

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

Design and validation of the Intervention and Multiphase Training Transfer Assessment Questionnaire (IMTTAQ) for individual, group, and leader occupational health psychology training interventions

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

Vasquez, C.A., Nielsen, K., Chiesa, R., Muschalla, B., Etzelmüller, A., Peláez Zuberbuhler, J., et al. (2026). Design and validation of the Intervention and Multiphase Training Transfer Assessment Questionnaire (IMTTAQ) for individual, group, and leader occupational health psychology training interventions. JOURNAL OF OCCUPATIONAL HEALTH PSYCHOLOGY, 31(1), 16-41 [10.1037/ocp0000420].

Availability:

This version is available at: <https://hdl.handle.net/11585/1050581> since: 2026-02-26

Published:

DOI: <http://doi.org/10.1037/ocp0000420>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

Abstract

Training transfer represents a crucial mechanism for understanding how occupational health psychology (OHP) training interventions achieve their effects, yet no validated instruments exist to systematically assess the implementation process surrounding transfer within wellbeing-focused contexts. This study develops and validates the Intervention and Multiphase Training Transfer Assessment Questionnaire (IMTTAQ), designed specifically to capture contextual elements, transfer mechanisms, and behavioral outcomes across the implementation phases of OHP training interventions. We tested the IMTTAQ's psychometric properties, convergent and discriminant validity, and predictive validity using data from a large multinational, multi-wave study across ten organizations (N = 1,966 at T1) in five European countries, with data collected over six time points. Results demonstrated: (a) strong psychometric properties, including robust factor structure and temporal reliability; (b) convergent and discriminant validity with expected relationships between subscales; and (c) predictive validity showing that pre-training contextual constructs and training experiences predicted post-training integration and wellbeing outcomes. The IMTTAQ provides a validated tool for systematically assessing training transfer across temporal phases (pre-training, during training, and post-training integration), enabling researchers and practitioners to evaluate the contextual constructs and mechanisms that facilitate successful workplace learning integration in OHP interventions.

Keywords: workplace training interventions; IGLO framework, questionnaire validation; wellbeing; process evaluation

Design and Validation of the Intervention and Multiphase Training Transfer Assessment Questionnaire (IMTTAQ) for Individual, Group, and Leader OHP Training Interventions

Occupational health psychology (OHP) interventions are defined as planned, theory-based actions to improve employee health and wellbeing (Fox et al., 2022). Such interventions range from individual stress management and mindfulness programs to organizational job redesign and leadership training initiatives (Van Agteren et al. 2021). Workplaces are critical settings for addressing mental health challenges that affect hundreds of millions of workers globally and cost trillions in lost productivity (WHO, 2022; Chisholm et al., 2016), as they provide direct access to adult populations and opportunities to address work-related stressors.

Despite their importance, recent meta-analyses and reviews demonstrate inconsistent effectiveness of OHP interventions across organizational contexts (Burgess et al., 2020; Dreison et al., 2018; Estevez Cores et al., 2021; Fox et al., 2022; Richardson & Rothstein, 2008; Van Agteren et al. 2021). This variability stems partly from evaluation approaches that focus primarily on pre-post changes in wellbeing outcomes while overlooking the processes by which intervention content is taken up, generalized, and maintained in practice (Hughes et al., 2020; Joyce et al., 2016; Nielsen & Shepherd, 2022). Training transfer, defined as the application, generalization, and maintenance of learned knowledge and skills (Baldwin & Ford, 1988), offers a promising lens for understanding variation in intervention outcomes.

However, existing transfer frameworks and instruments were originally designed for technical or performance-oriented training and therefore emphasize discrete, observable skills measured at single time points (Blume et al., 2019). These tools focus primarily on performance outcomes and largely neglect the psychological and interpersonal competencies central to OHP – such as stress management, emotional regulation, and collaborative behaviors – that develop gradually and are highly dependent on social and organizational

contextual variables (Bates et al., 2012; Blume et al., 2010; Day & Nielsen, 2017; Grossman & Salas, 2011). Moreover, current instruments tend to capture transfer only post-training, providing limited guidance on what to measure across different intervention phases or how contextual factors shape implementation processes at different time points (Blume et al., 2019; Cheng & Hampson, 2008; Nielsen & Shepherd, 2022). Evidence from sustainability studies highlights the practical consequences of these limitations: the effects of OHP interventions often diminish within months (Estevez Cores et al., 2021; Karabinski et al., 2021), reflecting insufficient attention to transfer mechanisms and contextual factors across the intervention lifecycle (Nielsen et al., 2023). As a result, researchers lack tools to understand why similar interventions succeed in some organizational contexts but not in others, and practitioners are left without diagnostic guidance to identify and address barriers to implementation success.

The present study addresses this gap by introducing and validating the Intervention and Multiphase Training Transfer Assessment Questionnaire (IMTTAQ), a novel instrument designed specifically for OHP interventions. The IMTTAQ systematically captures contextual elements, transfer mechanisms, and behavioral outcomes across four intervention phases. It integrates insights from three complementary research areas: established training transfer literature (Blume et al., 2010; Burke & Hutchins, 2007), realist evaluation approaches (Nielsen & Miraglia, 2017; Pawson & Tilley, 1997), and recent frameworks for OHP training intervention evaluation (Nielsen et al., 2023; Nielsen & Shepherd, 2022). We evaluated its psychometric properties using data from a large, multi-wave, multinational intervention study across five European countries (Czech Republic, Germany, Italy, the Netherlands, and Spain).

Our study makes three key contributions to OHP research and practice. First, we demonstrate the practical utility of the IMTTAQ for intervention evaluation. By identifying

specific contextual barriers and facilitators that influence implementation success, the IMTTAQ enables researchers and practitioners to design more effective interventions, adapt strategies to organizational realities, and diagnose why interventions succeed or fail beyond outcome measurement. Second, we develop and validate the IMTTAQ as a theoretically grounded instrument that addresses identified gaps in measuring training transfer for wellbeing-focused interventions. Unlike traditional instruments, the IMTTAQ captures contextual aspects, training processes, and integration mechanisms into a single validated questionnaire, the IMTTAQ enables researchers to systematically evaluate how intervention content is taken up and sustained in the workplace. Third, we provide empirical evidence on the temporal dynamics of transfer by showing how contextual factors, training experiences, and integration processes influence well-being outcomes differently across intervention phases. Using multi-wave data, we move beyond the predominant reliance on single time-point assessments in training transfer research (Blume et al., 2019) and reveal that the relative importance and influence of transfer mechanisms shift across intervention phases, shaping work engagement and burnout in distinct ways. This provides a concrete toolkit for improving the effectiveness of OHP interventions in the workplace.

