



Towards reliable primary data collection and harmonized set of indicators in S-LCA on the stakeholder worker

Arij Mohamad Radwan Omar Chabrawi¹ · Josemberg Moura de Andrade² · Cássia Maria Lie Ugaya³ · Marzia Traverso¹

Received: 12 March 2024 / Accepted: 21 October 2024
© The Author(s) 2024

Abstract

Introduction S-LCA has emerged within sustainability assessment in the last 20 years, aiming at assessing the social impacts of products and services. Several improvements on the methodological and theoretical backgrounds were reported in the literature, since its early beginning. However, there are still some gaps that must be filled: lack of consensus on the indicators, methodological limitations, and hegemony of secondary data, among others. Therefore, this research aimed to benefit from one area related to social sciences, named as psychometrics, to develop 14 questionnaires (scales) to collect primary data regarding all worker's impact subcategories.

Method In general terms, the methodological steps were as follows: (1) literature review of the impact subcategories; (2) definition of the constitutive and operational definitions; (3) run focus groups with workers from different economic sectors and positions to deep understand their reality; (4) creating scales' items; (5) run semantics analysis; (6) run specialists' analyses; (7) pilot and final application of the scales; and (8) run exploratory factor analysis (EFA) and confirmatory factor analysis (CFA).

Results and discussion Five different focus groups were run to better understand each subcategory, and 14 semantics analyses to check the understanding of the scales' items. Five different groups comprised of four different experts on S-LCA assessed the items' content. Kappa's Fleiss indicated from moderate to almost perfect agreement. The scales were applied in two rounds. On the first, EFA was run pointing out to factor retention solution identical or close to the literature review. CFA confirmed the models proposed with adequate adjustment indices. The analyses have shown that all 14 scales indicated

Communicated by Sara Russo Garrido.

Contextualization

This paper aimed to provide a general report on the development of 14 assessment tools known as multi-item scales. A multi-item scale is an assessment tool that measures the information of social or behavioral phenomena at the individual or societal levels in a reliable and valid way (Kline 2013). In psychometrics and in this paper it is referred to as "scale". A scale resembles a survey or a questionnaire format-wise, but differs in the method of development, where scales have validity and precision indices associated to its measurement accuracy. To collect primary data for the subcategories relating to the stakeholder "worker" in the Social Life Cycle Assessment (S-LCA) framework. Furthermore, it sought to (1) present some of the main gaps and challenges from previous as well as modern researches published on S-LCA, concerning data collection and methodological gaps; (2) briefly introduce the reader to the field of psychometrics as the scientific background to create the scales; (3) briefly present the methodology and results regarding the development of the scales, as specific papers for each subcategory shall follow this publication; (4) and lastly, provide recommendations to the scales application and how to analyze the data, for future social impact assessment on workers.

Extended author information available on the last page of the article

to have evidence of content validity and validity based on the internal structure. The reliability coefficients relating to the indicators appointed to a high precision on the measurement.

Conclusions This study contributes to overcoming some of the current theoretical and methodological gaps in S-LCA. Practitioners can benefit from 14 scales to collect primary data regarding worker in a reliable, scientific, and confidential way. Data can be used in both types of impact assessment methods. The scales may also support other research initiatives that aim at studying and improving working conditions for all workers, from any economic sector.

Keywords Social Life Cycle Assessment · Primary data · Worker · Data collection · Psychometrics · Scales

1 Introduction

The necessity to complement environmental Life Cycle Assessment (LCA) with social impacts dates back to over 20 years ago, leading, some years later, to the emergence of Social Life Cycle Assessment (S-LCA) (Norris 2006; Dreyer et al. 2006; Hunkeler 2006; Jørgensen et al. 2008; UNEP/SETAC 2009; Petti et al. 2018). In its early stages, S-LCA was included in the scope of LCA in recognition of the need to also address social and socio-economic impacts, beyond the scope of their environmental counterparts (McCabe and Halog 2018). The first inception of social aspects arose at the Society of Environmental Toxicology and Chemistry (SETAC) with the social welfare category in 1999 (Arcese et al. 2018).

Since then, scientific interest in S-LCA has grown with case studies and the development of approaches (Norris 2006; Dreyer et al. 2006; Hunkeler 2006; Jørgensen et al. 2008) to assess the social impacts of products, to promote sustainability (Petti et al. 2018). Soon, it became clear that including social aspects into the engineering framework of LCA would prove to be challenging (Sakellariou 2018). Sustainability researchers understood that it was not enough, and a new technique arose through different approaches and applications to address social matters (Huarachi et al. 2020).

S-LCA is defined as a methodology to assess the potential and actual social impacts of products and services, which can be either positive or negative, along their life cycle encompassing extraction and processing of raw materials, manufacturing, distribution, use, re-use, maintenance, recycling, and final disposal (UNEP/SETAC 2009). It provides information on social and socioeconomic aspects of decision-making, in the prospect of improving the social performance of an organization and ultimately the well-being of stakeholders (UNEP 2020). Therefore, the main objective of S-LCA is to promote the improvement of social conditions and the overall socioeconomic performance of a product throughout its life cycle, for all related stakeholders (UNEP/SETAC 2009).

S-LCA has been gaining increased interest in recent decades, but only after the Guidelines for Social Life Cycle

Assessment for Products (UNEP/SETAC 2009) were published has it gained more relevance and momentum from the sustainability community. With the publication of the Methodological Sheets for Subcategories in Social Life Cycle Assessment, guidance on data collection methods and sources was provided, followed by further application into case studies in subsequent years (Macombe et al. 2018). Therefore, the approach has evolved over time, and several methodological improvements for data collection have been made. The Guidelines and the Methodological Sheets were updated in 2020 and 2021 (UNEP 2020; 2021), respectively, reflecting consistent development by the scientific community.

Nevertheless, S-LCA still faces many constraints. Theoretical and methodological gaps have been pointed out by several authors so far—some of them are introduced in this work. Huarachi et al. (2020) state that S-LCA is in its best era, but there is still a way ahead to achieve full scientific maturity. However, Macombe et al. (2018) present S-LCA as the most effective technique to assess the (real and potential) social impacts of products throughout their life cycle. Different uses of S-LCA make the development of a consistent framework challenging (Arcese et al. 2018), claiming for future development directions towards a harmonized and robust approach to collect and assess social data.

To sum it up, this paper outlines some of the main gaps in the S-LCA framework pointed out by literature, at the inventory level: (1) lack of consensus on the social indicators (Wu et al. 2014; Kühnen and Hahn 2017; Petti et al. 2018; Huertas-Valdivia et al. 2020; Tokede & Traverso 2020); (2) lack of standardized instruments to collect data (Ugaya et al. 2015; Grubert 2021; Moltesen et al. 2018); and (3) primary data collection importance and challenges (Traverso et al. 2012; Macombe et al. 2013; Grubert 2021; Arcese et al. 2018). These gaps point to one of the more impactful shortcomings in the S-LCA methodology: primary data collection and analysis. Therefore, the present study aims to bridge the theoretical and methodological gaps in the social life cycle inventory phase by proposing a set of tools to collect

primary data regarding the stakeholder “worker” and its related subcategories, in a harmonized and scientific way.

2 Literature review on S-LCA challenges and gaps

2.1 Quantitative versus qualitative indicators

When assessing social aspects, more often than not, a researcher will come across subjective topics and qualitative data, from which questions arise, such as “How to avoid biases?”, “How to secure the validity of a subjective measure?”, and “How to quantify a qualitative phenomenon, without disregarding important aspects?”. Social topics, in contrast to the environmental and economic aspects, are intrinsically subjective and context dependent, varying in terms of socioeconomic development, culture, societal rules and norms, and different legal frameworks, for instance. Therefore, finding consensus in thresholds and indicators consists in a real challenge.

In this matter, Petti et al. (2018) affirm that one of the main methodological constraints that S-LCA faces comes from borrowing the impact assessment framework from LCA, as there is a huge difference when dealing with semi-quantitative and qualitative indicators in S-LCA. Kühnen and Hahn (2017) ran a state of the art on S-LCA finding out that approximately 63% comprised case studies, of which 50% employed an essentially quantitative approach versus 13% qualitative. In this sense, the authors question whether qualitative indicators are being neglected due to the natural subjectivity involved in social assessments, in favor of quantitative data, due to the ease of collection and analyzation. A recommendation proposed by Kühnen and Hahn (2017) is to run interviews with the stakeholders assessed, based on their worldly perceptions and subjective experiences (Kühnen and Hahn 2017).

Wu et al. (2014) portray the different types of indicators among quantitative, qualitative, and semi-quantitative, expressing that the indicators collected through interviews/surveys are often of a qualitative nature. In order to aggregate them into impact categories, the qualitative data should be processed into numeric data, representing a scoring schema that, more often than not, relies on the subjectivity and potential biases of the researcher (values, reasoning, cultural background, ideologies, etc.). Thus, when assessing subjective contents and qualitative data in S-LCA, the methodological process should count on impartial and scientifically robust procedures for data collection and analysis, which can be reproducible and repeatable. In the same direction, Grubert (2021) discusses research processes for human subjects, stressing the importance of utilizing methodological resources, such as qualitative interviews and quantitative

surveys, and including basic information on how subjects and/or informants are chosen or recruited, and how they differ from the population at large.

2.2 Primary and secondary data, system boundaries, and stakeholders

In terms of the data profiling, the majority of case studies applied methods for secondary data collection over primary data (Molnar and Waldekker, 2014; Petti et al. 2018; Huertas-Valdivia et al. 2020). This may be due to being less expensive, easier access to secondary data (UNEP/SETAC 2009), scarcity of reliable instruments developed according to scientific procedures from social sciences (Grubert 2021), challenges for measuring subjective data, collecting representative samples of data for certain populations, and potential sensitivity of some of the social indicators. Several authors concluded that “worker” is the stakeholder mostly assessed in case studies (Macombe et al. 2013; Kühnen and Hahn 2017; Petti et al. 2018; and Huertas-Valdivia et al. 2020).

