Author's response to reviews

Title: A rapid screening tool for psychological distress in children 3-6 years old: results of a validation study

Authors:

Caroline Marquer (caroline.marquer@epicentre.msf.org)
Caroline Barry (caroline.barry@inserm.fr)
Yoram Mouchenik (yoram.mouchenik@free.fr)
Sarah Hustache (sarah.hustache@epicentre.msf.org)
Douma M Djibo (djibo.doumamaiga@gmail.com)
Mahamane L Manzo (manzolic@yahoo.fr)
Bruno Falissard (Falissard_b@wanadoo.fr)
Anne Révah-Levy (anne.revah-levy@clubinternet.fr)
Rebecca F Grais (rebecca.grais@epicentre.msf.org)
Marie-Rose Moro (marie-rose.moro@cch.aphp.fr)

Version: 2 Date: 27 February 2012

Author's response to reviews: see over
February 27, 2012

Martina Ruf, MD
Contributing Editor
BMC Psychiatry

Reference: Submission of revised manuscript 1315152106606504: A rapid screening tool for psychological distress in children 3-6 years old: results of a validation study

Dear Dr. Ruf,

Thank you for reviewing our manuscript, “A rapid screening tool for psychological distress in children 3-6 years old: results of a validation study”

We are grateful to the editors and reviewers for their insightful comments, all of which have been carefully considered and incorporated in the revised version that we are enclosing.

We have changed the name of the tool used to facilitate comprehension and future use of the tool by rendering its name clearer (Rapid Screening Tool 22, RST-22).

Please also find a point-by-point response to each of the editors’ and reviewers’ comments.

We thank you for the consideration of our work and look forward to hearing back from you.

Sincerely,

Caroline Marquer
Epicentre
Epidemiology and Population Health
caroline.marquer@epicentre.msf.org
Editors’ Comments

We thank the editors for their comments, all of which have been considered and incorporated into the revised version enclosed.

**Comment 1:** Please describe the single steps of your study more detailed

**Response:** Following your comment, the methods section had been rewritten, including more details and clarifying each step of the study.

**Comment 2:** Please rework the introduction setting and give more explanations to your statements

**Response:** The introduction of the manuscript has been revised to provide additional information.

**Comment 3:** Please explain more in detail the different steps of your study (test-retest, external validity and inter rater)

**Response:** The different steps of the study were revised and the methods section rewritten accordingly.

**Comment 4:** Please give more information on the practical implications of your study

**Response:** Additional information on the practical implications of this study have been added to both the introduction and the conclusion.
Reviewer 1
Elisabeth Schauer

Comment 1: The authors are to be commended for having undertaken the development of a short screening tool for mental health needs of children in humanitarian settings. The need to develop screening tools for lay counselors is certainly important, given the many children who are exposed to family-, organised violence and complex emergencies around the globe.

Response: We thank the reviewer for this positive feedback.

Comment 2: Research in humanitarian settings shows that despite the often claimed ‘over focus’ on trauma, little has been done in terms of screening and treating trauma spectrum disorders in resource poor settings. So it is commendable and of key importance that trauma and its psycho-pathology is part of any screening tool that wants to adequately represent the mental health rehabilitation need in settings of humanitarian emergency in a given child population.

Response: We thank the reviewer for highlighting this important point.

Comment 3: Much of the paper however lacks sufficient explanation and description. As a reader it is hard to follow the logic of the methods part, through to results and discussion. For example: What was the reasoning to draw the specific sample? How exactly was test-retest carried out? How was validation carried out? Where is the result of the qualitative research reflected in the discussion? The paper should be reworked, better structured and much extended in explanation.

Response: This is an important point, thank you for highlighting the need for clarification. The specific steps of the study were clarified in the revised methods section. The structure of the article was revised to provide a more logical flow. Sample selection, test-retest, and the validation process have been rewritten and additional detail included. Concerning the qualitative component, this section has been revised, clarifying this step as a part in the validation process. The qualitative component of the cross-cultural validation process was implemented at the beginning of the study to translate and adapt the tool. In addition, the clinician spent time during this beginning period to gather information concerning the detection of psychological difficulties within the community.

Introduction

Comment 4: In general it would be good, if the paper would have a more elaborated background section on the situation of children in humanitarian settings and especially a more guiding explanation of the local setting from which data was drawn.
Response: Thank you for providing advice on the background section. These comments were included and the introduction revised.

Comment 5: The description of the tool, the RST-22, should be moved to the methods section.

Response: The description of the tool has been moved to the methods section.