Theoretical Framework

The IMTTAQ is theoretically grounded in the integration of three complementary approaches that together provide a comprehensive framework for understanding transfer in wellbeing-focused interventions: general training transfer research, OHP-specific models of intervention evaluation and transfer, and realist evaluation. The training transfer literature (Blume et al., 2010, 2019) shows that the transference of new skills depends on motivation and supportive environments, suggesting similar patterns for wellbeing interventions, and conceptualizes transfer as a cyclical, multiphase process. Building on these insights, the Integrated Training Transfer and Effectiveness Model (ITTEM; Nielsen & Shepherd, 2022)

links intervention characteristics to transfer outcomes across different phases of OHP interventions, while the Integrated Process Evaluation Framework (IPEF; Nielsen et al., 2023) extends the perspective by incorporating constructs measured at the individual, group, and leader levels. Grounded in realist evaluation principles (Nielsen & Miraglia, 2017; Pawson & Tilley, 1997), the IMTTAQ acknowledges that while some transfer elements are broadly applicable, others operate through context-dependent mechanisms specifically relevant to wellbeing outcomes. The following sections detail how these theoretical approaches inform the development and validation of the IMTTAQ within OHP contexts.

Fundamentals of Training Transfer and its Application to OHP Training

Baldwin and Ford's (1988) transfer process model remains among the most influential frameworks in organizational psychology (Brown & Sitzmann, 2011). It defines transfer as the application, generalization, and maintenance of learned knowledge and skills in the workplace, conceptualized through training inputs, transfer processes, and outputs. Multiple factors influence transfer effectiveness, including trainee characteristics (e.g., motivation, cognitive ability) and contextual aspects like the workplace environment and training design (Baldwin & Ford, 1988; Burke & Hutchins, 2007; Grossman & Salas, 2011). Blume et al.'s (2010) meta-analysis further demonstrated that motivation and supportive environments predict transfer more strongly for open, adaptive skills such as leadership development than for closed, technical skills. Because wellbeing interventions similarly target complex, open-ended competencies, they may follow distinct transfer patterns. Their Dynamic Transfer Model conceptualizes transfer as a cyclical process of attempts, feedback, and reapplication, supporting multiphase assessment approaches. Although validated tools exist (Holton et al., 2000; Bates et al., 2012), most were developed for performance-oriented contexts and may not suit OHP interventions, which require consideration of implementation processes and contextual factors related to wellbeing improvement. This limitation is addressed below.

The Difference Between OHP and Performance-Oriented Training Interventions

Despite growing interest in workplace wellbeing, transfer in OHP contexts remains underexplored (Estevez Cores et al., 2021; Karabinski et al., 2021; Nielsen et al., 2023). This gap is noteworthy because existing transfer models were largely developed for performance-oriented training and may not account for the distinctive aims and mechanisms of OHP interventions. Unlike performance-oriented programs, OHP interventions primarily aim to improve employee health and wellbeing (Day & Nielsen, 2017; WHO, 2022). They often involve complex psychological and behavioral changes that may rely on different transfer mechanisms, challenging assumptions of universality in traditional models. Empirical evidence supports this distinction: emotional factors predict outcomes in health training but not in technical skills training (Vignoli et al., 2018), and stress management transfer depends more on organizational climate than technical training does (Sørensen & Holman, 2014). Yet most evaluations of wellbeing interventions still focus on outcomes rather than transfer processes (Dreison et al., 2006; Joyce et al., 2016; Miguel et al., 2023), limiting insight into how employees apply learning in daily work.

Existing transfer instruments reinforce this gap. Widely used tools, such as the Training Transfer Questionnaire (Grossman & Salas, 2011) and Training Transfer Inventory (Bates et al., 2012), focus primarily on performance outcomes and largely neglect psychological and interpersonal competencies central to OHP. They also tend to capture transfer only post-training, providing limited guidance on what to measure before, during, and after training (Blume et al., 2019; Cheng & Hampson, 2008). As a result, these tools are not well-suited to capturing how wellbeing competencies are implemented, embedded, and are sustained over time in organizational settings. Evidence from sustainability studies further highlights this challenge: the effects of OHP interventions often diminish within months (Estevez Cores et al., 2021; Karabinski et al., 2021), reflecting insufficient attention to

transfer mechanisms and contextual factors (Nielsen et al., 2023). The IMTTAQ addresses this limitation by providing a structured, multiphase assessment that integrates general transfer principles with mechanisms unique to wellbeing-focused interventions.

Integrating Realist Evaluation with Training Transfer Models in OHP

Building on our discussion of existing transfer models' limitations, we now turn to realist evaluation as a framework to address these gaps. Realist evaluation (RE) examines "what works for whom in which circumstances" (Pawson & Tilley, 1997). Nielsen and Miraglia (2017) challenged the traditional view of randomized controlled trials (RCTs) as the gold standard in organizational intervention research (Burgess et al., 2020), noting that they often fail to explain inconsistent results or account for complex intervention elements (Deaton & Cartwright, 2018). Unlike RCTs, which aim to control variables, RE embraces real-world complexity through Context-Mechanism-Outcome (CMO) configurations. 'Context' refers to work environmental aspects influencing implementation; for example, participatory interventions require a foundation of trust between workers and management. According to Nielsen and Miraglia (2017), 'mechanisms' can be about interventions' content (relating to the intervention's core elements) and implementation (concerning delivery processes). Exploring how these mechanisms interact with specific contexts offers deeper insight into how and why interventions achieve intended outcomes.

Building on realist evaluation, Nielsen and Shepherd (2022) developed the ITTEM, a model that integrates training transfer and effectiveness frameworks to evaluate wellbeing intervention transfer. The ITTEM conceptualizes transfer across temporal phases, identifying that successful transfer requires trainees to actively translate newly acquired skills into behaviors at work rather than passively receiving information. Unlike traditional performance training, where success is measured by improved task execution and discrete performance metrics, OHP interventions often require transfer that manifests as subtle changes in complex,

multidimensional wellbeing outcomes. As noted, these outcomes typically emerge through indirect pathways, with transfer success dependent on social and psychological factors less relevant in performance-oriented training. Additionally, the voluntary nature of many OHP interventions fosters intrinsic motivation rather than compliance-driven engagement, requiring evaluation approaches centered on self-appraisal and reflective sense-making instead of objective performance metrics (Nielsen et al., 2025). Crucially, the ITTEM recognizes that in OHP contexts, transfer is a dynamic, cyclical process where wellbeing improvements themselves may recursively influence future transfer attempts, a dimension largely absent in traditional transfer models. Despite its theoretical contributions in bridging transfer and effectiveness literature for wellbeing interventions, the ITTEM presents practical challenges: it lacks precise alignment with intervention phases, emphasizes constructs over measurable variables, and introduces a framework that is simultaneously overly complex in structure yet potentially oversimplified in capturing transfer dynamics.