In terms of subcategory, health and safety was the one most assessed (Kühnen and Hahn 2017), possibly due to the fact of being more easily assessed, both with secondary and primary data, and having quantitative indicators with higher consensus as per national statistics in the theme. When it comes to indicators, Wang et al. (2016) identified “preventing forced labor practices,” “preventing child labor,” and “providing minimum and fair wages for workers” as the three most important indicators for assessing social impact. However, the finding can be discussable as the indicators’ importance varies dramatically from the impact category being assessed, as much as the socioeconomic and cultural context in which workers are involved. For system boundaries, Petti et al. (2018) describe that the scope of most studies is from cradle-to-gate (41%), followed by cradle-to-grave (32%).

2.3 Lack of consensus in indicators

Undoubtedly, the lack of consensus on social indicators is one of the greatest challenges faced by S-LCA and is confirmed by several authors throughout the last decade (Traverso et al. 2012; Wu et al. 2014; Kühnen and Hahn 2017; Grubert 2021; Petti et al. 2018; and Huertas-Valdivia et al. 2020). Kühnen and Hahn (2017) also emphasize the lack of experience in empirical data collection and handling for the sustainable and social performance measurement throughout the production chain’s life cycle. Huertas-Valdivia et al. (2020) indicate that a great effort has already been made to homogenize indicators, although consensus must be reached regarding the unresolved aspects of the complexity of social, cultural, and regulation topics when defining an indicator.

Furthermore, Wu et al. (2014) advocate for the use of indirect indicators as a means to better apprehend the reality at stake, as direct indicators often fail to explain the complexities associated with social issues. A practical example listed is the light labor typology of child labor, foreseen by the International Labour Office (ILO). For this indicator, specific criteria must be met, such as if the work is appropriate for the age and maturity level of the child in question, as well as if the child is provided with the opportunity to gain skills and learn a profession. Therefore, the authors drive attention to the complexity of how to capture and convey what social indicators should entail, without being oblivious in a S-LCA.

Another major constraint derived by the lack of consensus in indicators is how indicators are actively treated differently depending on the context of the case study, assuming beneficial or adverse categorizations for both individuals or society as a whole (Arvidsson et al. 2015; Costa et al. 2019; Gompf et al. 2020; Huertas-Valdivia et al. 2020). Arvidsson et al. (2015) ran an extensive review of indicators regarding child labor, working hours, and property rights in and outside S-LCA literature. They concluded that, in the context of S-LCA, these topics are at risk of appearing over-simplistic from a social science perspective, and, by extension, they raise concerns about how the indicators are currently applied in S-LCA in general. In the same direction, Huertas-Valdivia et al. (2020) found out that case studies on S-LCA are commonly diverse in terms of the methods employed to assess S-LCA studies resulting in a broad variety of approaches, particularly in the choice and formulation of indicators.

A clear example may be drawn from child labor, where commonly the indicator “number of children” working in a given facility does not portray the over complex context in which this subcategory should be analyzed. Child labor, depending on the circumstances, has different criticalities or may even be positive and endorsed. Moreover, addressing only the number of children working regardless of the socioeconomic reality may hinder fully targeting the issue within its complexities. Thus, “In the same way as LCA draws on the natural sciences, it would seem logical that S-LCA would draw heavily on the social sciences” (Arvidsson et al. 2015).

2.4 Methodological limitations

Given the aforementioned, it is possible to conclude that S-LCA also deals with methodological challenges, lacking a homogeneous code of practice, despite the significant efforts from UNEP Guidelines to pursue systematization and homogenization (Huertas-Valdivia et al. 2020; Tokede and Traverso 2020). Even with the update of the guidelines for S-LCA (UNEP 2020) and the methodological sheets (UNEP 2021), few progress has been made in the development of

scientific methods capable of assessing the individual’s subjective perceptions, for instance, in the working context (Waldekker & Molnar 2014; Grubert 2021; Tokede and Traverso 2020). It is possible to verify an extensive variety of methods and different proposals in S-LCA that still face some methodological challenges. Case studies employ different inputs, procedures, and impact assessment methods, generally based on the researchers’ priorities, particularly when formulating the indicators (Huertas-Valdivia et al. 2020).

Wu et al. (2014) argue that for practitioners, the data types and methods for data collection remain a critical issue when conducting a S-LCA. Ugaya et al. (2015) list some practical limitations of S-LCAs, such as primary data collection methods, the guarantee of secrecy, and reliability of information. In the same vein, Grubert (2021) reflects on the challenge of rigor in S-LCA, distinguishing two major settings: (1) how data is collected and (2) whether the fundamental questions and approaches in S-LCA are scientifically bound to the methods applied. Additionally, she speaks in favor of further grounding S-LCA with the rigorous methodological frameworks coming from social sciences, especially concerning the benefits gained from standardized data collection procedures, particularly for surveys and interviews (Grubert 2021).

Regarding the first topic, it is possible to find a profusion of new questionnaires for data collection and analytical approaches for scoring responses and impact assessments, suggesting that we have yet to achieve significant value both from standardization and from looking at examples existing in other literature (Grubert 2021). The second topic of challenge in rigor relies on the fact that inquiries and the lenses used for interpreting the results are not justifiable, being focused on the S-LCA practitioner’s choice, for instance, the selection of indicators and impact subcategories (Grubert 2021). Several alternatives are coming from social sciences and, especially, psychology to deal with theoretical and methodological constraints in S-LCA (Chabrawi et al. 2023).

Huertas-Valdivia et al. (2020) highlight another order of challenges inherent to the social character of S-LCA, such as cultural issues and the associated methodological impediments, such as limited reliable data availability and lack of technical know-how in social assessment methods. Another common gap in this area is low data availability, often used as an excuse to not conduct or only partially conduct an S-LCA or another type of sustainability assessment (Groiß-Fürtner et al. 2023).

In this sense, Grubert (2021) advocates for benefitting from the methodological background coming from social sciences to fill the gap in primary data collection in S-LCA. Scientific methods, such as surveys and standardized assessment scales, among other applications, would allow researchers to collect their data, considering socioeconomic

and geographic specificities, reliably. These tools are capable in assessing subjective and qualitative data by applying scientifically validated approaches. In doing so, they yield numeric and, thus, quantitative values reflecting the indicator score in an objective way, exempt from potential researchers' bias (Chabrawi et al. 2023).

3 Objective and scope

An absolute consensus in the literature on S-LCA is that many efforts and advancements were made on a wide variety of fronts aiming at filling the theoretical and methodological gaps, although it is still a methodology under development. Therefore, it still requires resolution and harmonization in terms of indicators, scientifically sound methods for data collection and impact assessment, and developing/improving databases (Wu et al. 2014; Petti et al. 2018; Huertas-Valdivia et al. 2020; Tokede & Traverso 2020).

Therefore, this study aimed at building upon psychometrics, an area of study within psychology, that specializes in the measurement or assessment of social phenomena, to create reliable and standardized tools to collect primary data relating to the stakeholder “worker.” This stakeholder category is assessed most often in S-LCA case studies and consists of the most important yet challenging profile for data collection. The research also sought to theoretically identify the indicators falling under every “worker” subcategory using the ILO international labor standards, both for developing and developed countries. Furthermore, it aimed to empirically verify evidence for the theoretically mapped out indicators. ILO conventions, recommendations, and reports consisted of a reliable and robust baseline to conceptualize and identify the thresholds for the indicators for each subcategory.

When applicable, literature deriving from social sciences was applied to set a sound epistemological background for defining and setting up the list of indicators, following the indication of several authors thus far (Arvidsson et al. 2015; Iofrida et al. 2014; Arcese et al. 2018; Grubert 2021; Petti et al. 2018; Huertas-Valdivia et al. 2020; and Tokede & Traverso 2020). It is noteworthy that the standardized scales and harmonized list of both, theoretically and empirically, identified sets of indicators were intrinsically built compliant with the updated version of the Guidelines for Social Life Cycle Assessment of Products and Organizations (UNEP 2020).

The stakeholder “worker” is comprised of 11 subcategories, namely (1) freedom of association and collective bargaining, (2) child labor, (3) fair wages, (4) working hours, (5) forced labor, (6) equal opportunities/discrimination, (7) health and safety, (8) social benefits/social security, (9) sexual harassment, (10) employment relationship, and (11) small producers, including small farmers. The

last subcategory was not covered in this research, since it is a cross-cutting topic that was necessary to cover further aspects concerning the smallholders which are not easy to assess with the usual subcategories under workers.

However, other subcategories of paramount importance for assessing the social context of work were missing in the worker scope. ILO (2020) resorts to the phenomenon of globalization to explain the growing competition and pressure on all workers, resulting in long hours of work and increased workload. Such scenarios contribute to the deterioration of work relationships, stress at work (Robert 2018), risks of cardiovascular diseases (Xu et al. 2019), and depression (Rugulies et al. 2012). In addition, issues of relational orders at work, such as exaggerated demands, and bullying at work, among others, also impact the volume of absenteeism, turnover, poor image of organizations, and accidents at work, among other physical and psychological illnesses related to work. Therefore, it affects the general ability of organizations to perform well and be competitive. Thus, eliminating violence and stress at work should be a priority for companies, managers, and employees (Di Martino & Musri 2001).

In this sense, new subcategories were suggested that focus on the human relationships pertaining to the organizational working environment: (11) working relationship (Dejours 1993; 2011; 2012; Lancman 2011); (12) recognition at work (Dejours 2012; Clot 2010); and (13) bullying at work (Hirigoyen 2002, 2008; Heloani 2004; Freitas 2007; Soboll 2008). It is important to endorse that the proposed new subcategories are not merit of a psychosociological evaluation alone, but are topics related to workers under the S-LCA framework where the organization practices may them both impact positively or negatively.