Comment 6: The authors claim that their paper is the first assessment tool designed for humanitarian setting and to be applied by local lay practitioners. This is not up-to-date current knowledge on peer-reviewed journal publications. Earlier such instruments have successfully been validated and presented in peer-reviewed papers, e.g., by Elbert et al. (2009) for Tamil children, Ertl et al., in a sample of former child soldiers or by Odenwald et al. for a Somali context, as well as others. It would be useful if authors include existing research and their findings.

Response: This is an important point and this has been revised. We meant that a tool including general psychological difficulties did not yet exist and have reworded this section so that it does not state that no tool has been validated previously in such context for trauma screening. The section was revised to point this important comment and propose a clearer understanding of the previous research concerning instruments for young children, both in the text and by including additional references.

Comment 7: There are a number of statements in the introduction, which remain unclear to the reader. For example: “Psychological distress in young children is particularly difficult to evaluate by non-specialists, because of its specific psychopathology”. What exactly is ‘specific psychopathology’? And why if lay persons were appropriately trained, could they not conclude with a diagnosis? It would be useful if authors could explain such statements in more detail.

Response: We thank the reviewer for this comment. The introduction has been revised to clarify such statements. Specific psychopathology means, “Further, even when mental health professionals are present, they are rarely specialists in young children. Psychological distress in young children is particularly difficult to evaluate by non-specialists requiring knowledge of normal child development as many behaviors are normal at certain ages but not at others. Young children, between 3 and 6 years old, are in a vulnerable psychological period, which can have consequences on the quality of their emotional, cognitive, and physical capacities (5,6). Although the psychological response children depends on their individual, family, and social environments among many other factors, recognizing the diversity of potential psychological responses is essential to provide appropriate intervention (7–9)”

The question of the lay persons is important. The tool was created to screen for potential psychological difficulties who are then referred for diagnoses and
treatment if needed. In humanitarian contexts, there are few, if any specialists in the mental health of young children able to diagnose appropriately. As a result, this tool was developed for screening, to be able to orient children who need a comprehensive assessment, diagnosis and care.

Methods

Comment 8: It is stated by the authors that the present aim was to develop a tool for a complex humanitarian context, yet the tool was tested on a non-Conflict-exposed part of the Niger population. It thus might not accurately serve now for populations of children who have been exposed to conflict or humanitarian hardship in the past or present. The aim of the paper should therefore be rephrased, since the tool was not validated in a population who suffers from a humanitarian crisis.

Response: Thank you for this comment. This has been reworded in the manuscript and clarified further. The entire cross-cultural validation process includes 3 countries. Within these 3, we needed to include both non-violent conflict and violent conflict area to be able to test the post-trauma component of the tool. To be sure that the post-traumatic component could be screened, the validation process necessarily had to include a “non conflict area”. In this case, “non conflict area” means political stability of the country as well as no current war, conflict. The two others countries are Colombia and Kenya, two countries where children had been recently confronted with violence.

Comment 9: Several studies have shown that especially mothers have very poor judgment of the mental health status of their own children. In some studies the data of mother have been withdrawn due to misrepresentation. Best results tend to be obtained through direct child report and adjunct third person report of non-family members, e.g. classroom teachers, or family members but not the primary care taker. Can authors please clarify why they have chosen mothers’ as key informants?

Response: In this context, young children, less than 6 years old, are most often raised by their mothers, who are the best informants for information on daily activities (for example, sleeping and eating habits). In extremely poor rural settings such as this one, we could have included other adult female relatives as informants, but unless they are the direct caregiver, they would not have sufficient knowledge of the daily activities of the child. Most children do not attend school before 5 or 6 years of age. The studies that included the judgment of the mental health status by informants other than the mother, concern school-age children. Here, we esteemed that mothers were the most pertinent key informants. However, as suggested by the reviewer, other key informants could be used if the context allowed. This paragraph had been rewritten to add more details and explanation.

Comment 10: Can authors please explain how exactly the ‘random sample’ was drawn?
The study population included a random sample of 580 caregivers recruited between November 2009 and July 2010. A comprehensive information and awareness campaign was organized before study implementation to inform the population about the aims and objectives of the study. Caregivers were selected using cluster-based sampling, with population proportional to size weighting, a standard methodology in populations where population data is either limited or inaccurate (56). Within the region, 68 villages or city districts were randomly chosen (22 districts in Maradi city, and 23 villages each in the districts of Madarounfa and Guidan Roumdji) (57). The random sample in the villages followed the traditional method by spinning a pen, as random starting direction from a central location in the cluster. Households lying on this transect from the center to the border of the cluster were counted and then one of them was chosen at random. Proximity selection was then used to select subsequent households as the "next nearest" until the desired sample size was reached (56)(58).