While the ITTEM provides a temporal framework for understanding transfer, OHP interventions face an additional challenge: diverse implementation across multiple organizational levels (Martin et al., 2016; Nielsen et al., 2023). Comprehensive interventions often employ multilevel approaches addressing individual employees, work groups, leaders, and organizational policies simultaneously (Boyd et al., 2023; WHO, 2022). Day and Nielsen (2017) conceptualized this through the IGLO (Individual, Group, Leader, and Organizational) model. Individual-level interventions develop employee abilities through mindfulness or job crafting programs. Group-level interventions improve team processes, while leader-level interventions enhance employee wellbeing through leadership training. Organizational-level interventions address work organization and design (Nielsen, 2022).

The IPEF (Nielsen et al., 2023) addresses this multilevel complexity by proposing systematic process evaluation across different intervention levels. Unlike traditional transfer

models focused primarily on individual aspects (Baldwin & Ford, 1988; Blume et al., 2010), IPEF recognizes that transfer mechanisms may operate differently depending on intervention level. For example, leader-level intervention transfer depends on both the leader's application of skills and their effect on team dynamics (Hammer et al., 2021; Tafvelin et al., 2019).

To address both the temporal limitations of the ITTEM and the multilevel considerations of the IPEF, we developed the IMTTAQ as a concrete measurement instrument that provides researchers with a validated tool to assess key transfer factors at specific timepoints throughout OHP training interventions. The IMTTAQ uniquely focuses on constructs at individual, group, and leader levels, excluding the organizational level. This reflects that these three levels typically involve training where transfer is key, whereas organizational-level interventions target structural changes and follow different implementation processes. Some mechanisms and contextual factors operate across all levels, while others depend on the intervention target. For example, subordinate support is critical in leadership training (Nielsen & Daniels, 2012), whereas colleague support may be more important for individual training (Blume et al., 2010). By capturing adaptations at the individual, group, and leader levels, the IMTTAQ goes beyond existing questionnaires (Bates et al., 2012; Holton et al., 2000), providing researchers with a comprehensive yet practical tool to evaluate training transfer in OHP interventions.

Integrating Theoretical Foundations: The IMTTAQ Measurement Approach

The following sections present the IMTTAQ constructs organized by intervention phase, covering their theoretical basis, empirical support, and application to occupational health psychology contexts, reflecting our temporal approach to measuring wellbeing intervention training transfer. Construct selection was systematically guided by the ITTEM (Nielsen & Shepherd, 2022), IPEF (Nielsen et al., 2023), and established training transfer literature (Baldwin & Ford, 1988; Blume et al., 2010), as detailed in the Methods section and

Table 2. Figure 1 provides an overview of the main constructs across intervention phases.

INSERT FIGURE 1 AROUND HERE

Constructs Measured by the IMTTAQ

Pre-OHP training interventions: Informed by training transfer (Blume et al., 2010) and organizational intervention frameworks (Nielsen et al., 2023; Nielsen & Shepherd, 2022), our research identifies key organizational conditions as critical pre-intervention elements influencing OHP training effectiveness. These conditions align with realist evaluation principles and represent elements across the IGLO framework (Individual, Group, Leader, Organization), shaping the context for successful transfer.

Drawing on ISO standards (2018) and training transfer literature (Burke & Hutchins, 2007; Jackson et al., 2018), we strategically adapted two dimensions from the Workplace Integrated Safety and Health (WISH) assessment (López Gómez et al., 2021): leadership commitment and dialogue processes (adapted from participation). These dimensions were selected based on evidence that they influence training transfer in psychosocial intervention contexts (Nielsen & Shepherd, 2022; Nielsen et al., 2023), allowing us to develop a focused, theory-driven tool for assessing pre-training context conditions in OHP interventions.

These two dimensions were operationalized to capture key contextual aspects influencing training transfer in OHP interventions. First, *leadership commitment* reflects employee perceptions of how organizational leaders at all levels prioritize, communicate, and allocate resources mental health initiatives. It includes explicit commitment to wellbeing, resource allocation, integration into organizational goals, and consistent demonstration through actions. As Jackson et al. (2018) noted, such commitment fosters a safe environment for applying new skills. Second, *dialogue processes* encompass collaborative mechanisms through which managers and employees jointly plan and evaluate wellbeing initiatives, including employee involvement in decision-making and psychologically safe environments

for voicing concerns. These dimensions operate across all IGLO levels by facilitating meaningful exchange between organizational stakeholders (WHO, 2022).

Beside leadership commitment and dialogue processes, a third important organizational factor is *communication* about training interventions, which refers to how clearly the intervention's relevance, goals, and participant expectations are articulated to potential trainees. Jackson et al. (2018) identified such communication as a significant predictor of training transfer, as it ensures participants understand the purpose and personal requirements of the intervention. Effective communication about interventions enhances employee awareness, enables cognitive appraisal of the change, and facilitates co-learning processes that support implementation (DeJoy et al., 2010; Nielsen et al., 2014; Nielsen & Randall, 2012), all of which improve implementation and outcomes (Roodbari et al., 2023).

Beyond these organizational factors, participants' orientation toward the intervention also significantly affects transfer outcomes. *Readiness for change*, classified by Nielsen et al. (2023) as a pre-intervention context factor in the IPEF framework, reflects participants' cognitive and emotional orientation toward the intervention. Following Randall et al. (2009) and Weiner (2020), we define this as confidence in applying the intervention, expectations of personal wellbeing benefits, and willingness to accept resulting changes. This construct is particularly relevant for wellbeing interventions, where personal benefit rather than professional advancement typically drives participation (Jackson et al., 2018). Our measurement captures readiness for change at the pre-training stage across all three intervention levels, consistent with realist evaluation's emphasis on pre-existing conditions that influence how intervention mechanisms operate (Pawson & Tilley, 1997).

During the OHP training interventions: In this phase, as training interventions are implemented (Nielsen et al., 2023), several key constructs influence training transfer, functioning as mechanisms within our realist evaluation framework. These during-training

variables represent critical components in the progression from pre-training conditions to post-training application. First, *training design* plays a key role in facilitating transfer (Baldwin & Ford, 1988) and functions as a critical training mechanism that brings about learning (Nielsen & Shepherd, 2022). We conceptualize this as the extent to which training activities, exercises, and examples help trainees apply learning to their work context, with emphasis on facilitators' understanding of the trainees' work environment. OHP training materials should mirror real-life situations (Saks & Belcourt, 2006), helping trainees connect training to daily work (Salas et al., 2012). For example, in a leadership training about managing employee time pressures, Stansfeld et al. (2015) found leaders appreciated accessible, understandable, and relevant materials. In addition, the wider literature confirms training design as a predictor of transfer success (Gegenfurtner et al., 2009).