Moreover, another topic emerged in this study: (14) remote work, covered in the working hours subcategory, according to the definition and parameters established by ILO and WHO (Messenger 2019; WHO 2021). It has recently been addressed by ILO due to its trivialization after the COVID-19 pandemic, becoming an important source of working complexities, care, and attention (Messenger 2019; WHO 2021). Therefore, it should be addressed as an individual topic due to the number of indicators it comprises but is still considered within the working hours subcategory.

4 Method

The method for building scales and finding validity evidence for the 14 individual instruments (questionnaires) was structured on the psychometrics framework. International references such as Standards for Educational and Psychological Testing (AERA et al. 2014) and guidelines from the International Testing Commission were considered, in addition

to other references such as Pasquali (2010), Andrade and Valentini (2018), and Damásio and Borsa (2017).

4.1 A brief introduction to psychometrics

Psychometrics is a field of study that covers the theory and technique of measuring mental processes, especially applied in the fields of psychology, education, and social sciences in general. The field is grounded in the theory of measurement in sciences, in the quantitative method where the major concern is to represent subjective contents in a precise way (Pasquali 2009). Apprehending a reality requires measuring some quality connected with it. Thomson (1894) used to affirm that when a researcher can measure a phenomenon and express it in numbers, then he knows something about it. Still, if it is not possible to measure it, nor express it in numbers, then the knowledge about the phenomena is of an unsatisfactory kind, and would scarcely advance to the stage of science. For Spearman (1937), the measurement method is so vital that, if lacking, any study may be thought of as not scientific in the full sense of the word.

Furr (2021) expresses that psychometrics is the science dedicated to evaluating the attributes of psychological tests. In particular, three of these attributes are of interest: (1) the type of information (usually scores) generated by the use of data yielded from the test or scale; (2) the reliability of the data; and (3) aspects concerning the validity of the data. The author explains that the word “test” has often been used as a synonym for “measure,” “instrument,” “scale,” “inventory,” etc. Within this study, the words “scale” and “instrument” are used and should be understood as synonyms.

The psychometric theory is a well-recognized approach for developing measurement and standardized tools (Kim-O and Embretson 2010). Modern psychometrics can be divided into two fields: classical test theory (CTT) and item response theory (IRT). Although it is not a concern of this paper to portray the differences between them, it is worth mentioning that the scales were developed using CTT. Currently, psychometrics has evolved to cover all statistical methods that apply to studying behavioral and social sciences.

Measurement is intrinsic to the activity of science and involves the collection, quantification, and evaluation of data (DeVellis and Thorpe 2021). Quantitative science has existed since ancient times, and some examples of measured attributes are velocity, temperature, and length (Mitchell et al. 2017). Different instruments are used to measure observable attributes in the physical world, such as a scale to measure weight or unobservable attributes, e.g. a clock to measure time. Likewise, social scientists also measure observable phenomena, for instance, behaviors or facial expressions, or unobservable phenomena, such as intelligence, depression (Furr 2021), or the individual’s agreement or perception towards a certain topic. When measuring unobservable

phenomena, social scientists commonly use psychological tests or multi-item scales (DeVellis and Thorpe 2021).

Within this context, a measure can be understood as a single-item indicator or a multi-item instrument that quantifies the information of social or behavioral constructs at the individual or societal levels. A multi-item scale, or just “scale,” comprises some items (statements) that measure a variable of interest (construct) reliably and validly, and yield parametric data (Kline 2013). It is also relevant to mention that a rating scale is commonly confused with a multi-item scale. The first measures the response itself in a rating schema, for instance, a Likert scale consisting of 5 points. In contrast, a psychometric scale measures participants against a focal variable using multiple items (DeVellis 2006).

Social scientists refer to the theoretical concepts to be measured as hypothetical constructs or latent variables. Depending on the variable, they might comprise different characteristics, attributes, processes, or states that cannot be directly observed, and they include things such as knowledge, personality traits, and attention (Furr 2021). A social construct may assess aspects related to social systems as culture and norms that affect individuals and groups, whereas behavioral constructs capture neurocognitive processes and behavioral action (Klemke et al. 1980). When using a scale, scores are extracted from the items’ responses to measure the attributes related to the construct. Most of the scores derived from a scale are denominated as effect indicators since the way a participant responds to a scale is partly a reflection of the hypothetical construct (Furr 2021). The responses to the effect indicator reflect a latent variable (theoretical construct) that is the common cause of the responses (Murray and Booth 2018).

A scale should be population-specific, as it is not possible to assume that a given measure works across diverse populations living in different contexts. However, a scale can be adapted using specific procedures related to a technique named cultural adaptation, where not only the scale is translated to another language, but other aspects such as cultural, socioeconomic, and the norms and values associated with another culture are addressed. The more a scale can be validated, tested, and used across different populations, the more robust the measure (Borsa et al. 2012). A scale may be of a self-report format of response (individual response about his or herself) or hetero-report (individual response related to the perception of other people). All scales developed in this study were on a self-report basis, exceptions being the child labor and forced labor scales, due to the fact that people under these circumstances would seldom be reachable to respond to the scales.

In practical terms, to set up the scope of what the construct should measure, a literature review on the latent variable is necessary and a small theory that shall (1) establish the system or scope of the measure, (2) characterize the properties

of the construct, (3) identify the theoretical dimensions, (4) elaborate the constitutive and operational definitions, (5) use operationalization which is to create phrases on how the construct is manifested through behavioral sentences that will constitute the items of the scale in a later step, and (6) analyze the items in terms of its content on the construct, with experts, and semantics, with the target audience (Pasquali 2010; Furr 2021). See the theoretical pole in Fig. 1.

All the scales developed under this study considered ILO's conventions, recommendations, and related reports as a means to build the theories regarding each one of the subcategories that were identified here as the "construct" to be measured, except for the "working relationship" and "recognition at work" scales that were built upon correspondent and widely accepted theoretical backgrounds coming from labor psychology. The use of ILO as a main reference is because a comprehensive and consolidated understanding of the subcategories is found due to the accumulated knowledge derived from a systematic study of the labor world during the last decades as well as for the international standards on labor practices and working conditions for both developing and developed countries.

A construct may be unidimensional, bidimensional, or multidimensional. This means that when defining the scope of a measure, it is possible to find different attributes regarding the construct, comprising one or more forms of its manifestation within a theory of choice. Each of the ways in

which a construct is manifested, considering a given theory or literature on the theme, is called theoretical dimension. Therefore, a theoretical dimension is a descriptive model that represents a specific form of manifestation of the construct to be measured.

To measure the theoretical dimension, a set of operational definitions (items) are created to assess every dimension. An operational definition is a sentence that represents the theoretical dimension in terms of behavior. The set of operational definitions will later be assessed in empirical terms when applying the scale and running the statistical investigation known as exploratory factor analysis (EFA), which will indicate whether the dimensions are empirically confirmed and consistent with the theory (Pasquali 2010).

Another main statistical investigation that takes place when building a measure is the confirmatory factor analysis (CFA). It helps to further confirm the factors found in the EFA step and checks the internal structure of the scale. The latter is related to the notion that a scale, unlike a questionnaire, is not a partitionable instrument, and has to be applied as a whole, since the items influence themselves varying in how accurate the corresponding indices of adequacy and reliability of the measure (Furr 2021). When a theoretical dimension is empirically confirmed in the development of a scale, after running the EFA and CFA, it is then named as a factor. As aforementioned, the factors are the dimensions

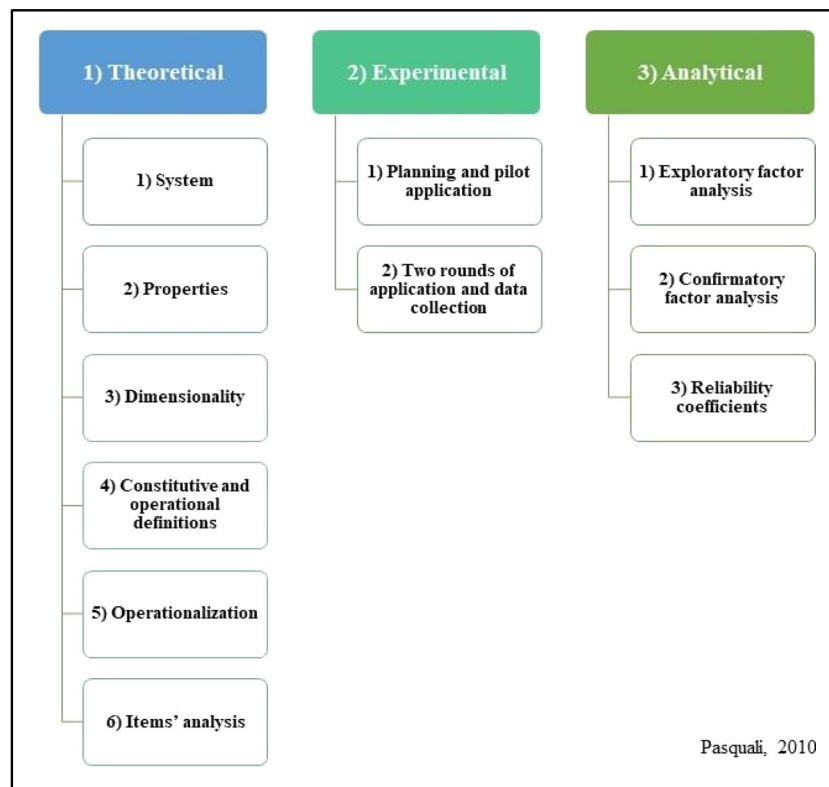


Fig. 1 Poles and methodological procedures for developing scales

of the measured construct, which are further investigated and not only theoretically found, but empirically given. In the scope of the present work, a factor is understood as the indicator related to an impact subcategory. In a parallel way, just as the indicator is a measure of the subcategory, the factor is a measure of the construct, where factor and indicator reciprocate and are synonyms within the scope of this study.