In each household, caregivers needed to have at least one child aged 3 to 6 years old, be Hausa speakers, and resident in the study area (Maradi city, Madarounfa, Guidan Roumdji). As mothers were the most knowledgeable of the daily activities and behaviors of children, they were asked to respond to the questionnaire. Children identified by the mother/caregivers as presenting a mental disorder, mental retardation, development disorder, and/or psychosis were excluded. Interviewers also excluded children who met visible, recognizable criteria of mental retardation or grave development disorders. These children were referred immediately for free care. All mothers and caregivers were read an information letter describing the study and asked for written informed consent before enrollment.

Comment 11: Can authors please explain why they did not ask children and their caretakers for distinct trauma experiences, including questions re family violence, conflict experiences, human suffering? Latest research shows that children exposed to humanitarian/complex/conflict settings, suffer an extra amount of family (parents/step parents, extended family e.g. grand parents, foster families) and community violence (teachers, neighbors), which adds to the development of psychopathology caused by experiences of war/organised violence/forced migration.

Response: Thank you for this important and interesting comment. The tool was designed to detect general psychological difficulties including the post-traumatic experience. The caregiver answers questions related to the daily behavior of the child. As the reviewer points out, trauma that the child may be exposed cannot be simply summarized by an event, but also includes other sources of trauma such as those due to the family, and in older children, to the school environment. As a screening tool, difficulties are detected, but at the diagnostic stage, a professional
would go into greater detail on the child's environment and potential sources of trauma if applicable.

Comment 13: Can authors please further clarify how many clinical psychologists carried out clinical diagnosis?

**Response:** Two clinical psychologists carried the interviews. The sentence had been changed to make it clearer.

Comment 14: Was this a direct and confidential child interview?

**Response:** Yes, it was a direct and confidential child interview. In most cases, the mother was the caregiver interviewed. Unless the mothers wanted someone else to participate, the individual interview included the mother/the child/the translator and the psychologist.

Comment 15: Was an interpreter present? How were interpreters trained?

**Response:** A French/Hausa interpreter (translator) was present. She was trained in the same manner as the interviewers for the study and then trained by the psychologist on the specific task of translating "word by word" without interpretation. The sentence had been revised to provide additional clarification.

Comment 16: Local interviewers were trained for 2 days only? Who trained them and what was the content of the training?

**Response:** Local interviewers as well as the translator were trained during 3 days. Two days included presentation of the tool, role plays, background information child development and on the sampling methodology employed for household selection. One day in the field was also used as pilot phase, to observe the local interviewers in their interaction with mothers and provide additional guidance if needed. The section has been revised as follows:

“All interviewers were selected by meeting the main criteria, experience in questionnaire administration, fluent in Hausa and French and without a background in mental health. They were trained over a three-day period on how to administer the questionnaire. After a general presentation of the tool, items were presented and discussed one by one. Role-plays were used to simulate interviews and training provided on the information and informed consent procedure. Interviewers were also trained on study exclusion criteria. The exclusion criteria for interviewers were based on visible, recognizable difficulties and interviewers were trained on the basics of mental retardation and the severe classifications of the ICD-10. A one-day pilot phase followed the theoretical training to assure standardization of
administration and reinforce the theoretical training. The translator participated in
the interviewer’s training and also performed an additional day of training on the
translation process itself (44). Interviewers were supervised during the duration of
data collection to respond to any difficulties or questions.”

Comment 17: It remains unclear why more severely affected children were withdrawn
from the study sample and who diagnosed a ‘psychosis’ in a child of less than 6 years?

Response: We thank the reviewer for this comment and have added additional
clarification. Children presenting easily recognizable, visible signs of severe
impairment, such as repetitive movement, drooling, incoherent vocalization
were de facto referred for further evaluation. The score on the scale would be high
for these children and the children were withdrawn from the study sample and
directly referred to care. As pointed out by the reviewer, the use of the term
psychosis to describe these children has been rewritten to clarify that no diagnosis
occurred.

Comment 18: Can authors please explain who carried out the validation interviews and
with which instruments? It is mentioned that the CGI was applied. We need a good
description of the CGI-S, including its psychometric properties.

Response: The clinical interviews were carried out by clinical psychologists trained
in child development and cross-cultural psychiatry. One of the essential components
of the validation process is comparison to clinical interviews, as the domain of
validation is more clinometric than psychometric. The psychometrics properties
were analyzed and described. A table, additional text and references have been
added to clarify concerning this point.

Comment 19: For validation, the last 255 interviews were used, why that? Why did authors
not draw a randomized sub-sample?

Response: This explanation has been rewritten to provide additional clarification
on the validation methodology. During the validation process, the selection of
children was conducted as a two-steps procedure. After a first-analysis, performed
to provide preliminary results on psychometric properties. As the tool had never
undergone a cross-cultural validation, this preliminary step was included to ensure
that results were promising enough to continue. The 255 interviews concern those
children who were administered the questionnaire after the preliminary analysis.
This section had been rewritten to provide additional clarification which is also
included in the study flow-chart [Figure 1].
Results:

Comment 20: There are far too little results presented. The reader would like to see more calculations and also tables, to better understand the results.