Second, *acceptability of training content* serves as a key training mechanism in the learning process. We define this as participants' satisfaction with the training experience, including the extent to which training met their expectations, the perceived appropriateness of format and materials, and willingness to recommend the training to others (Coffeng et al., 2013). While previous OHP studies did not directly link acceptability to transfer (Carolan et al., 2017; Martin et al., 2020; Stansfeld et al., 2015), the training transfer literature has linked perceived utility to transfer success (Grossman & Salas, 2011); trainees who feel interventions address their needs and have a positive evaluation of the content are more likely to integrate learning into daily work (Burke & Hutchins, 2007). This factor operates across Individual, Group, and Leader-level interventions, though specific aspects of acceptability may vary depending on the intervention target and content.

Third, trainees' *perceptions of training facilitators* influence their motivation to change and retention of key messages (Andersen et al., 2022), representing another important training mechanism in our model. We define this construct as participants' assessment of how

well facilitators valued their participation, demonstrated understanding of their situation, and encouraged active engagement in training assignments. Effective facilitators engage trainees and demonstrate how learning applies to daily work practices (Burke & Hutchins, 2007). Research shows that facilitators who empower trainees and challenge learning mindsets enhance training transfer (Creon & Schermuly, 2022).

Finally, *intention to transfer* occupies a unique position in our framework. The ITTEM identifies it as a post-training individual mechanism crucial for training transfer (Nielsen & Shepherd, 2022), while the IPEF acknowledges its role when discussing how implementation elements influence transfer attempts (Nielsen et al., 2023). We define it as participants' belief that what they learned will help them at work, their perception of having developed helpful skills, and their recognition of acquiring work-relevant knowledge. Intention to transfer represents the first stage in Foxon's (1993) five-stage transfer process and a critical component often missing from traditional transfer models (Cheng & Hampson, 2008). Empirical evidence confirms it as a significant predictor of actual transfer behavior (Gegenfurtner, 2013; Cheng et al., 2015), consistent with the theory of planned behavior (Ajzen, 2002). In our framework, intention to transfer serves as a factor associated with the other mechanisms described above regarding training and precursor to post-training transfer behaviors, illustrating the dynamic progression through temporal phases that distinguishes our approach from static transfer models.

OHP training interventions integration: While initial reactions and intentions are important, the true test of training effectiveness occurs in the weeks and months following completion (Blume et al., 2019; Nielsen & Shepherd, 2022). During this integration phase, trainees must incorporate learning into their daily working lives. The key mechanism in this phase is *training transfer*. As established in our theoretical framework, this process involves the application, generalization, and maintenance of knowledge and skills (Baldwin & Ford,

1988; Blume et al., 2010) specifically within OHP contexts where wellbeing rather than performance is the primary outcome. Operationally, transfer encompasses the extent to which trainees use the knowledge and skills gained in training in their everyday work, successfully apply training contents, and transfer learning back to their actual job (Nielsen et al., 2023; Nielsen & Shepherd, 2022). Within our framework, transfer represents the key mechanism through which training produces health and wellbeing outcomes.

Our framework also acknowledges that this mechanism will be more likely activated within a context that promotes transfer. *Opportunities to integrate* training into practice represent a crucial contextual factor that facilitates transfer. We define this as having the necessary tasks, resources, and information to apply training content. Successful transfer requires trainees to practice newly acquired knowledge, skills, and attitudes and maintain these changes over time (Burke & Hutchins, 2007).

Beyond these practical opportunities, the social environment also shapes transfer outcomes. The training transfer literature emphasizes the importance of social support from various workplace stakeholders (Hughes et al., 2020; Nielsen & Daniels, 2012; Shann et al., 2019). While traditional training transfer literature typically assesses support through specific questions about job task changes (Burke & Hutchins, 2007; Schoeb et al., 2021), this approach may not suit OHP training interventions, which often target changes in cognitions and emotions rather than direct job tasks. To address these unique needs, we included differentiated support measures across intervention levels. For individual training, we incorporated measures of *peer support* and *supervisor support* capturing assistance with work difficulties, making work easier, and listening to problems (Frese et al., 1999). For leadership training, we included measures of *subordinate support*, assessing similar dimensions from the perspective of leaders receiving support from their team members. For group training, we developed measures of *group support* as a whole, capturing the group's collective

encouragement and appreciation for utilizing training (Holton et al., 2000).

In summary, the training integration phase represents the critical transition from training to workplace practice. Training transfer functions as the central mechanism, supported by contextual aspects including opportunities to integrate and differentiated social support (peer, supervisor, subordinate, and group). By measuring these constructs across Individual, Group, and Leader levels, the IMTTAQ captures how OHP training translates into improved workplace practices and ultimately enhances wellbeing outcomes.

OHP training outcomes: After trainees have had the opportunity to transfer learning, the training's intended outcomes should be evaluated. Von Thiele Schwarz et al. (2017) introduced the concept of *capacity building* as a key outcome of organizational interventions, demonstrating how interventions can foster enhanced capability across multiple organizational levels. More recently, Nielsen et al. (2023) emphasized the importance of capturing whether individuals, groups, leaders, and organizations have developed their abilities to manage mental health issues, representing a form of learning that extends beyond individual skill acquisition to build sustainable organizational capability. We adopt *capacity building* as a proximal outcome because, in OHP contexts, transfer operates as a dynamic and cyclical process through which wellbeing improves indirectly, shaped by psychological and social factors that revolve around self-appraisal and activation.

Drawing on the IGLO framework (Nielsen et al., 2023), we operationalize *capacity building* as participants' retrospective assessments of what they have gained from the intervention across individual, team, leadership, and organizational levels. This retrospective approach is complementary to, rather than a replacement for, traditional pre-post measurements of capacity development. While pre-post designs capture objective changes in specific capabilities, retrospective assessments offer distinct advantages: they capture participants' integrated sense of growth across the intervention period and their confidence in

newly developed capabilities, which are themselves predictive of sustained application and future engagement with wellbeing initiatives. Von Thiele Schwarz et al. (2017) demonstrated that different levels can interact synergistically to improve well-being. Accordingly, capacity building is assessed using a formative measurement model (Fleuren et al., 2018) where distinct learning domains combine to form participants' overall evaluation of intervention gains rather than reflecting a single underlying construct. This approach aligns with training transfer literature emphasizing participants' perceptions of learning acquisition (Baldwin & Ford, 1988), and with Nielsen et al.'s (2025) argument that capturing participants' perceptions of intervention success alongside pre-post outcomes is essential, as organizational members' sensemaking of change influences both future engagement with initiatives and whether intervention effects are sustained over time.