The procedures regarding the pilot and final application are considered in the experimental pole, while EFA, CFA, and reliability coefficients are addressed on the third step, or analytical pole. The question of reliability and validity is of great concern, being regarded through several methodological steps, and analyses to identify whether a measure is robust or not (AERA, APA, NCME, 2014). Validity is the most fundamental consideration in developing and evaluating tests. Whereas unreliable scales prevent the consistent measurement of a variable, scales low in validity may indeed not be measuring the variable at all (Cook and Beckman 2006). “Validity refers to the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests” (AERA, APA, NCME, 2014, p. 11).

Hence, validity is the degree to which the measure reflects what is supposed to be measured other than something else (Hays et al. 1993). Currently, there are five types of validity: (1) content validity, (2) evidence based on the internal structure, (3) evidence based on relations to other variables, (4) validity based on response process, and (5) evidence for validity and consequences of testing (AERA, APA, NCME, 2014). The types of validity assessed for the scales comprised the two first ones, as these are the most applicable to the characteristics of the scales. Content validity comprises the experts and semantics analysis (further detailed in topic 4.2).

Reliability coefficients speak about the precision of the scale, being defined as the correlation between scores on two same or equivalent forms of the instrument, assuming that one form does not affect the performance of the second form (AERA, APA, NCME, 2014). It refers to the extent to which a measure results in the same score each time it is applied, all else being equal (Hays et al. 1993). In other words, an errorless measurement means that the same instrument applied on the same subjects, but at different moments in time, produces identical outcomes. Hence, the correlation between the two measurements would be equal to 1 (Pasquali 2009). Therefore, reliability coefficients vary from 0 to 1.

5 Methodological steps

The study was structured in three main poles, namely (1) theoretical procedures, (2) experimental procedures, and (3) analytical procedures (Pasquali 2010).

5.1 Theoretical pole

Pasquali (2010) proposes that the theoretical pole is comprised of initiatives concerning the literature review and the mapping of the references regarding each one of the subcategories to deepen the understanding of each one. Then, the review was used in the definition of the theoretical framework adopted as an explanatory model of the phenomenon that was intended to be measured. Thereby, the following steps are conducted in the theoretical pole: (1) system definition, (2) property, (3) dimensionality, (4) constitutive and operational definitions, (5) operationalization, and (6) analysis of the items.

In this step, the scope of analysis (system) was defined, being characterized by the properties and aspects identified in the literature. Subsequently, the dimensions or parts that make up the subcategory to be assessed were described according to the selected theoretical backgrounds. Finally, the construct to be measured was defined, conceptually and then operationally. The operational definitions indicate the action or representation related to the assessed phenomenon, and were used in the formulation of the items in a later stage called “operationalization.” It is noteworthy that extreme care was taken to avoid double counting when elaborating the items, due to the similarities of topics that some subcategories have. Five focus groups were run to deepen the understanding of the themes studied in the context of work, addressing all the themes foreseen in the scales, except for child labor and forced labor, due to the sensitivity of the subject and the difficulty in finding workers who could report situations related to these two subcategories.

The last stage of the theoretical pole comprised the analysis of the items, and it carried out procedures to find validity evidence based on content through semantic analyses and judges’ analyses. The semantic analysis consisted of bringing together individuals from the target audience to verify their understanding as well as the clarity of the items, making minor, partial, or major changes to the items’ wording. Approximately five individuals participated in each of the six focus groups for semantic analyses. In the judges’ analyses, three specialists in S-LCA and a psychometrist were asked to assign to which theoretical dimension the items of each scale belonged, as well as the clarity of each item, assigning a score from 0 to 5. The agreement analysis was carried out using a method known as Kappa’s Fleiss, which is a statistical analysis that calculates the agreement between judges, being the most appropriate method since the analysis comprised four different judges for each scale (Furr 2011).

5.2 Experimental pole

The experimental pole included planning and application of the scales. The target audience of this step comprised of

workers over 18 years old, from any hierarchical level and position, linked to any type of work organization, formal and informal, and from any company size and economic sector. Also, the format, instructions, and systematic application of each of the scales were defined.

5.3 Analytical pole

The first and second phases of the scales' application aimed respectively at running EFA and CFA, where the same following procedures were applied. The scales were organized into five different sets, consisting of three to five scales, to receive a greater amount of participation from the individual respondents. Therefore, the sets of scales were organized as follows: (1) bullying at work, sexual harassment and sexual harassment in the organizational context (subscale), and equal opportunities and discrimination; (2) working relationship, recognition, employment relationship, and social security and social benefit; (3) fair salary, working hours, and remote work; (4) forced labor, child labor, and freedom of association and collective bargaining; and (5) the five scales related to health and safety—qualification and training in health and safety, health and safety promotion, risk and accident prevention and protection, organizational policy and compliance to the laws of the country, and working conditions in the context of health and safety.

As strategies for applying the scales, a partnership was established with a national company that has a user email database composed of more than five million workers from different economic sectors and hierarchical levels. Online links were sent via email to the user database, asking them to respond to the survey related to assessing decent work. For each user, only one link was sent, consisting of 3 to 5 scales—depending on the themes. In addition, the child labor and the forced labor scales were disseminated among governmental and non-governmental organizations that are focused on monitoring and tackling these forms of work. Finally, the links to respond to the scales were also sent to a convenience sample.

The links to which participants were invited to respond comprised of the following instruments: (a) a Free and Informed Consent Form, where the confidentiality of the information obtained was assured to the participants; (b) sociodemographic questionnaire (age, sex, education, religion, sexual orientation, etc.) and work organization information (job position, hierarchical level, economic sector, employment relationship, etc.); and (c) the developed scales covering the subcategories of the worker stakeholder.

The analytical pole comprises the EFAs and CFAs. The EFAs aimed at detecting latent structure which verifies whether the proposed theoretical model would be empirically identified for the given population to which data was collected. EFAs were performed using Factor software

version 12.03.02, and comprised a total of 791 participants. All analyses implemented incorporated Pearson's correlation matrix and robust diagonally weighted least squares (RDWLS) extraction method (Asparouhov and Muthen 2010). As for the method for retaining factors, the parallel analysis technique was used with a random permutation of the observed data (Timmerman and Lorenzo-Seva 2011). In some cases, the eigenvalues method was considered because it is more compatible and equivalent to the theoretical and empirical model employed. Concerning rotation, Robust Promin was applied to all cases except for the child labor scale (varimax), since its factors were uncorrelated (Ferrando and Lorenzo-Seva 2018). The samples obtained for each scale varied in numeric and descriptive terms.

At a later step, the CFAs were run using data from the second application, to confirm the factor structure identified in the EFAs. CFAs comprised of, overall, 1023 participants, and the JASP software version 0.16.4 was employed to run the analyses. The estimation method used was RDWLS, suitable for categorical data (DiStefano and Morgan 2014; Li 2016). The adjustment indices analyzed both for EFA and CFA were χ^2 (chi-squared), χ^2/df (degrees of freedom), Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), standardized root mean residual (SRMR), and root mean square error of approximation (RMSEA). Finally, the following reliability coefficients were assessed for each scale's factor(s): Cronbach's alpha, McDonald's omega, Guttman's lambda, and composite reliability. A brief explanation of the abovementioned indices and how to interpret them is found in the respective content in the "Results" section.

6 Results

The results were of a large proportion and were synthesized to be contained in this paper. They were organized according to the first and second methodological poles (theoretical and analytical) because the experimental consists of a preparation for the following step.

6.1 Theoretical pole

The system was defined as "worker in the organizational context," common to all constructs that in turn coincided with the subcategories. The theoretical dimensions (Table 1) were defined according to the properties of each construct, identified in the literature review, both in ILO's conventions, recommendations, reports, and in the literature coming from labor psychology. The definition of each indicator will follow with the publications concerning each scale. Regarding the semantics analyses, all the corrections of items were made during the focus groups, so the participants could assess the new texting, discuss, and reach together a

Table 1 System (stakeholder), constructs (subcategories), and factors (indicators)

System	Constructs (subcategories)	Theoretical dimensions	Factors (indicators)
Worker in the Organizational Context	Freedom of Association and Collective Bargaining	1) Trade unionism 2) Collective bargaining	1) Trade unionism 2) Collective bargaining
	Child Labour	1) Light work 2) Hazardous labour 3) The worst forms of child labour	1) Light work 2) Hazardous labour 3) The worst forms of child labour
	Forced Labour	1) Involuntary nature of forced labour 2) Work exacted under the menace of a penalty	Involuntary nature and maintenance of forced labour
	Fair Salary	1) Living wage 2) Equanimous wage 3) Equal wage	1) Living wage 2) Equanimous wage 3) Equal wage
	Working Hours	1) Workday 2) Paid rest 3) Overtime	1) Workday 2) Paid rest 3) Overtime 4) Working hours regulation
	Remote Work	1) Remote workday 2) Balance between professional and personal life 3) Working conditions in the context of remote work	1) Remote workday 2) Balance between professional and personal life 3) Working conditions in the context of remote work
	Equal Opportunities and Discrimination	1) Discrimination 2) Equal opportunities 3) Equal treatment 4) Discrimination in the organizational context	1) Discrimination 2) Equal opportunities 3) Discrimination in the organizational context

common agreement in a transparent and harmonized way. The modifications on the items were classified as 31% minor, 51% partial, and 17% major corrections.

The judges' analysis resulted in Kappa's Fleiss values that ranged from moderate to near perfect agreement on all scales: being 26.3% of them almost perfect agreement, 47.3% of strong agreement, and 26.3% of moderate agreement. The average clarity of the items was 4.56 out of 5, which double confirmed the clear semantics' understanding of the sentences (items). As referred before, Kappa's Fleiss is a statistical analysis that verifies the level of agreement between three or more experts. The higher the agreement, the higher the common understanding of a specific assessment. In our case, the higher the agreement, the better the items were framed to assess a certain indicator.

6.2 Analytical pole

The exploratory factor analysis and parallel analysis were run, resulting in a final factorial structure for each scale,

listed in Table 1. Parallel analysis is a Monte Carlo simulation technique based on random data simulation to determine the number of factors of a data sample. The analysis simulates a dataset based on the real one, and compares the simulated outcomes with the actual ones, suggesting the ideal number of factors to be considered (Ledesma and Mora 2007). As mentioned before, the number of factors is the same as the number of indicators comprised under an impact subcategory, according to the studied population.