Response: We thank the reviewer for this comment and have provided additional results. We now include Table 2 which presents additional details on the characteristics of the children in the study sample. Table 2 presents the means and standard deviations of the RST-22 in different sub-groups (across gender, ages, CGI severity, etc.). Figure 2 presents the distribution of the RST-22 score in the two samples. Figure 3 presents the frequency of positive responses (sometimes, often) between children needing care and others: this is a way to see graphically which item is discriminant and which is not. Finally, Figure 4 provides the ROC curve of the RST-22 compared to the question: “does the child need psychological/psychiatric care?”.

Comment 21: Were the interview results of mothers’ compared to clinicians’ results? or were the interview results of mothers’ compared to the clinical results of child interviews? Especially, test-retest, inter-rater reliability and external validation should be presented in clear structure and with tables if possible.

Response: Interview results were compared to the clinicians’ results. The score of the tool was compared to clinical evaluation answer to the question: “does the child need psychological care: yes /no? “as well as to the CGI. A score of more than 17 was expected link to correspond to a “yes” response of the clinician. Additional information is now provided on test-retest, interrater reliability and external validation.

Comment 22: Also measures of objective validation are missing. A small child with a mental health disorder, presenting with symptoms of distress, should also present with symptoms of cognitive delay/dysfunction, with problems in motor control and physical challenges, such as stunting and weight problems. These features should be considered in future validation studies.

Response: This is an important point. Hopefully after this first proposed scale, additional research will include more indicators. For this study the main objective was to be able to propose a tool for children in the general population, those who do not have any major developmental issue since birth, as the care for this children can be major, although of course essential. These are critical concerns and the RST-22 is simply the first step in hopefully future developments of this area.

Discussion
Comment 23: The paper lacks a clear description of its practical implications. Authors should explain why rapid non-specific screening is important in settings of complex emergencies, since a screening tool, like the one that is presented, which is general and does not help elucidate on the specific origin of children’s distress might be redundant. Massive evidence has been provided in the past years on the mental health impact of children in circumstances of e.g. forced migration, war, and organized violence around the globe. We know what children in humanitarian settings who have been exposed to trauma suffer from; the challenge is to understand the specific issues of the chosen child population at hand.

Response: We have provided additional explanation and justification in the manuscript. In humanitarian contexts, it is relevant to use trauma scales for screening but also important to propose a general scale, such as the scale proposed here, to be able to comprehensively screen all children potentially in need as trauma is only one component of potential suffering.

“The mental health needs of young children in humanitarian contexts often remain unaddressed (1–4). During the acute phase of a humanitarian emergency, and in humanitarian contexts in general, psychological care of children may come far down on the list of priorities. The limited number of both local and international medical professionals combined with the relative lack of mental health professionals in these settings also hinders the implementation of mental health activities (14). Further, even when mental health professionals are present, they are rarely specialists in young children. Psychological distress in young children is particularly difficult to evaluate by non-specialists requiring knowledge of normal child development as many behaviors are normal at certain ages but not at others. Young children, between 3 and 6 years old, are in a vulnerable psychological period, which can have consequences on the quality of their emotional, cognitive, and physical capacities (5,6). Although the psychological response children depends on their individual, family, and social environments among many other factors, recognizing the diversity of potential psychological responses is essential to provide appropriate intervention (7–9). The training and infrastructure needs in children’s mental health remain the ultimate goal, but in their absence, tools that help identify children who require further assessment would help to rationalize scare resources and orient children to care in humanitarian contexts.

The lack of cross-culturally valid instruments, and data about child psychological difficulties, is a public health concern in humanitarian contexts (2,10–13). Although scales exist for general psychological difficulties, none of them concern children aged 3 to 6 years in humanitarian contexts (15–17). Further, before use, screening tools should be cross-culturally validated for specific contexts (10,18–20). Difficulties with mental health assessment include lack of consistent assessment tools for measuring psychological distress, lack of cross-cultural validation research and variation in methods for validity testing and differences in methods of translation (21). Typically, examining mental health cross-culturally involves
transporting Western assessment tools with no examination of their validity (22). As result, children remain unscreened or evaluated using a scale not designed for either the specificities of childhood psychological distress or the context. The lack of a validated, rapid and simple tool for screening, combined with few mental health professionals able to accurately diagnose and provide appropriate treatment, mean that young children may remain without appropriate care.