While capacity building represents the final construct within the IMTTAQ framework, a comprehensive evaluation of OHP training interventions should also include wellbeing outcomes. These outcomes, while outside the IMTTAQ, represent the ultimate targets of OHP training. OHP interventions aim not only to prevent illness and distress, but also to promote quality of working life (Fox et al., 2022). In our validation study, we focused on two widely established indicators of work-specific wellbeing: work engagement and burnout. Work engagement represents a positive, fulfilling work-related state characterized by vigor, dedication, and absorption (Schaufeli et al., 2002). Burnout, in contrast, captures a state of emotional exhaustion, cynicism, and reduced professional efficacy (Maslach et al., 2001). Importantly, wellbeing is a complex, multidimensional construct that can be approached through various conceptual and measurement frameworks (Iasiello et al., 2024; Van Agteren et al., 2021). Researchers implementing the IMTTAQ may select wellbeing indicators that align with their intervention targets, participant characteristics, or theoretical approach. For example, leadership-focused training might assess leader wellbeing outcomes

(Kaluza et al., 2020), team-oriented training might measure team climate, and individual interventions might evaluate quality of life or job satisfaction (Iasiello et al., 2024). This flexibility allows the IMTTAQ to accommodate diverse wellbeing metrics while maintaining its core focus on training transfer processes.

The present study

This paper introduces the IMTTAQ, which assesses key contextual elements and process mechanisms of OHP training interventions across multiple timepoints. While not included in the IMTTAQ, broader working conditions provide important validation context. We validate the questionnaire through three analyses: (1) examining structural validity (Hypothesis 1) by testing whether proposed constructs form distinct factors aligned with temporal phases and demonstrate measurement invariance across time and languages; (2) assessing convergent and discriminant validity (Hypothesis 2) through relationships with theoretically related constructs including working conditions; and (3) evaluating predictive validity (Hypothesis 3) by testing whether pre-training constructs and during-training perceptions positively relate to training transfer and whether transfer positively associates with wellbeing outcomes. We elaborate on each hypothesis below.

For Hypothesis 1, we examine the factorial validity of the IMTTAQ's measurement structure across four distinct temporal phases of the OHP training process. Rather than conducting a single factor analysis with all constructs, we analyze the factor structure within each temporal phase separately, as this approach: (1) aligns with our theoretical framework emphasizing distinct processes at different intervention stages, (2) provides more meaningful information about how constructs function within each phase, and (3) accommodates our multi-wave data collection strategy. Specifically, we propose that the pre-training phase constructs will form four distinct factors: leadership commitment, dialogue processes, communication, and readiness for change. The during training phase constructs will comprise

four distinct factors: acceptability of content, training design, perceptions of facilitators, and intention to transfer. The training integration phase constructs, being more complex, will form three core factors (training transfer, opportunities to integrate, and manager support) and three level-specific factors (peer support at the individual level, group support at the group level, and subordinate support at the leader level). The training outcomes phase will include one factor: capacity building. As part of Hypothesis 1, we examine temporal measurement invariance of training integration scales across three timepoints to ensure observed changes reflect true differences rather than measurement artifacts. We also test cross-linguistic invariance to validate the IMTTAQ's reliability across different language groups.

In Hypothesis 2, we examine the convergent and discriminant validity of our IMTTAQ scales. For convergent validity, we expect pre-training scales (leadership commitment, dialogue processes, communication, and readiness for change) to correlate moderately with conceptually related working conditions. Specifically, pre-training constructs should correlate positively with autonomy and team support (as both represent supportive organizational elements), and negatively with workload, job insecurity, and work-home conflict (as these represent organizational stressors that typically undermine supportive environments). These specific working conditions were selected based on established job resources-demands models (Bakker & Demerouti, 2017) that identify them as key aspects influencing employee wellbeing. For discriminant validity, we expect IMTTAQ scales to demonstrate distinctiveness in two ways. First, correlations between IMTTAQ scales and working conditions should be moderate rather than strong, indicating related but distinct constructs. Second, the pattern of intercorrelations among IMTTAQ scales should reflect their theoretical temporal relationship: scales within the same temporal phase (e.g., all pre-training constructs such as leadership commitment and dialogue processes) should correlate more strongly with each other than with scales from different phases (e.g., during training

constructs like training design and facilitator effectiveness, or post-training constructs like training transfer). This pattern would support that our measurement model captures distinct constructs rather than a single underlying dimension, while still acknowledging their theoretical relationships within the training transfer process.

In Hypothesis 3, we examine predictive validity. We hypothesize that pre-training constructs (leadership commitment, dialogue processes, communication, and readiness for change) and during training variables (training design, acceptability of content, perception of facilitators, and intention to transfer) will be positively related to training transfer (Hypothesis 3a). This hypothesis is grounded in our integrated framework (Baldwin and Ford's, 1988; Nielsen & Shepherd, 2022; Nielsen et al., 2023), which identify these pre-training constructs and training delivery mechanisms as critical antecedents to successful application of training content. Particularly important is intention to transfer, which Foxon (1993) positioned as the first stage in the transfer process.

Additionally, we hypothesize that training transfer will be positively related to capacity building, work engagement, and burnout (Hypothesis 3b). The IPEF framework (Nielsen et al., 2023) proposes capacity building as an important outcome related to successful interventions, while improved wellbeing represents the ultimate goal that OHP training interventions aim to address. These hypothesized relationships allow us to examine key temporal associations in our data, considering training transfer in relation to other variables at the time of measurement.

Methods

Open Science Practices

The study was funded by the European Union's Horizon 2020 research and innovation program (reference withheld for anonymity). Data included in our analyses are available on Zenodo, an open-access repository indexed in the EU OpenAire (reference

withheld for anonymity). Access to the dataset can be requested. The study received initial ethical approval by the lead university's Ethics Committee (reference withheld for anonymity). In addition, all partners involved in data collection activities received ethical approval from their respective Ethics Committees.

Sample and procedure

The study employed a multi-wave design, with data collected through an online survey across six time points. The first wave (T1) captured pre-training constructs at baseline (month 0), before the interventions began. Before completing the baseline survey, participants received written information about the intervention's aims, procedures, and their right to withdraw at any time. The baseline questionnaire was distributed to all employees in participating workplaces. Subsequent waves (T2-T5) focused exclusively on individuals participating in the training. The T2 questionnaire measured perceptions during training and was distributed immediately after participants completed their final training session (e.g., after the sixth session in a six-session training program). T3-T5 questionnaires focused on training integration and were distributed at three, six, and nine months after the T1 baseline. Due to staggered training implementation, these questionnaires were only sent to participants who had completed their training, resulting in varying completion rates across the three time points. The final wave (T6) measured training outcomes, including capacity building, work engagement, and burnout. The T6 questionnaire was distributed to all employees within the participating organizations, as we reached out again to everyone who received the T1 survey. Figure 1 provides an overview of measures at each time point.

Participants

The sample consisted of 1966 employees (at T1) in ten organizations, five of which were private and five public, located in Germany, Czech Republic, Italy, Spain, and the Netherlands. An overview of response rates and demographics is provided in Table 1.