The minimum and maximum values of the adjustment indices obtained for the EFA and CFA are displayed in Table 2, while Table 3 provides the thresholds to interpret each index. The specific indices, item loads, and their commonalities, both in EFA and CFA, as well as the specific reliability indices for each scale's factor(s), could not be included in this paper and will be further detailed in upcoming publications.

The indices portrayed in Table 2 are known as fit indices, measuring different aspects of model fit (Hair et al. 2009). In other words, they indicate how fit, or adjusted, the created

Table 1 (continued)

System	Constructs (subcategories)	Theoretical dimensions	Factors (indicators)
	Social Security and Social Benefits	1. Access to social security 2. Access to social benefits	1. Access to social security 2. Access to social benefits
	Employment Relationship	1. Employee 2. Self-employed – formal 3. Self-employed – informal	1. Employee 2. Self-employed – formal 3. Self-employed – informal
	Bullying at Work	1. Organizational bullying 2. Descending bullying at work 3. Ascending bullying at work 4. Horizontal bullying at work 5. Bullying at work in the organizational context	1. Descending bullying at work 2. Ascending and horizontal bullying at work 3. Bullying at work in the organizational context
	Recognition	1. Intrapersonal 2. Interpersonal	1. Intrapersonal 2. Interpersonal
	Working Relationship	1. Cooperation 2. Practical intelligence 3. Public space for discussion 3. Vertical relationships 4. Real and Prescribed at work	1. Cooperation 2. Practical intelligence 3. Vertical and institutional relationships
	Health and Safety	1. Training and capacity building 2. Health and safety promotion 3. Prevention and protection against risks and accidents 4. Organizational policy and compliance to the laws of the country 5. Working conditions in the context of health and safety	1. Training and capacity building 2. Health and safety promotion 3. Prevention and protection against risks and accidents 4. Organizational policy and compliance to the laws of the country 5. Working conditions in the context of health and safety
	Sexual Harassment	1. Verbal harassment 2. Physical harassment 3. Non-verbal harassment 4. Sexual harassment in the organizational context (subscale)	1. Sexual harassment manifestations at work 2. Sexual harassment in the organizational context (subscale)

model is to measure a certain phenomenon in a given population. They should be analyzed altogether and their final interpretation indicates if the developed instrument (scale) is adjusted or not. To better portray what these indices assess, a brief explanation follows (Boateng et al. 2018): the chi-square value is a test statistic that checks the goodness of fit of a factor model, comparing the observed covariance matrix with a theoretical one. TLI is a comparison between the proposed factor model to another in which no interrelationships are assumed among any of the items, whereas CFI is an incremental relative fit index to which a comparison between the fit of the proposed model and the baseline is made to measure the relative improvement.

RMSEA measures the estimated difference between the covariance matrices per degree of freedom from the population responding to the instrument and the model-implied population. SRMR is a measure of the mean absolute correlation residual and allows to assess the magnitude of the discrepancies between the observed and expected correlations. The explained variance refers to how much the model can explain the variation in the data, or how much the scale measures the intended phenomena. In social sciences, where data are more subjective and information is less precise, it is commonly considered a solution that accounts for 60% of the total variances, and in some cases, even less, considered satisfactory (Hair 2014).

Table 4 shows the minimum and maximum values obtained for the reliability coefficients for all of the respective derived from the 14 scales. Table 5 provides the thresholds to interpret each coefficient. Reliability is a measure varying from 0 to 1 and indicates the precision and replicability of the factor. Thus, not only the adjustment of the scales is assessed, but their accuracy as a measurement tool. A brief example of three out of the 14 scales and a few of their items is provided in Tables 6, 7, and 8 to illustrate their format and content. They pertain to the scales of child labor, remote work, and sexual harassment, respectively. It is relevant to mention that while the child labor scale is a hetero-reported instrument (replied by other stakeholders, such as local community and other workers and not the children themselves), the remote work and sexual harassment scales are self-reported.

The items are responded to in terms of agreement on a Likert scale that varies from 1 (I fully disagree) to 5 (I fully agree). The selected items comprise only one of the indicators encompassed in each scale, and a definition of the indicator is provided in the first row for a better understanding of this work. However, the definition of the indicator is not presented to workers responding to the scales. It is noteworthy to endorse that this is a mere exemplification and a direct translation from the scales originally developed in the Portuguese language. A cultural adaptation of all scales will be followed for the English version, which will be applicable to workers fluent in English and published subsequently.

7 Discussion

The results from the content analysis (semantics and experts) showed that the 14 scales had evidence of content validation, as items were adjusted in terms of semantics according to the focus groups with the target audience, and the experts' analyses indicated by the Kappa's Fleiss with most results varying from strong to almost perfect levels of agreement. Both semantics and content analysis point out the clear understanding of the sentences of the scales, and the

Table 2 Minimum and maximum values obtained for EFA and CFA indices

Ranges of EFA indices	Ranges of CFA indices
χ^2/gl : 1.1–2.54	χ^2/gl : 1.3–3.51
CFI: 0.97–1.00	CFI: 0.97–1.00
TLI: 0.97–1.00	TLI: 0.97–1.00
SRMR: 0.00–0.04	SRMR: 0.03–0.08
RMSEA: 0.00–0.8	RMSEA: 0.04–0.10
Explained variance: 54–74%	-

Table 3 Adjustment indices and respective thresholds for EFA and CFA

Adjustment indices	Threshold
χ^2/gl	< 5, preferably < 3 (Brown 2015)
CLI	> 0.90, preferably, 0.95 (Hu & Bentler 1999)
TLI	> 0.90, preferably, 0.95 (Hu & Bentler 1999)
RMSEA	≤ 0.05 close fit ≤ 0.08 fair fit (Browne, 1993)
SRMR	≤ 0.10 (Hair et al. 2009)

pertinence of their content to assess the indicators within the framework set in S-LCA.

In terms of the search for evidence based on the internal structure of the scales, it was possible to observe that all EFA ran for the 14 scales indicated good to excellent fit indices to detect the latent structure, i.e., the association of the items with the factors (indicators). Parallel analysis was run to identify the retention solutions for the number of factors for each scale (Table 1), resulting in an identical solution foreseen in the theoretical dimension for 9 scales and different solutions for five scales, which were equally plausible in theoretical terms. The scales that had a different factorial solution were as follows: (1) forced labor, unifactorial rather than bifactorial, not discriminating between the two basic elements of forced labor; (2) working hours, indicated four instead of three factors, however with an agreeable structure and convergent with the theory; (3) equal opportunities and discrimination, three instead of four factors, as one was theoretically overestimated and it was empirically verified that equal treatment was comprised within the equal opportunities factor; (4) sexual harassment, unifactorial rather than three factors, converging to a single typology of manifestations of sexual harassment; (5) working relationship, three factors instead of the five initially foreseen, indicating that the perception of hierarchical relations and the institutional relations are intertwined; and (6) bullying at work, three instead of five factors, demonstrating that the manifestation of harassment is more characterized by who is the agent than the victim, endorsing the perception of power in the expression of bullying at the workplace.

As for the confirmation of the structure of the scales, 12 reached adjustment indices of the proposed assessment model to the data samples with good to excellent values for all indices. The only exceptions were the working hours

Table 4 Reliability coefficients

Reliability coefficient
Composite reliability: 0.85–0.98
Cronbach's alpha: 0.81–0.98
McDonald's omega: 0.82–0.97
Guttman's lambda 2: 0.83–0.98

Table 5 Reliability coefficients and respective thresholds

Reliability coefficients	Threshold
Composite reliability	> 0.80 (Raykov 1997)
Cronbach's alpha	≥0.70 acceptable ≥0,80 good ≥0,90 excellent (Cortina, 1993)
McDonald omega	≥0.70 acceptable ≥0,80 good ≥0,90 excellent (Cortina, 1993)
Guttman's lambda	≥0.70 acceptable ≥0,80 good ≥0,90 excellent (Guttman 1945)
H index	> 0.80 (Ferrando & Lorenzo-Seva 2018)

Table 6 Excerpt from the child labor scale

Light work: related to paid child work, referred to as light work, which does not harm the health and development of the child, and working hours that do not affect school attendance. It may be carried out by children aged between 13 and 15 years, depending on the national legislation of each country. It is regulated and, therefore, the employer checks the age of the children before hiring them

I've seen children...

- 1 Having their age checked by the company before they start working
- 2 Receive training before they start working
- 3 Having food provided by the company
- 4 Working with protective equipment
- 5 Working, but also attending school
- 6 Having their work supervised by adults

Table 7 Excerpt from the remote work scale

Remote workday: remote working arrangement that refers to: (1) working shift, workload, and number of hours worked; (2) type of remote work carried out: occasional, additional, partial, or substitutive; and (3) diverse options for flexible working hours, and whether positive or negative for the worker

When I do remote work...

- 1 I work flexible hours
- 2 I work on weekdays and at weekends
- 3 I usually have to work outside business hours to complete my tasks
- 4 I feel I need to pay attention to work-related phone calls and emails, even outside working hours
- 5 I feel that I must be available anytime to people from work who need me
- 6 I can set my own working hours
- 7 I work without rest or meal breaks
- 8 My boss or co-worker(s) often contact me outside of working hours

and recognition at work scales that indicated a marginal RMSEA value. However, they both presented values considered excellent for all other adjustment indices, and, therefore, their models were considered adjusted (Hair et al. 2009; Bentler 2007; Hayduk, et al. 2007). As for the reliability indices, all showed precision rates with values above 0.80 for all factors, ranging from good to excellent, attesting to a high level of accuracy for all indicators from all instruments.