In addition, interest in humanitarian settings has focused primarily on trauma rather than other disorders or psychological difficulties (1,11,23–25). This focus is based on the assumption that exposure to violence frequently entails post-traumatic symptoms. This premise has been criticized recently, as well as the use of only a post-traumatic scale for screening (10,26). Research addressing the cross-cultural validity of Western diagnostic classification of psychological difficulties in such contexts remains essential to ensure appropriate care is provided (10,27–30). Recent research has shown the importance of tools able to detect and orient children in need, but has focused on children older than 6 years with an emphasis on post-traumatic stress disorder (22,31–34). In addition to PTSD, recent studies have also shown the importance of addressing depression and anxiety disorders (34)."

Comment 24: Since screening for mental distress of children also carries the ethical obligation to build-up psychological treatment structures (with clinical experts as well as lay counselors), it is unclear why the questionnaire was designed to only test for distress and does not allow clinical diagnosis. Clinical interviewing by local lay counselors, resulting in diagnosis and thereby allowing to gain a clear picture of the present problems, has been successfully developed and implemented in a good number of (post-) conflict settings in recent years. It would be helpful if the authors could reference those. Examples by Catani et al, 2008 & 2010 in a samples of war affected Tamil children in Sri Lanka, by Catani 2009 in a sample of Hazara school children in Kabul, Afghanistan; by Elbert 2009 in a sample of war-affected Sri Lankan school children; by Ertl 2009 in a sample of war-affected children and youngsters in Northern Uganda; by Neuner 2006 in a sample of war and tsunami affected children in Sri Lanka, among other.

Response: This is a crucial point and we thank the reviewer for emphasizing this issue. Additional text has been added for further clarification along with the references suggested. The tool was not created for diagnoses, but to screen for psychological difficulties, provide a rapid needs assessment and to orient referrals. The tools mentioned above are similarly important, but address a different issue and are now references. As the reviewer points out, the ethical considerations are substantial. We felt that the potential benefits of screening, prior to diagnosis, would allow for an additional approach to the problem, which may be beneficial in some settings. In other settings, direct clinical diagnosis may be more relevant.
Comment 25: The authors might further assume that based on the intended outcome of the questionnaire “does the child need psychological/psychiatric care” (yes – no) the children might now be referred to accessible mental health treatment facilities, where proper clinical diagnosis will follow. However reality on the ground shows that those structures (e.g. locally based clinical psychologists, clinical counselors and psychiatrists) almost never exist in complex humanitarian settings. Having screened only for ‘psychological difficulties’ now, does not really allow the development of counseling or treatment services that can possibly fully address children’s needs in a given population. Can authors please clarify the usefulness of such a tool and the practical implications it might yield nevertheless?

Response: We thank the reviewer for this comment and the text has been revised to reflect these concerns. The tool, created in a complex humanitarian setting, addresses the need, for when resources are limited, how children can be screened and oriented towards limited care. In some instances, it will be NGOs providing appropriate care in humanitarian contexts, especially when there are no local structures, which is often the case.
Reviewer 2  
Mark Jordans

Comment 1: Overall, the authors present a much needed study into the validation of a screening tool for young children in humanitarian settings. The conducted study is of importance in the effort to increase availability of contextually valid tools to assess psychosocial distress of children. While of importance, there are several key issues that need to be addressed.

Response: We thank the reviewer for this positive feedback.

Major compulsory revisions:

Comment 2: The authors do not present a clear rationale for why one should conduct this type of screening, nor do they convincingly describe what happens after scoring above a validated cut-off score. This is of crucial importance, because it determines the need to go through the effort of development and validation of instruments in the first place. Not in the least place because of ethical concerns. One of the major criticism in the field has been that there is no use in screening if there are no clear interventions available that follow-up on the screening outcomes. The authors do not provide a clear framework for how to deal with this important public health criteria for screening. There are examples of programs for children in humanitarian settings that include systematic screening, yet this is always linked to subsequent service provision. The authors are invited to elaborate on this point. Of importance here is further the lack of clarity what intervention one might do for children with generic psychosocial distress of that young age (3-6 years).

Response: Thank you for this important comment. This section had been rewritten to provide more details and clarification on the justification of a screening tool. Ethical concerns have also been incorporated into the text as suggested in the introduction.

“The mental health needs of young children in humanitarian contexts often remain unaddressed (1–4). During the acute phase of a humanitarian emergency, and in humanitarian contexts in general, psychological care of children may come far down on the list of priorities. The limited number of both local and international medical professionals combined with the relative lack of mental health professionals in these settings also hinders the implementation of mental health activities (5). Further, even when mental health professionals are present, they are rarely specialists in young children. Psychological distress in young children is particularly difficult to evaluate by non-specialists requiring knowledge of normal child development as many behaviors are normal at certain ages but not at others. Young children, between 3 and 6 years old, are in a vulnerable psychological period, which can have consequences on the quality of their emotional, cognitive, and physical capacities (6,7). Although the psychological response children depends on their
individual, family, and social environments among many other factors, recognizing the diversity of potential psychological responses is essential to provide appropriate intervention (8–10). The training and infrastructure needs in children’s mental health remain the ultimate goal, but in their absence, tools that help identify children who require further assessment would help to rationalize scarce resources and orient children to care in humanitarian contexts.