INSERT TABLE 1 AROUND HERE

OHP training interventions

A multilevel intervention approach was implemented, with interventions identified from literature and the research team's experience (reference withheld for anonymity). Each organization underwent a comprehensive needs analysis, comprising focus groups with employees and their representatives, semi-structured interviews with middle and senior managers, and senior management completion of relevant Workplace Integrated Safety and Health Assessment (WISH) dimensions (Sorensen et al., 2018). The research team analyzed these data and prepared reports for stakeholder meetings where employee representatives, management, and Human Resources collaboratively prioritized interventions. Organizations were encouraged to select interventions targeting at least two IGL levels.

After the interventions were selected, the T1 assessment was distributed to all employees in participating companies. All selected interventions were training-based, with previous evidence of improving work engagement and reducing burnout; hence, our focus on these outcomes in testing predictive validity. Research team members served as facilitators across countries, following a standardized protocol. The number of interventions implemented varied across organizations, ranging from one to five, with a median of two interventions per organization. This approach of implementing multiple complementary interventions is supported by Bernadino and Curado's (2020) research, who found that exposure to multiple training interventions enhances learning outcomes. In total, 32 interventions were implemented across the ten organizations: 13 at the individual level, 9 at the group level, and 10 at the leader level. While participation was generally voluntary, group-level interventions required the involvement of all members within a respective work group. For a detailed overview of implemented interventions, see FigShare URL (A provisional version is provided in the appendix section for review purposes).

Development and adaptation of IMTTAQ measures

We followed Hinkin's (1998) guidelines in developing our scales. Building on the theoretical constructs identified and justified in the framework section, the identification of measures drew from multiple sources: the IPEF (Nielsen et al., 2023), the ITTEM (Nielsen & Shepherd, 2022), existing training transfer measures from organizational learning literature (including validated questionnaires by Bates et al., 2012; Holton et al., 2000), and constructs identified in meta-analyses (Blume et al., 2010) and systematic reviews (Baldwin & Ford, 1988; Burke & Hutchins, 2007; Cheng & Hampson, 2008; Ford et al., 2018; Grossman & Salas, 2011; Jackson et al., 2019; Schoeb et al., 2021).

A key challenge in measure selection was that many existing training transfer scales focus on job performance, career progression, and specific job tasks (e.g., Bates et al., 2000; Holton et al., 2000). These foci were not well aligned with OHP training interventions, which target changes in emotions, cognitions, and generic behaviors (e.g., mindfulness training). As detailed in our theoretical framework, this distinction between performance-oriented and wellbeing-oriented transfer highlights the need for specialized measurement approaches. We selected measures based on their face validity for OHP training interventions, their alignment with the theoretical mechanisms and contextual elements identified earlier, and their established psychometric properties. As existing measures from organizational learning literature did not adequately capture the temporal phases (pre-training, during training, integration) or the intervention levels (Individual, Group, Leader), we both adapted existing measures and developed new ones.

The measure identification process involved three steps. First, we identified potential measures to operationalize the constructs in our theoretical framework. Second, the research team evaluated whether these measures could adequately capture the constructs within occupational health training interventions. Third, we finalized the selection for each temporal

phase, and where appropriate, reformulated items to enable assessment at the Individual, Group, and Leader levels.

Table 2 presents the IMTTAQ constructs, response scales, references, level-specific assessment details, and each measure's role in validity testing, indicating whether measures were adapted, newly developed, or retained in their original form. Several measures were adapted for OHP training interventions (e.g., leadership commitment, dialogue processes, communication, content acceptability, training design, and facilitator perception), and others were self-developed (e.g., intention to transfer, opportunities to use). The table also notes which measures have prior psychometric validation. All adaptations focused on aligning scales with mental health and wellbeing outcomes, resulting in the first comprehensive measurement tool validated for OHP training implementation. For detailed items and adaptations, see FigShare (URL).

INSERT TABLE 2 AROUND HERE

Data analysis

We employed a three-step analytic strategy to test our hypotheses. First, to test Hypothesis 1, we examined the factor structure of our measures using Exploratory Factor Analysis (EFA) followed by Confirmatory Factor Analysis (CFA). Separate analyses were conducted for constructs within each temporal phase, consistent with our theoretical framework. EFAs were performed using principal axis factoring with Direct Oblimin rotation in SPSS Statistics 29. We determined the number of factors using multiple criteria (Costello & Osborne, 2019; Watkins, 2018). Items were retained with primary factor loadings $\geq .40$ and cross-loadings $< .30$. Subsequently, we performed CFAs using Mplus version 8 (Muthén & Muthén, 1998-2017) to confirm the EFA structures. We assessed model fit using four indices: Confirmatory Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residual (SRMR) (Hu &

Bentler, 1999; Kline, 2010). Model comparisons used the likelihood ratio test (difference in χ^2 between nested models).

Measurement invariance. We tested language (Italian, Spanish, Czech, Dutch, and German) and temporal (T3-T5) measurement invariance of our factor models. These tests assess whether the factor structure remains equivalent across languages and time points, indicating construct stability (Little et al., 2007). Following established procedures (Brown, 2006; Cheung & Rensvold, 2002), we tested a sequence of increasingly constrained nested models: configural invariance (same factor structure across groups), metric invariance (equal factor loadings), and scalar invariance (equal intercepts). Model comparisons were evaluated using Chi-Squared Difference Test and CFI comparisons (Cheung & Lau, 2012). At a minimum, partial metric invariance was required to validate cross-group comparisons.

Second, to test Hypothesis 2, we assessed convergent and discriminant validity through correlation analyses. For convergent validity, we examined bivariate correlations between pre-training constructs (T1) and established working conditions (e.g., workload, autonomy), interpreting coefficients of .10, .30, and .50 as small, medium, and large effects (Cohen et al., 2003). For discriminant validity, we expected correlations below .70, as higher values may indicate construct redundancy rather than related but distinct constructs (Rönkkö & Cho, 2022). We anticipated correlations in the .30-.50 range, given that our focal constructs (e.g., leadership commitment, readiness for change) are theoretically related to but conceptually distinct from working conditions (e.g., workload, autonomy, team support). Finally, we examined the pattern of intercorrelations among IMTTAQ scales across temporal phases, expecting stronger correlations within the same phase (e.g., all pre-training constructs) compared to those across phases (e.g., pre- and during-training constructs).

Third, to test Hypothesis 3 regarding predictive validity, we tested predictive validity using hierarchical multiple regression analyses. For Hypothesis 3a, we tested whether pre-

training constructs (measured at T1) and during training variables (measured at T2) were positively related to training transfer (measured at T3-T5). For Hypothesis 3b, we tested the associations between training transfer (measured at T3-T5) and outcome variables (capacity building, work engagement, and burnout), all measured at the same time points. For all regression analyses, we examined unstandardized beta coefficients and reported adjusted R^2 values to indicate the proportion of variance explained. Due to anonymity constraints in our questionnaire design, we were unable to aggregate some measures at the group level and conduct multilevel analyses, a limitation addressed in the discussion section.