The scales gathered the required criteria to allow comparability across S-LCA studies when addressing the same subcategories from the worker stakeholder. Not only that, it allows for comparability among different materials, products, countries, etc., which will further support one of the main goals of S-LCA: being a methodology to support

decision-making and improving social strategies towards human well-being (Benoît et al. 2013). The application of the scales to different populations working in the same organization, for instance, will shed more light on the differences and details concerning working conditions for different types of workers (e.g., women or migrants), and fully grasp the social risks taking place (Huertas-Valdivia et al. 2020).

Since the worker is one of the most impacted stakeholders in S-LCA, developing mechanisms that foster employee voice could be pivotal (Guimarães et al. 2018) to promote more socially sustainable practices (Huertas-Valdivia et al. 2020). Hearing directly the voices of workers would ensure a representative sample of respondents, and may avoid

Table 8 Excerpt from the sexual harassment scale

Sexual harassment in the organizational context: set of actions taken by the organization to prevent, investigate, combat, and punish sexual harassment at work, as well as to promote a healthy, respectful, and ethical environment among workers

- 1 The organization provides means for reporting sexual harassment at work
- 2 The organization has a policy to combat sexual harassment at work
- 3 The policy to combat sexual harassment at work is widely publicized among workers
- 4 The policy to combat sexual harassment at work is easy to understand
- 5 The policy to combat sexual harassment at work is practical and applicable
- 6 People who commit sexual harassment and have been convicted are punished
- 7 There are actions that provide support for people who are victims of sexual harassment
- 8 The organization pays little attention to complaints of sexual harassment
- 9 There are anonymous ways to report sexual harassment at work
- 10 After the complaint, the work organization investigates the facts impartially and discretely
- 11 I'm afraid to report the harasser because I don't trust the organization to handle the case appropriately

situations in which those in authority are motivated to provide incorrect information (Grubert 2021). Moreover, data collection from self-reported instruments can be used for triangulation and provide a wider understanding and validation of other types of data, whether from primary or secondary sources.

8 Recommendations

8.1 How to apply the scales

First and foremost, it is relevant to mention that psychometrics was only a method to develop the measurement tools (scales) that were tailored primarily to S-LCA, according to the framework established in the guidelines for S-LCA (UNEP 2020). Therefore, the proposed indicators were first conceived based on the theoretical background and working international standards from ILO and related literature, where after applying the scales on workers, in two different rounds, the statistics analysis of the EFA and CFA demonstrated whether the reality of the world of work endorsed, or not, the indicators found in the literature, tying theory with real life. In this way, the proposed scales are accurate tools to measure different social aspects related to workers. It is noteworthy that each indicator is assessed by multiple sentences that cover a large variety of aspects of the social phenomena, different from the common practice of one “single question” to evaluate an indicator, which can sometimes be complex and multi-faceted. Moreover, the results of the indicators using the scales can be fully applied to the same methods currently found in the literature to assess social impact.

The scales can be applied independently, which means that one or more scales can be applied to the same population of workers. They can be applied at the same time, or in different moments. However, it is also important to bear in

mind that social phenomena are dynamic and the circumstances causing a specific effect can change abruptly over time, depending on the type of phenomenon itself and the surrounding circumstances. Therefore, it is recommended to prevent long periods between one application and another of different scales to avoid a mismatch of the circumstances and the assessment's results. In the same way, the scales can be applied online or in person, although processing and analyzing data can be easier when in the online format.

In terms of defining which subcategory to assess, the guidelines for S-LCA (UNEP 2020) recommend three steps for prioritizing data collection: (1) literature review of the studied system; (2) identify the most active or intensive activities/unit processes; and (3) identify the hotspots in the product's life cycle, for risk and opportunity assessment. There are also various procedures focusing on materiality assessment to screen which stakeholders and/or topics are most relevant and need to be assessed (Wu et al. 2014; UNEP 2020). Indicators relating to the same stakeholder are usually intertwined and it is possible that under the same company, workers in different positions and from different backgrounds have different working conditions and face different realities. Therefore, when addressing social aspects, the subcategory indicators should be assessed thoroughly, as much as possible, to better portray the reality within a product system or specific organization.

With regard to the definition of the functional unit (FU) to model the product system, there are no differences when collecting data using the scales, and the S-LCA practitioner may use the methods that suit better the goal and scope of the study, although it is advisable caution as this is not an easy task. When defining an FU, applying an activity variable is necessary to measure the process output and relativize the social indicators according to each unit process (Benoit-Norris and Norris 2018). It does not represent an impact but an elementary flow used to compare the processes and

aggregate the impact results (UNEP 2020). Currently, the most commonly applied activity variable in S-LCA studies is worker hours; however, just a few social indicators are related to it, making no sense to translate all social indicators.

Another noticeable challenge of linking FU to social impacts relies on the fact that social impacts are more related to the behavior of a work organization than the function delivered by a product (Dreyer et al. 2006; Jørgensen 2013; Hosseini et al. 2014). In the same direction, Zamagni et al. (2011) endorse the company-based perspective in the S-LCA framework, and advocate to connect the social impacts to the conduct of a company and not to the function given by a product. It is also true that not all social impacts assessed in a S-LCA are related exclusively to a company's practices. Externalities such as lack of governance and education, poverty, and corruption also play a decisive role in impacting, either positively or negatively, the affected stakeholders. Far from proposing a solution for the problem of FU in S-LCA, this study merely reinforces the previous concerns of other authors and emphasizes the need to better understand the dynamics of social phenomena. Thus, it is expected that future studies may develop a more consistent and socially based approach to represent and aggregate social impacts.

All the scales can be used to assess any product or service, from any industry sector. In the same way, the scales are applicable for both type I (reference scale) and type II (impact pathway) social impact assessment approaches. For the reference scale approach, the currently available methods such as subcategory assessment method (SAM) (Ramirez et al. 2014) may be applied using the results from the different scales. Due to the format of the response (5-point Likert scale), the indicators will vary from 1 to 5. Reference values should be developed, based on a country's regulations, and socioeconomic and cultural conditions to compare the outcomes to the benchmark values and verify compliance, or performance under or above a given reference. Likewise, corresponding scale varying from $+X$ to $-X$ can be used to translate the potential impacts.

In the case of impact pathways, data yielded from the scales can be especially applicable for statistical analysis when causality is implied, for instance, in regression models and structural equation modeling. However, when using the impact pathway approach, a reverse logistic perspective about the impact to be evaluated is also of utmost relevance to secure consistency of the study. In other words, it is important to conduct a literature review about the impact itself to better identify which indicators might be related to it. Otherwise, the causal relationship may not be found, or result in a biased outcome. Without a deep understanding of the theoretical background of the impact and its pathway or chain, it is not possible to guarantee which are the relevant indicators that must be assessed.

Data collected through the scales can be used also for triangulation, or it can be triangulated with other data sources, e.g., secondary data (national statistics, databases, literature, etc.) and other primary data (human resources department, company records, etc.). Nevertheless, when a study seeks to collect primary data regarding workers, the workers themselves should be heard to the fullest extent, since they are the stakeholder in question. It is recommended that, in cases where the target population is semi-literate or illiterate, possible strategies can be applied for data collection, namely (1) identify someone trusted by the workers and/or the community (depending on the scope of research) and train him/her on the application of the scale(s) and (2) use the scale items as a structured or semi-structured questionnaire for direct application with workers.

The scientific and robust approach that the scales bring will ease primary data collection for small or large samples of workers. When the population is known and small, it would allow collection of data (ideally) from the entire population under investigation. There is no doubt that the more data is collected, the more informative a case study will be (Lakens 2022). On the other direction, when the sample size is too small for its population, it might not yield valid results that represent adequately the reality of the entire population at stake (Kibuacha 2021). In situations in which the organization is large and it is not possible to collect data from every worker due to limited financial and time resources, a sample calculation is an alternative to secure a reliable representation of the studied population and results of the analysis.

When doing so, it is important to previously know a set of information to determine the sample size required. One of them is the overall number of the specific population. It is relevant to bear in mind that one work organization can have different populations, due to different characteristics or working conditions. For instance, if the study aims to compare the working conditions of migrants and locals, the total number of these populations is different and two sampling calculations would be necessary. Otherwise, considering only the total number of employees of a company is enough to determine the sample size. Another example of the information required to calculate the sample size is the confidence interval and confidence level. Currently, there are online calculators and software such as GPower that make this task easier (Lakens 2022).

Data yielded from the scales are collected in a completely confidential way as no information regarding the same or other forms of identification is required. A survey including questions about age, gender, ethical origin, job position, and education level, among others, can be included in the assessment to better stratify the population and widen the scope of analysis. Finally, the scales can be applicable for the foreground and background processes, as depending on

the relationship with the supply chain companies, collecting primary data using an online tool can be a reachable possibility. When collecting primary data relating to workers is not possible for the background processes, it is advisable to use secondary data while collecting primary for the foreground processes.

8.2 How to calculate the indicators

The 14 instruments were created with a 5-point Likert rating scale, assessing either agreement or frequency, depending on the topic assessed. After the application, scores can be created for substantive analysis, using (1) weighted or (2) unweighted procedures. The unweighted approach consists of using the standardized item scores or raw item scores, which means employing the raw worker's responses, for calculating the factor score (indicator) by adding them up or calculating the average among the items corresponding to a factor (Armor 1973). When the scale has only one factor, both summing the items' scores and computing the average is recommendable. When the scale has one or more factors and an equal amount of items to each factor, also both calculations can be applicable.

However, when a scale has more than one factor, but a different amount of items per factor, taking the average from the items' raw scores would be advisable. However, when the researcher is applying more than one scale, for easier interpretation and comparison purposes, computing the average is the best practice as all subcategories and their respective indicators would be at the same rating scale. A final scale score is also possible by calculating the indicators' average. It is also possible to weigh the indicators, according to their relative importance, by applying different calculation approaches.

The weighted procedure, which does not relate to weighting in S-LCA, is the calculation of the factor score by applying statistical software programs such as Mplus, R, SAS, SPSS, and Stata. (Boateng et al. 2018). The statistical procedure analyzes the different contributions of each item of an instrument, according to the sample's responses, and is represented in the calculated item score (Armor 1973), with the factor loading value representing the weight (Hair 2014). However, there is not a significant difference in the performance of the scale if it is computed as weighted or unweighted items (Boateng et al. 2018).