The lack of cross-culturally valid instruments, and data about child psychological difficulties, is a public health concern in humanitarian contexts (2,11–14). Although scales exist for general psychological difficulties, none of them concern children aged 3 to 6 years in humanitarian contexts (15–17). Further, before use, screening tools should be cross-culturally validated for specific contexts (11,18–20). Difficulties with mental health assessment include lack of consistent assessment tools for measuring psychological distress, lack of cross-cultural validation research and variation in methods for validity testing and differences in methods of translation (21). Typically, examining mental health cross-culturally involves transporting Western assessment tools with no examination of their validity (22). As result, children remain unscreened or evaluated using a scale not designed for either the specificities of childhood psychological distress or the context. The lack of a validated, rapid and simple tool for screening, combined with few mental health professionals able to accurately diagnose and provide appropriate treatment, mean that young children may remain without appropriate care.

In addition, interest in humanitarian settings has focused primarily on trauma rather than other disorders or psychological difficulties (1,12,23–25). This focus is based on the assumption that exposure to violence frequently entails post-traumatic symptoms. This premise has been criticized recently, as well as the use of only a post-traumatic scale for screening (11,26). Research addressing the cross-cultural validity of Western diagnostic classification of psychological difficulties in such contexts remains essential to ensure appropriate care is provided (11,27–30). Recent research has shown the importance of tools able to detect and orient children in need, but has focused on children older than 6 years with an emphasis on post-traumatic stress disorder (22,31–34). In addition to PTSD, recent studies have also shown the importance of addressing depression and anxiety disorders (34).

Our aim was to begin to respond to one of the gaps in addressing the mental health needs of young children in humanitarian contexts. Although there are many valid models of screening, evaluation and care, the need for a simple, rapid screening scale administered by non-specialists would fill one of the many gaps in responding to the mental health needs of young children in humanitarian contexts. We report the results of a study to cross-culturally adapt and assess the reliability, validity and psychometric properties of the Rapid Screening Tool 22 (RST-22) following cross-cultural validation. Although the RST-22 had been validated in several populations (35,36), it had never undergone but had never undergone a rigorous cross-cultural validation process. The entire validation process for the RST-22 includes three
validations (one principal and two secondary), with the results of the principle validation presented here. Selection of study sites were based on the political context and the population of children expected to be exposed to conflict to facilitate the evaluation of the post-traumatic component of the screening scale.

Comment 3: A second crucial point of needed revisions, in my opinion, is that currently the authors seem to combine three different objectives into one overarching objective of ‘validation’, i.e. adaptation of the instrument, concurrent validation and item reduction. However, these represent two separate and distinct phases and objectives: instrument development and adaptation (including the item reduction question and cultural adaptations question) and validation (assessment of the psychometric properties of the final instrument). In the present form, the paper combines these processes. Furthermore, if item reduction is aimed for, it is strange that the authors do not use any factor analyses in the process. Similarly, for the item adaptation question, the authors do not present clearly how they have gone about this (beyond the translational issues and FGDs – and even with the conducted FGDs it is not entirely clear how the data was used in the item adaptation process).

Response: We thank the reviewer for pointing out the need to clarify further the methodology used in this study. We have now rewritten the methods section and included a study flow-chart to highlight these different steps. The FGD were used for different objectives in the process of validation. FGD were implemented to adapt the tools to the context first. Then, when the tool was translated, to assure that all items were understood by the community. We felt that it was important to present these results even if it was one step of the cross-cultural validation.

Comment 4: While I think that the emphasize on non-diagnostic or non-disorder-focused instruments is very valid in low-income settings where resources are unavailable to provide disorder specific care, it is still necessary to provide a framework for what is being assessed. The authors refer to generic psychological distress, but fail to define this. Furthermore, from the interpretation of the results it appears that it is a collection of domains including depression, PTSD and anxiety, however it is not clear whether these are subscales or whether it is a list of items that corresponds with symptoms from these diagnostic clusters. Nor is there information on how this was perceived for the original instrument. The authors refer to ‘several areas and domains of psychopathology’ but do not clarify this statement, and at other times refer to it as ‘traumatic exposition’ (e.g. Page 11). Finally, the authors refer to using the RST-22 scores to make inferences of psychological difficulties, followed by a list of specific disorders ranging from depression to PTSD (page 5). In addition, it is striking that the authors have not included a construct validity component to this study, nor make mention of the lack thereof in the limitations section (i.e. what is being measured?). This is important in light of the theoretical issue above, but also for methodological reasons (i.e. can one use composite total scores using this scale?)
Response: We thank the reviewer for these important comments. As part of the answer, we have revised the methodology section of the paper to provide a clearer explanation. To address further what we mean by generic psychological distress, we have added additional explanation on the rationale for the construction of the original instrument. Initially, the tool was created as a list of items, rather than separate sub-scales.