Results

Stability of the IMTTAQ Factor Structure

Hypothesis 1 proposed that the IMTTAQ would demonstrate distinct measurement structures across four temporal training phases.

Exploratory and Confirmatory factor analyses of the IMTTAQ.

We present our factor analyses following the temporal sequence of the training process phases (pre-training, during training, integration, and outcomes). We conducted separate analyses for constructs within each temporal phase rather than one omnibus factor analysis. This approach allowed for more precise evaluation of construct distinctiveness within conceptually related timepoints while accommodating our multi-wave design.

Time 1 – Pre-training. This phase comprised four dimensions: leadership commitment, dialogue processes, communication, and readiness for change. EFA extracted three factors, explaining 69.70% of the variance, with items from leadership commitment and dialogue processes loading onto a single factor (Table 3). CFA, however, confirmed that a four-factor model aligned with our theoretical framework ($\chi^2(71) = 554.870$, CFI = .969, TLI = .961, RMSEA = .067, SRMR = .028) provided a significantly better fit than both the three-factor EFA-based solution and a one-factor solution (see Appendix A, Table A1 for model

comparisons); standardized factor loadings from the CFA are presented in Table 3. These results indicate potential overlap between leadership commitment and dialogue processes, which is also reflected in their strong correlations detailed in the Convergent and Discriminant Validity section below. We return to this issue in the Discussion section to consider its theoretical and practical implications.

INSERT TABLE 3 AROUND HERE

Time 2 – During training. This phase included training design, acceptability of content, perception of facilitators, and intention to transfer. EFA initially extracted two factors (58.64% variance explained), with items from training design, acceptability of content, and intention to transfer loading onto one factor, while perception of facilitator items loaded separately (Table 4). When a four-factor solution was forced, all items loaded onto their theoretically expected factors. CFA confirmed that the four-factor model ($\chi^2(71) = 204.320$, CFI = .974, TLI = .967, RMSEA = .057, SRMR = .035) fitted significantly better than both alternative models (See Appendix A, Table A2; CFA factor loadings in Table 4), supporting the theoretical distinctiveness of these constructs despite their conceptual relatedness.

INSERT TABLE 4 AROUND HERE

Time 3-5 - Training integration. To account for the varying sources of information and levels of measurement, we conducted four separate analyses. This approach was necessary because while some integration variables are standard across all Individual, Group, and Leader (IGL) levels, others are level-specific.

Solution 1 examined three measures common across all levels (training transfer, opportunities to integrate, and management support). EFA extracted two factors explaining 64.27% of variance, with training transfer and opportunities to integrate loading onto a single factor (Table 5). When a three-factor solution was forced, all items loaded onto their expected

factors. CFA confirmed that the three-factor model ($\chi^2(24) = 125.714$, CFI = .987, TLI = .980, RMSEA = .057, SRMR = .025) fitted significantly better than alternatives.

Solution 2 added peer support (individual level) to the common measures (training transfer, opportunities to integrate, and management support). EFA extracted three factors explaining 64.15% of variance, with manager support and peer support loading onto a single factor, suggesting conceptual similarity in social support measures (Table 5). CFA confirmed a four-factor structure ($\chi^2(48) = 247.762$, CFI = .976, TLI = .967, RMSEA = .057, SRMR = .040) as superior to alternatives.

Solution 3 incorporated group support with the common measures. EFA extracted three factors explaining 68.29% of variance, with training transfer and group support loading onto one factor, and opportunities to integrate and manager support loading onto another (Table 5). However, CFA confirmed a four-factor structure ($\chi^2(48) = 217.765$, CFI = .982, TLI = .975, RMSEA = .052, SRMR = .028) as theoretically and empirically distinct.

Solution 4 examined subordinate support (leader level) with the common measures. EFA extracted three factors explaining 64.36% of variance, with manager support and subordinate support loading together (Table 5). However, CFA confirmed the theoretically proposed four-factor structure ($\chi^2(48) = 173.972$, CFI = .984, TLI = .978, RMSEA = .045, SRMR = .045) with all constructs distinctly measured.

Across all solutions, the theoretically proposed models demonstrated superior fit compared to alternatives (see Appendix A, Table A3). CFA factor loadings for the combined Time 3-5 training integration are reported in Table 6, and factor loadings for each time point separately are provided in Appendix A, Table A4. These findings support the construct validity of both core and level-specific factors in the integration phase.

INSERT TABLES 5 AND 6 AROUND HERE

Time 6 - Training outcomes. The capability building measure comprised a single

factor explaining 56.19% of variance with good factor loadings ($\geq .61$). CFA confirmed this one-factor model with good fit: $\chi^2(2) = 19.558$, CFI = .984, TLI = .953, RMSEA = .106, SRMR = .021 (Table 7).

INSERT TABLE 7 AROUND HERE

Measurement Invariance.

Measurement invariance across languages

We conducted measurement invariance tests for each temporal phase's factor structure across language groups with sufficient sample sizes (see Table 8)¹. Results demonstrated excellent goodness-of-fit for the configural model with no statistically significant differences compared to the metric model for most constructs (Table 8). While full metric invariance was not achieved for all constructs, partial metric invariance was established by freeing constraints on specific items: at T1, items 2-4 of leadership commitment and item 4 of readiness for change; at T3-5, items 2-3 of opportunities to integrate; and at T6, item 4 of capability building. In practical terms, this means that these specific factor loadings were allowed to differ across language versions while all other loadings remained constrained to equality.

Scalar invariance was not achieved for any model, which is a common pattern in cross-cultural research (Funder & Gardiner, 2024; Van De Schoot et al., 2015). These results indicate that while the basic factor structure and most factor loadings remain consistent across languages, some factor loadings and all item intercepts vary. Given that full metric invariance was not achieved, we did not test for partial scalar invariance, following the hierarchical nature of measurement invariance testing (Vandenberg & Lance, 2000). These findings support using the IMTTAQ to examine relationships between variables within each

¹ In our invariance analysis, we excluded groups with small sample sizes (e.g., German, see Table 11) to ensure robust results. This decision was based on established guidelines for minimum sample size requirements in CFA related to model complexity (Koran, 2020; MacCallum et al., 1999). While there are no universally agreed-upon cutoff, larger samples are generally recommended for complex models to achieve stable parameter estimates.

language group. However, researchers should avoid direct comparisons of raw scores or latent means across language groups, or interpret such comparisons with caution.

