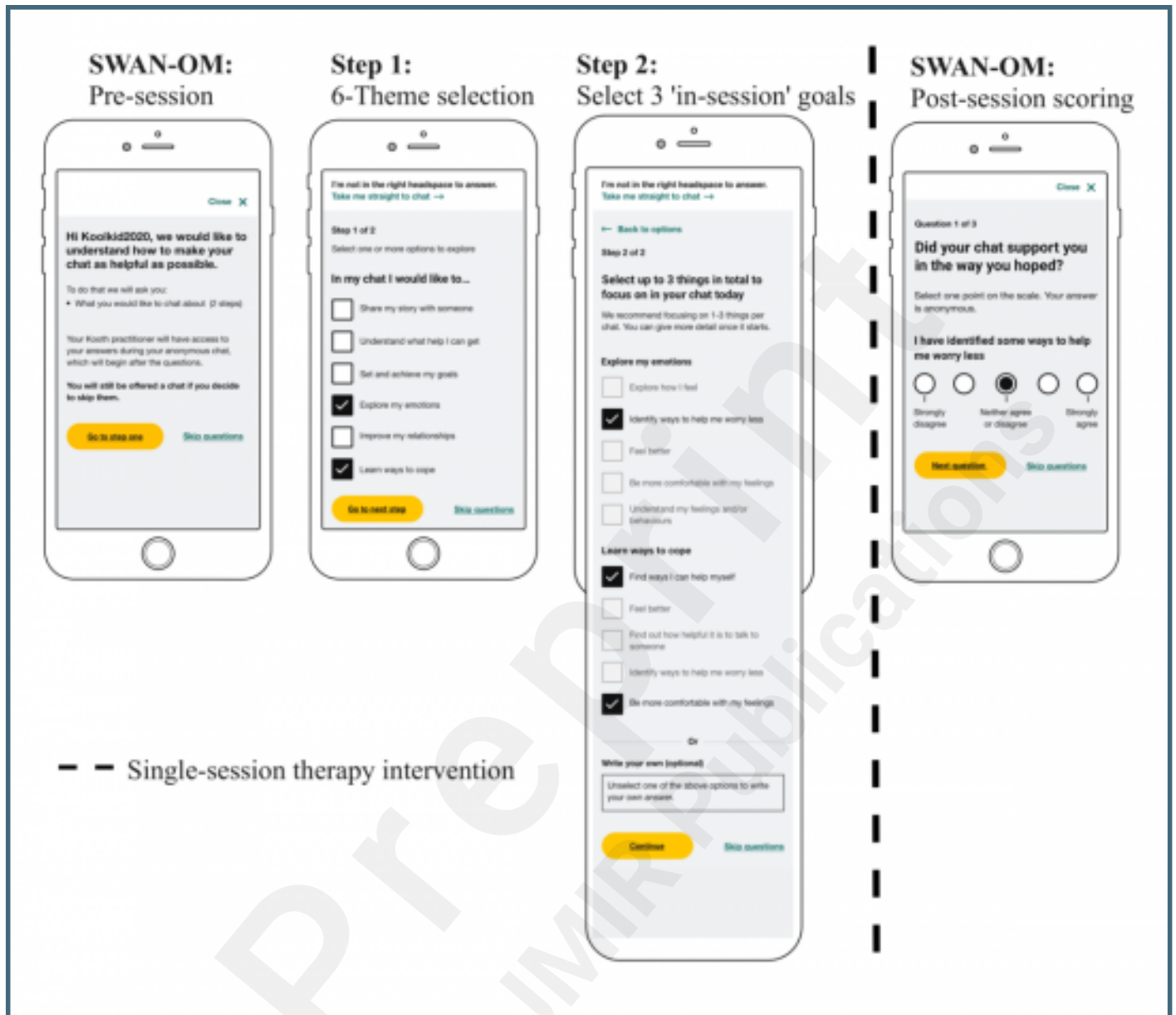
Untitled.