Comment 5: Finally, the use of the gold standard is not clear. The authors describe a combination of clinical interview and administration of the CGI-S, but it is not clear how the indication for need for psychological/psychiatric care was made. Furthermore, given that this is the gold standard it would be good to know how this method has been applied in other settings. Finally, it is unclear to me what type of psychological/psychiatric care the psychologist is screening for (again, especially given the age group 3-6 years old)? And while the presented AUC with sensitivity and specificity are presented, it would be good to also include other metrics of concurrent validity, such as positive and negative predictive values.

Response: Clinical interviews were used as the gold standard. The need for psychological care was made after a classical clinical interview provide by a trained clinical psychologist. After answering to the question, the psychologist was asked to give his impression in the severity of the disorder presented or not. Concerning the results, the explanation is provide comment 8 below. We have rewritten the text to provide additional clarification on this issue.

Minor compulsory revisions:

Comment 6: In the results section of the abstract the authors refer to construct validity; however, this study does not assess construct validity, only concurrent validity. This should be corrected.

Response: Thank you for the comment and the terminology has been changed in the text.

Comment 7: A few times the authors present the conducted study as the first validation of a screening scale with cross-cultural validation for use in a humanitarian context. That is not correct (given also the references that the authors use, i.e. Jordans et al.). The authors should correct this for the fact that this might be the case for the 3-6 years old age group.

Response: We apologize for the confusion and have reworded the text accordingly. We have reworded to state that to our knowledge, this is the first cross-cultural validation of a screening tool for psychological difficulties in children 3 to 6 years old.
Comment 8: The authors report a method for assessing test-retest reliability that seriously allows recall bias. This should at least be included as a limitation.

Response: We agree with the reviewer and this is a frequent limitation when validating clinical scales: when the test-retest interval is too short, recall bias is possible; while when the test-retest interval is too long the clinical state of the patient may be different therefore no longer a test-retest, assessment. As recommended by the reviewer we have added the following sentence in the discussion:

“Third, test-retest and interrater reliability was estimated from interviews performed on the same day. Time between the different interviews is problematic since short interval are prone to recall bias, while long intervals risk being associated with the clinical evolution of the subject evaluated”

Comment 9: It would be good to include a table with overview of main validation results.

Response: We appreciate this comment from the reviewer as this was not an oversight, but our choice. Indeed, the lack of validation results can appear embarrassing and we should have discussed this option. The traditional validation process (item analysis, factor analysis, Cronbach, etc.) of a psychometric tool was developed because these instruments do not benefit from a gold standard so that the validation process used with objective measurements is not straightforward. In our case, there is a gold standard, which is the answer that the clinician gave to the question: “does the child need psychological/psychiatric care?”. The objective of the RST-22 is to screen subjects who need psychiatric/psychological care and in an ideal situation this is a clinician who has to decide after an interview if such care is required. For this reason we chose to validate the RST-22 as compared to the above question and classical statistics like sensitivity and specificity are proposed in the text. Most other screening tools in medicine are based on the same methodology; this is the case for example for scores of gravity in intensive care units like the APACHE score (Statistical validation of a severity of illness measure. Wagner DP, Knaus WA, Draper EA. Am J Public Health. 1983 Aug; 73(8):878-84). This last score has been validated as compared with mortality and not with psychometric tools. In fact, we conducted the traditional psychometric analyses and the scree plot with parallel analysis (simulations) and factor analysis are presented below for reference.
## Table 1: Results of the Factor Analysis: Components and factor loadings