Finally, there are scales with reverse-scored items that go in the opposite direction of the majority of items. For instance, in the remote work scale, most of the items have positively worded sentences that assess the potential benefits of remote work, such as "I can balance work time with my family duties," and a reverse scored item would be "My boss or co-worker(s) often contact me outside working hours." For the latter item, the scoring runs in the opposite direction

and if rated by the respondent as "5" (strongly agree), it should be assigned as "1" (strongly disagree), i.e., the reverse score, when computing the indicators or the final composite score of the scale. All the reverse-scored items will be signalized in the published scales.

9 Conclusions

The scales comprise a powerful tool and source for investigation, study, and understanding of the different contexts related to work, and can be applied collectively or individually since they are independent and aim to cover different topics. Beyond the scope of S-LCA, the scales may also support other research initiatives that seek to create diagnostics and improve working conditions for all workers, from any economic sector. The results yielded from the scales can also serve to identify potential topics of greater criticality for research prioritization and intervention. One major aspect of social topics is how dynamic they may be. Depending on the phenomena, a minor change in an indirect variable can have a great effect on a social indicator. Thus, they may change fast and must be assessed preferably with primary data to truly grasp what has been taking place in a working organization, within a specific timeframe.

It is expected that the present study may contribute to overcoming some of the aforementioned methodological and theoretical gaps in the S-LCA: (1) facilitate and streamline primary data collection in a standardized way; (2) propose a list of harmonized and empirically evidenced set of indicators; (3) promote comparability among studies; (4) allow a deep mapping of sensitive and confidential information that, without the confidentiality of a scale, would be seldom achieved; and (5) provide with reliable data to assess social impact regarding worker for both types of impact—reference scale and impact pathways.

This study faced some limitations. The child labor and forced labor subcategories were not covered in the focus group, due to the natural constraints of reaching out to people under these conditions to deepen the understanding of both contexts. To maximize the number of participants in the two rounds of applications, the scales were organized in a set of 3 to 5 instruments to be responded to at once. This strategy may have posed difficulties to some participants to finish responding to the scales and a reduced number of cases was registered, although the samples were enough in number to achieve the statistical assumptions of factor analysis and the adjust indices were met.

In terms of the research agenda, it is relevant to mention that the first set of scales was developed in the Portuguese language, although the English version is already under development, and a cultural adaptation to other languages, such as Spanish, German, and French, is expected to

be delivered next, in order to further allow application and comparability throughout studies from different geographic and cultural backgrounds. Furthermore, the next step may be the use of data yielded from the scales in statistics modeling in which causality is implied, such as regression and structural equation modeling to analyze the relationship between the indicators and how much they influence a given impact regarding the worker stakeholder.

Funding Open Access funding enabled and organized by Projekt DEAL.

Data availability Data will be made available from the authors upon reasonable request.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). Standards for educational and psychological testing. AERA, APA, NCME.
- Andrade JM de, Valentini F (2018) Diretrizes para a Construção de Testes Psicológicos: A Resolução CFP nº 009/2018 em Destaque. *Psicol, Ciênc Prof (Impr)*, 38(1):28–39. <https://doi.org/10.1590/1982-3703000208890>
- Arcese G, Lucchetti MC, Massa I, Valente C (2018) State of the art in S-LCA: integrating literature review and automatic text analysis. *Int J Life Cycle Assess* 23:394–405
- Armor DJ (1973) Theta reliability and factor scaling. *Sociol Methodol* 5:17–50
- Arvidsson R, Baumann H, Hildenbrand J (2015) On the scientific justification of the use of working hours, child labour and property rights in social life cycle assessment: three topical reviews. *Int J Life Cycle Assess* 20:161–173
- Asparouhov, T., & Muthen, B. (2010). Simple second order chi-square correction. Unpublished manuscript.
- Benoît Norris C, Traverso M, Valdivia S, Vickery-Niederman G, Franze J, Aзуero L, Ciroth A, Mazijn B, Aulisio D (2013) The methodological sheets for sub-categories in social life cycle assessment (S-LCA). United Nations Environment Programme (UNEP) and Society for Environmental Toxicology and Chemistry (SETAC)
- Bentler PM (2007) On tests and indices for evaluating structural models. *Pers Individ Dif* 42(5):825–829
- Boateng GO, Neilands TB, Frongillo EA, Melgar-Quinonez HR, Young SL (2018) Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Front Epidemiol* 6:149
- Borsa JC, Damásio BF, Bandeira DR (2012) Cross-cultural adaptation and validation of psychological instruments: some considerations. *Paidéia (Ribeirão Preto)* 22:423–432
- Brown TA (2015) Confirmatory factor analysis for applied research. The Guilford Press
- Browne MW (1993) Testing structural equation models. (No Title), 136.
- Chabrawi AMRO, Andrade JMD, Traverzo M, Ugraya CML (2023) Bridging the gaps in social life cycle assessment: theoretical and practical contributions. *Psic Teor Pesq (Online)* 39:e39501
- Clot Y (2010) *Le travail à coeur*. Paris: La Découverte.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Int J Appl Psychol* 78(1):98–104. <https://doi.org/10.1037/0021-9010.78.1.98>
- Cook DA, Beckman TJ (2006) Current concepts in validity and reliability for psychometric instruments: theory and application. *Am J Med* 119(2):166–e7
- Costa D, Quinteiro P, Dias AC (2019) A systematic review of life cycle sustainability assessment: current state, methodological challenges, and implementation issues. *Sci Total Environ* 686:774–787
- Damásio BF, Borsa JC (Eds) (2017) *Manual de desenvolvimento de instrumentos psicológicos*. Vetor.
- Dejours C (1993) Coopération et construction de l'identité en situation de travail. *Futur Antérieur* 16(2):41–52
- Dejours C (2011) Entre sofrimento e reapropriação: o sentido do trabalho. In S. Lancman & L. Sznclwar. (Orgs.), *Christophe Dejours: da psicopatologia à Psicodinâmica do Trabalho*, (pp. 433–448). Rio de Janeiro: Fiocruz Brasília: Paralelo 15.
- Dejours C (2012) *Trabalho vivo: trabalho e emancipação* (Tomo 2, F. Soudant trad.). Brasília: Paralelo 15.
- DeVellis RF (2006) Classical test theory. *Med Care* 44(11):S50–S59
- DeVellis RF, Thorpe CT (2021) *Scale development: theory and applications*. Sage publications
- DiStefano C, Morgan GB (2014) A comparison of diagonal weighted least squares robust estimation techniques for ordinal data. *Struct Equ Modeling* 21(3):425–438. <https://doi.org/10.1080/10705511.2014.915373>
- Dreyer L, Hauschild M, Schierbeck J (2006) A framework for social life cycle impact assessment. *Int J Life Cycle Assess* 11(2):88–97
- Facas EP (2013) *Protocolo de avaliação dos riscos psicossociais no trabalho: Contribuições da psicodinâmica do trabalho* [Tese de Doutorado, Universidade de Brasília]. Repositório Institucional da Universidade de Brasília. <https://repositorio.unb.br/handle/10482/15420>
- Ferrando PJ, Lorenzo-Seva U (2018) Assessing the quality and appropriateness of factor solutions and factor score estimates in exploratory item factor analysis. *Educ Psychol Meas* 78(5):762–780. <https://doi.org/10.1177/0013164417719308>
- Freitas MED (2007) Quem paga a conta do assédio moral no trabalho?. *RAE eletrônica*, 6. <https://doi.org/10.1590/S1676-56482007000100011>
- Furr RM (2011) *Scale construction and psychometrics for social and personality psychology*. SAGE Publications Ltd. <https://doi.org/10.4135/9781446287866>
- Furr RM (2021) *Psychometrics: an introduction*. SAGE publications.
- Gompf K, Traverso M, Hetterich J (2020) Towards social life cycle assessment of mobility services: systematic literature review and the way forward. *Int J Life Cycle Assess* 25:1883–1909
- Groß-Fürter D, Mair-Bauernfeind C, Hesser F (2023) Proposing a multi-level assessment framework for social LCA and its contribution to the sustainable development goals. *Progress in Life Cycle Assessment* 2021. Springer International Publishing, Cham, pp 103–129