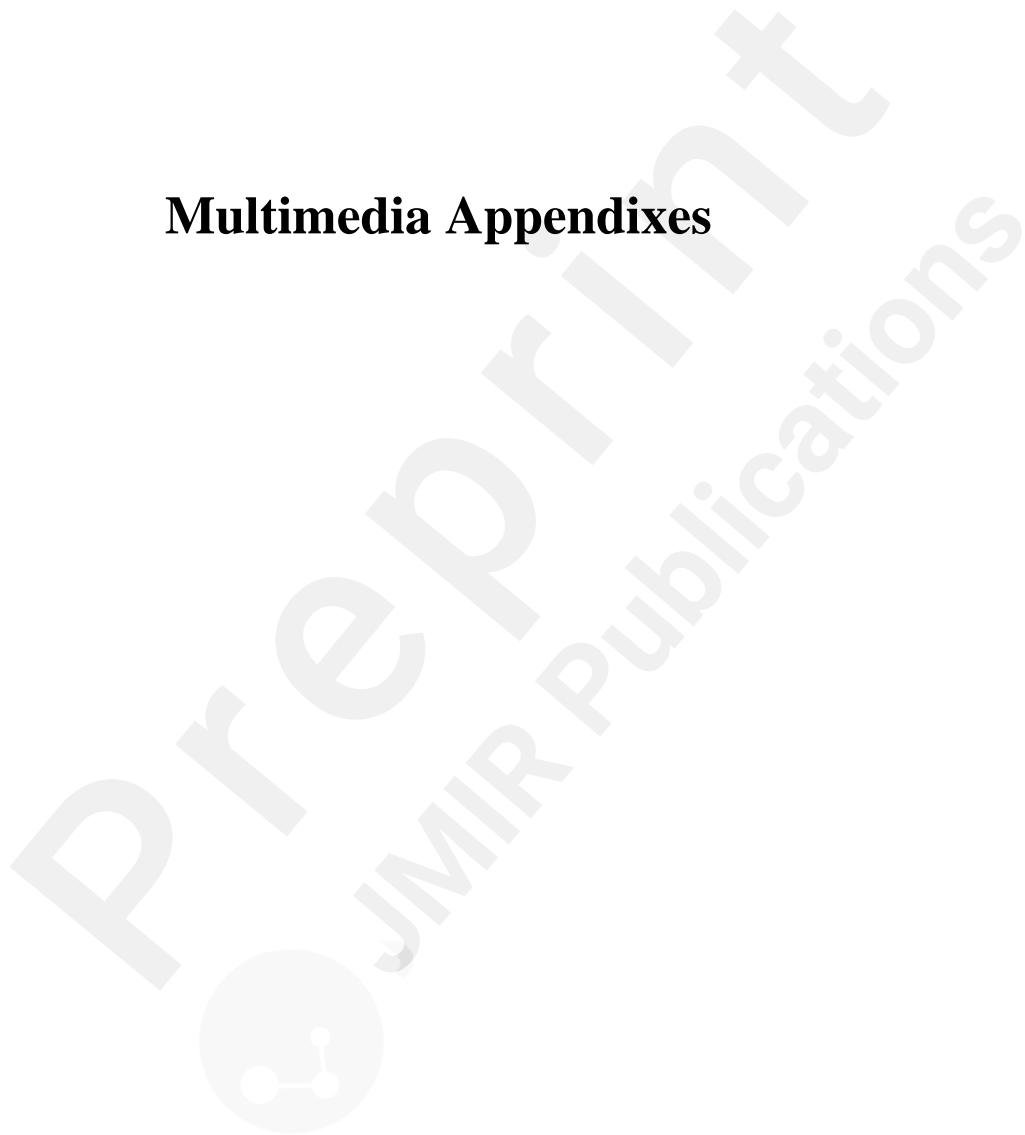
Figures



Smartphone wireframes of SWAN-OM at Kooth.com.



Multimedia Appendixes



Session Wants And Needs Outcome Measure [SWAN-OM].

URL: <http://asset.jmir.pub/assets/5937e527222813160b416ca10badee80.docx>

Supplementary Tables.

URL: <http://asset.jmir.pub/assets/08ae201cb7a1f0ab9fb66889aa1cad60.docx>