<table>
<thead>
<tr>
<th>RST-22 Items</th>
<th>anxiety</th>
<th>Externalized depressive disorder</th>
<th>Oral language disorder</th>
<th>Autistic behaviors</th>
<th>Isolation</th>
</tr>
</thead>
<tbody>
<tr>
<td>6. The child is frightened, worried, anxious</td>
<td>0.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. The child had a bad dream or a nightmare that comes often</td>
<td>0.61</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. The child wakes up frequently, insomnia</td>
<td>0.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13. The child has outbursts, have uncontrolled movements for no apparent reason</td>
<td>0.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. - The child refuses to eat repeatedly</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. The child has difficulty to be clean (pee, poop, ...)</td>
<td>0.23</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
15. The child is unable to sit still, he moves constantly 0.51
18. The child's behavior is really too aggressive, he is violent (at home and/or outside) 0.48
21. The child plays repetitive games or activities 0.47
20. The child is easily overwhelmed by his emotions anger, sadness, fraternal jealousy etc. 0.44
17. The child is tired, discouraged 0.36
8. The child refuses to separate with one of his parents, siblings etc. 0.26
10. The child does not speak or very little, his language is very different from children of his age 0.88
1. The child stutters 0.62
4. The child is absent, seems somewhere else or in "his world", has difficulties to interact with you 0.81
9. The child eats too much 0.57
16. The child refuses to leave the household 0.58
19. The child isolates himself or often moves away from others 0.41
12. The child has difficulty falling asleep 0.33
22. The child runs away or avoids sounds, images or specific situations 0.32
11. The child refuses to eat certain foods and chooses what to eat at every meal 0.19
14. The child complains of pain or complains about his body without obvious medical reason 0.18

The structure is in favor of a 5-factor solution. The shape of the scree plot corresponds to what has been recently found in a meta-analysis of the structure of various psychiatric scales (Sample size requirements for the internal validation of psychiatric scales. Rouquette A, Falissard B. Int J Methods Psychiatr Res. 2011 Dec; 20(4):235-49)

We would prefer not to present these results because we believe they may be misleading and prefer to present them in a separate paper describing the factor structure of the RST-22 in 3 different settings corresponding to 3 different clinical conditions. We presume that the structure will be different and that this difference could be interesting from a clinical point of view, especially in a transcultural perspective.
We propose to add the following to the discussion section:

“An additional apparent limitation is that we do not present a traditional psychometric validation. The traditional validation process (item analysis, factor analysis, etc.) of a psychometric tool was developed in the absence of a gold standard (69). The objective of the RST-22 is to screen subjects who need further evaluation for psychiatric/psychological care. In an ideal situation, this is a clinician who decides, after an interview, if such a care is required. Our gold standard, the clinician’s answer to the question: “does the child need psychological/psychiatric care?”. For this reason, we validated the RST-22 as compared with the above question and classical statistics. Most other screening tools in medicine are based on the same methodology; this is the case for example for scores of gravity in intensive care units such as the APACHE score (70), validated against mortality rather than with psychometric tools. After secondary evaluations have been completed, the factor structure of the RST-22, corresponding to different clinical conditions should be investigated. The internal consistency may be viewed as a limitation from a psychometric perspective. However, as the RST-22 is a screening tool for psychological difficulties, the scale is not one-dimensional to ensure the detection of psychological difficulties in several area of psychopathology. As the RST-22 includes several domains, this is not unexpected”.

Discretionary revisions:

Comment 10: Why was the tool initially developed in French?

Response: The tool was created by francophone clinicians.

Comment 11: Why was language itself a reason for population selection, the fact that it is a language spoken in different parts of sub Saharan Africa does not entail that there are no cultural variations between these settings. I would opt for omitting this reason for sample selection. In addition, the second reason for choosing this population is not clearly formulated and could deserve some reformulation.

Response: We thank the reviewer for this comment. Language selection was an important component. We felt that the more common languages should be addressed first in the validation process, to provide as many opportunities as possible for use of the tool. We agree that common language certainly does not entail cultural uniformity, but this was a pragmatic choice on our part. The next languages for validation include Swahili and Spanish, but there is a clear need for additional languages and contexts to be explored. This section has been reworded to provide additional clarification on the rationale behind our choice of language.

Comment 12: The sampling strategy makes mention of exclusion criteria being MR, development disorder and psychosis; but how were these identified?
Response: We thank the reviewer for this need for additional clarification, which was also raised by reviewer 1.

“All interviewers were selected by meeting the main criteria, experience in questionnaire administration, fluent in Hausa and French and without a background in mental health. They were trained over a three-day period on how to administer the questionnaire. After a general presentation of the tool, items were presented and discussed one by one. Role-plays were used to simulate interviews and training provided on the information and informed consent procedure. Interviewers were also trained on study exclusion criteria. The exclusion criteria for interviewers were based on visible, recognizable difficulties and interviewers were trained on the basics of mental retardation and the severe classifications of the ICD-10. A one-day pilot phase followed the theoretical training to assure standardization of administration and reinforce the theoretical training. The translator participated in the interviewer’s training and also performed an additional day of training on the translation process itself (44). Interviewers were supervised during the duration of data collection to respond to any difficulties or questions.”

Comment 13: The first of the presented limitation seems more like strength of the study.

Response: We presented this comment as a limitation because translation-back translation, is presented as the stronger method. We agree with the review that this may indeed be strength of our study, but as there is debate as to which is preferable, we present it as a limitation.