- Grubert E (2021) The need for social scientists in developing social life cycle assessment. Energy impacts: A multidisciplinary exploration of North American energy development, 74
- Guimarães AG, Vaz-Fernandes P, Ramos MR, Martinho AP (2018) Co-processing of hazardous waste: the perception of workers regarding sustainability and health issues in a Brazilian cement company. *J Clean Prod* 186:313–324
- Guttman L (1945) A basis for analyzing test-retest reliability. *Psychometrika* 10:255–282. <https://doi.org/10.1007/BF02288892>
- Hair JF, Black WC, Babin BJ, Anderson RE, Tatham RL (2009) *Análise multivariada de dados*. Bookman.
- Hair JF (2014) *Multivariate data analysis*. Bookman.
- Hayduk L, Cummings G, Boadu K, Pazderka-Robinson H, Boulianne S (2007) Testing! testing! one, two, three—testing the theory in structural equation models! *Pers Individ Dif* 42(5):841–850
- Hays RD, Anderson R, Revicki D (1993) Psychometric considerations in evaluating health-related quality of life measures. *Qual Life Res* 2:441–449
- Heloani R (2004) Assédio moral: um ensaio sobre a expropriação da dignidade no trabalho. *RAE eletrônica*, 3.
- Hirigoyen MF (2008) Assédio moral: a violência perversa no cotidiano. Bertrand Brasil, Rio de Janeiro
- Hirigoyen MF (2002) *Mal-estar no trabalho: redefinindo o assédio moral*. Bertrand Brasil.
- Hosseiniyou SA, Mansour S, Shirazi MA (2014) Social life cycle assessment for material selection: a case study of building materials. *Int J Life Cycle Assessment* 19:620–645
- Hu L, Bentler PM (1999) Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Modeling* 6(1):1–55. <https://doi.org/10.1080/1070519909540118>
- Huarachi DAR, Piekarski CM, Puglieri FN, de Francisco AC (2020) Past and future of social life cycle assessment: historical evolution and research trends. *J Clean Prod* 264:121506
- Huertas-Valdivia I, Ferrari AM, Settembre-Blundo D, García-Muñia FE (2020) Social life-cycle assessment: a review by bibliometric analysis. *Sustainability* 12(15):6211
- Hunkeler D (2006) Societal LCA methodology and case study. *Int J Life Cycle Assess* 11(6):371–382
- International Labour Office (ILO) (2020) Teleworking during the COVID-19 pandemic and beyond: a practical guide. https://www.ilo.org/wcmsp5/groups/public/---ed_protect/---protrav/---travail/documents/publication/wcms_751232.pdf
- Iofrida N, De Luca AI, Strano A, Gulisano G (2014) Social life cycle assessment in a constructivist realism perspective: a methodological proposal. In: *Social LCA in progress*. Pre-proceedings of the 4th international seminar in social LCA. Montpellier, France
- Jørgensen A (2013) Social LCA—a way ahead? *Int J Life Cycle Assess* 18:296–299
- Jørgensen A, Le Bocq A, Nazarkina L, Hauschild M (2008) Methodologies for social life cycle assessment. *Int J Life Cycle Assess* 13(2):96–103
- Kibuacha F (2021) How to determine sample size for a research study. Available in: <https://www.geopoll.com/blog/sample-size-research/>
- Kim-O MA, Embretson SE (2010) Item response theory and its application to measurement in behavioral medicine. *J. Behav. Med*, 113–123.
- Klemke ED, Hollinger R, Rudge DW, Kline AD (1980) Introductory readings in the philosophy of science.
- Kline P (2013) *Handbook of psychological testing*. Routledge
- Kühnen M, Hahn R (2017) Indicators in social life cycle assessment: a review of frameworks, theories, and empirical experience. *J Ind Ecol* 21(6):1547–1565. <https://doi.org/10.1111/jiec.12663>
- Lakens D (2022) Improving Your Statistical Inferences. Retrieved from https://lakens.github.io/statistical_inferences/. <https://doi.org/10.5281/zenodo.6409077>
- Lancman S (2011) O mundo do trabalho e a psicodinâmica do trabalho. In S. Lancman & L. Sznelwar. (Orgs.), *Christophe Dejours: da psicopatologia à Psicodinâmica do Trabalho*, (pp. 31–44). Rio de Janeiro: Fiocruz Brasília: Paralelo 15.
- Ledesma DR, Mora PV (2007) Determining the number of factors to retain in EFA: an easy-to-use computer program for carrying out parallel analysis. *Practica Assess Res Evaluation* 12(2):1–11
- Li CH (2016) Confirmatory factor analysis with ordinal data: comparing robust maximum likelihood and diagonally weighted least squares. *Behav Res Methods* 48(3):936–949. <https://doi.org/10.3758/s13428-015-0619-7>
- Macombe C, Leskinen P, Feschet P, Antikainen R (2013) Social life cycle assessment of biodiesel production at three levels: a literature review and development needs. *J Clean Prod* 52:205–216. <https://doi.org/10.1016/j.jclepro.2013.03.026>
- Macombe C, Loeillet D, Gillet C (2018) Extended community of peers and robustness of social LCA. *Int J Life Cycle Assess* 23:492–506.
- Di Martino V, Musri M (2001) *Guidance for the prevention of stress and violence at the workplace*. Ministry of Human Resources Malaysia and International Labour Organization, Kuala Lumpur.
- McCabe A, Halog A (2018) Exploring the potential of participatory systems thinking techniques in progressing SLCA. *Int J Life Cycle Assess* 23:739–750
- Messenger J (2019) *Telework in the 21st century*. Edward Elgar Publishing. <https://doi.org/10.4337/9781789903751>
- Mitchell DJ, Tal E, Chang H (2017) The making of measurement: editors' introduction. *Stud Hist Philos Sci* 65:1–7
- Moltessen A, Bonou A, Wangel A, Bozhilova-Kisheva KP (2018) Social life cycle assessment: an introduction. In *Life Cycle Assessment* (pp. 401–422). Springer, Cham.
- Norris GA (2006) Social impacts in product life cycles: towards life cycle attribute assessment. *Int J Life Cycle Assess* 11(1):97–104
- Pasquali L (2009) Psychometrics. *Rev Esc Enferm U S P* 43:992–999
- Pasquali L (2010) Testes referentes a construto: Teoria e modelos de construção. Em L. Pasquali, *Instrumentação psicológica: Fundamentos e práticas* (pp. 165–198). Artmed.
- Petti L, Serreli M, Di Cesare S (2018) Systematic literature review in social life cycle assessment. *T Int J Life Cycle Assess* 23:422–431
- Ramirez PKS, Petti L, Haberland NT, Ugaya CML (2014) Subcategory assessment method for social life cycle assessment. Part 1: methodological framework. *Int J Life Cycle Assess* 19:1515–1523
- Raykov T (1997) Estimation of composite reliability for congeneric measures. *Appl Psychol Meas* 21(2):173–184. <https://doi.org/10.1177/01466216970212006>
- Robert F (2018) Impact of workplace bullying on job performance and job stress. *J Manag Info* 5(3):12–15
- Rugulies R, Madsen IE, Hjärsbech PU, Hogh A, Borg V, Carneiro IG, Aust B (2012) Bullying at work and onset of a major depressive episode among Danish female eldercare workers. *Scand J Work Environ Health*, 218–227.
- Sakellariou N (2018) A historical perspective on the engineering ideologies of sustainability: the case of SLCA. *Int J Life Cycle Assess* 23:445–455. <https://doi.org/10.1007/s11367-016-1167-9>
- Soboll LAP (2008) Assédio moral no Brasil: a ampliação conceitual e suas repercussões. Em L. A. P. Soboll (Org.). *Violência psicológica e assédio moral no trabalho: pesquisas brasileiras* (pp. 23–65). São Paulo: Casa do Psicólogo.
- Spearman C (1937) *Psychology down the Ages*, vol 1. Macmillan, London
- Thomson W (1894) *Popular lectures and addresses* (Vol. 2). Cambridge University Press.
- Timmerman ME, Lorenzo-Seva U (2011) Dimensionality assessment of ordered polytomous items with parallel analysis. *Psychol Methods* 16(2):209–220. <https://doi.org/10.1037/a0023353>

- Tokede O, Traverso M (2020) Implementing the guidelines for social life cycle assessment: past, present, and future. *Int J Life Cycle Assess* 25(10):1910–1929
- Traverso M, Finkbeiner M, Jørgensen A, Schneider L (2012) Life cycle sustainability dashboard. *J Ind Ecol* 16(5):680–688
- Ugaya CML, Zortea RB, Haberland NT, Ramirez PKS, Souza AM (2015) Proposals for improvements in social life cycle assessment. VI International Conference on Life Cycle Assessment, Cilca, Lima. https://www.researchgate.net/publication/281407758_Proposals_for_improvements_in_a_social_Life_Cycle_Assessment
- United Nations Environment Programme (UNEP) (2020) Guidelines for social life cycle assessment of products and organizations. Benoît Norris C, Traverso M, Neugebauer S, Ekener E, Schaubroeck T, Russo Garrido S, Berger M, Valdivia S, Lehmann A, Finkbeiner M, Arcese G (eds)
- United Nations Environment Programme (UNEP) (2021) Methodological sheets for subcategories in social life cycle assessment (S-LCA).
- United Nations Environment Programme and Society for Environmental Toxicology and Chemistry (UNEP/SETAC) (2009) Guidelines for social life cycle assessment of products. UNEP-SETAC Life-Cycle Initiative.
- Waldekker B, Molnar S (2014) Social LCA at SP: challenges and opportunities. SP Technical Research Institute of Sweden Report. <https://www.diva-portal.org/smash/get/diva2:962861/FULLTEXT01.pdf>
- Wang SW, Hsu CW, Hu AH (2016) An analytic framework for social life cycle impact assessment—part 1: methodology. *Int J Life Cycle Assess* 21:1514–1528
- World Health Organization (2021) Healthy and safe telework: technical brief-Geneva 2021.
- Wu R, Yang D, Chen J (2014) Social Life Cycle Assessment Revisited. *Sustainability* 6(7):4200–4226. <https://doi.org/10.3390/su6074200>
- Xu T, Magnusson Hanson LL, Lange T, Starkopf L, Westerlund H, Madsen IE, Rod NH (2019) Workplace bullying and workplace violence as risk factors for cardiovascular disease: a multi-cohort study. *Eur Heart J* 40(14):1124–1134
- Zamagni A, Amerighi O, Buttol P (2011) Strengths or bias in social LCA? *Int J Life Cycle Assess* 16:596–598

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Arij Mohamad Radwan Omar Chabrawi¹  · Josemberg Moura de Andrade² · Cássia Maria Lie Ugaya³ · Marzia Traverso¹

✉ Arij Mohamad Radwan Omar Chabrawi
arij.chabrawi@inab.rwth-aachen.de

Josemberg Moura de Andrade
josemberg.andrade@gmail.com

Cássia Maria Lie Ugaya
cassiaugaya@utfpr.edu.br

Marzia Traverso
marzia.traverso@inab.rwth-aachen.de

¹ Institute of Sustainability in Civil Engineering (INaB), RWTH Aachen University, Mies-Van-Der-Rohe-Straße 1/ Gebäude 2130, 52074 Aachen, Germany

² Universidade de Brasília, Instituto de Psicologia, Departamento de Psicologia Social, Do Trabalho E das Organizações (PSTO), Brasília, Brazil

³ Universidade Técnica Federal Do Paraná (UTFPR), Curitiba, Brazil