Temporal invariance for training integration (Time 3-5)

We tested temporal invariance for a model including the three common subscales measured across Times 3-5: training transfer, opportunities to integrate, and manager support. The results demonstrated excellent goodness-of-fit across all levels of invariance testing (Table 9). The configural model showed strong fit ($\chi^2 = 196.943$, $df = 72$, $CFI = .984$, $TLI = .976$, $RMSEA = .063$, $SRMR = .032$). The metric model maintained excellent fit with no statistically significant degradation compared to the configural model ($\Delta\chi^2(12) = 6.589$, $p = .884$, $\Delta CFI = .001$). Similarly, the scalar model demonstrated strong fit with no significant difference from the metric model ($\Delta\chi^2(12) = 3.842$, $p = .986$, $\Delta CFI = .001$). Table 9 displays the factor loadings for each time point separately.

These results provide robust evidence for strong temporal invariance, indicating that the factor structure, factor loadings, and item intercepts remain consistent across the three time points. This consistency allows for meaningful comparisons of latent means across time points and supports the reliability of any observed changes in these constructs over time. This finding is particularly valuable in training evaluation contexts, as it suggests that observed changes in these constructs represent true changes rather than measurement artifacts.

INSERT TABLES 8 AND 9 AROUND HERE

Internal consistency of the IMTTAQ

Table 10 presents Cronbach's alpha coefficients for the 15 IMTTAQ subscales measured across different time points. The scales demonstrated excellent internal consistency, with alpha values exceeding .75 for the entire sample. At the country level, results were generally consistent, with a few exceptions of lower reliability in smaller subsamples (e.g., perception of facilitators in the German subsample, $\alpha = .47$, $N = 21$).

These results, combined with our factor analyses and invariance testing, provide substantial evidence supporting the psychometric properties of the IMTTAQ scales (Hypothesis 1). The proposed factor structure demonstrated significantly better fit than alternative models across all temporal phases, with items showing strong loadings on their intended factors. While measurement invariance analyses showed metric invariance across languages and strong temporal invariance for integration phase measures, scalar invariance across languages was not achieved. Each subscale demonstrated good internal consistency overall, with variations in country-specific reliability attributable to smaller sample sizes.

INSERT TABLE 10 AROUND HERE

Convergent and Discriminant validity

Hypothesis 2 proposed that IMTTAQ subscales would demonstrate both convergent and discriminant validity, with pre-training constructs correlating moderately with related working conditions and scales within the same temporal phase correlating more strongly with each other than with scales from different phases. Table 11 presents these correlations. Detailed results for each phase are presented below.

Time Point 1: Pre-training. Table 11 shows correlations between pre-training constructs and working conditions. Leadership commitment demonstrated significant positive correlations with conceptually similar constructs (e.g., dialogue processes, $r = .80, p < .001$), moderate correlations with less related constructs (e.g., autonomy, $r = .38, p < .001$), and negative correlations with potentially adverse aspects (e.g., work-home conflict, $r = -.32, p < .001$). Similar patterns were observed for other pre-training subscales, with correlation strengths varying according to theoretical proximity.

Time Points 2-5: During training and integration. As shown in Table 11, our discriminant validity expectations were supported with variables measured at the same time point showing stronger correlations (e.g., training design and acceptability of content at T2, r

= .70, $p < .001$) than those measured at different time points. Cross-temporal correlations followed theoretically expected patterns, with earlier constructs such as intention to transfer (T2) showing significant positive relationships with subsequent integration phase constructs, including opportunities to integrate (T3, $r = .53$, $p < .001$), training transfer (T4, $r = .53$, $p < .001$), and manager support (T5, $r = .16$, $p = .03$). These patterns further support the temporal structure of our measurement model while confirming the interconnected nature of the training transfer process across phases.

These results demonstrate that relationships between IMTTAQ subscales and external constructs are largely theoretically consistent, partially supporting both convergent and discriminant validity (Hypothesis 2). Most correlations between IMTTAQ scales and working conditions were below .70, supporting discriminant validity. An exception was the strong correlation between leadership commitment and dialogue processes ($r = .80$, $p < .001$), which exceeded this threshold. As expected, scales within the same temporal phase correlated more strongly with each other than with those from different phases, and cross-temporal correlations followed theoretically consistent patterns. Implications for construct distinctiveness are addressed in the Discussion.

INSERT TABLE 11 AROUND HERE

Predictive validity

Hypothesis 3 examined relationships between pre-training constructs, during training variables, subsequent training transfer (H3a), and between training transfer and outcomes (H3b), following our proposed temporal structure. For H3a, Table 12 shows regression results linking pre-training context constructs to training transfer. Leadership commitment showed a significant positive association with training transfer at T3 ($\beta = .37$, $p = .006$). Readiness for change demonstrated consistent associations across all time points, with significant relationships to training transfer at T3 ($\beta = .36$, $p < .001$), T4 ($\beta = .22$, $p = .02$),

and T5 ($\beta = .32, p = .006$). Table 12 presents regression results between during training variables and subsequent training transfer. Intention to transfer was significantly associated with training transfer at both T3 ($\beta = .42, p < .001$) and T4 ($\beta = .29, p < .001$). Training design showed positive relationships with training transfer at T4 ($\beta = .20, p = .042$) and T5 ($\beta = .29, p < .025$). These findings indicate that both pre-training context constructs and during training perceptions relate to subsequent application of learning.

For H3b, Table 13 presents regression results examining relationships between training transfer and outcomes. Training transfer was positively associated with work engagement at T6 across all measurement points (T3, $\beta = .47, p = .002$; T4, $\beta = .36, p = .009$; T5, $\beta = .22, p = .048$) and negatively associated with burnout at T6 (T3, $\beta = -.11, p = .24$; T4, $\beta = -.14, p = .069$; T5, $\beta = -.15, p = .014$). However, training transfer did not show significant associations with capacity building at T6 (T3, $\beta = -.11, p = .28$; T4, $\beta = .05, p = .53$; T5, $\beta = -.03, p = .61$), suggesting differential relationships with various outcomes. These results support H3a, demonstrating significant relationships between pre-training context aspects, during training variables, and subsequent training transfer. H3b was partially supported, with significant associations between training transfer and wellbeing outcomes (work engagement and burnout) but not capacity building.

INSERT TABLES 12 AND 13 AROUND HERE

Discussion

In this study, we developed the IMTTAQ by synthesizing Blume et al.'s (2010) meta-analytic findings with the ITTEM (Nielsen & Shepherd, 2022) and IPEF (Nielsen et al., 2023) models. Rather than developing entirely new constructs, the IMTTAQ strategically integrates and adapts established measures within a comprehensive tool specifically tailored to OHP contexts. Grounded in realist evaluation principles (Nielsen & Miraglia, 2017; Pawson & Tilley, 1997), it addresses the critical gap between training and outcomes in health

and wellbeing training interventions. This instrument captures both the temporal dynamics of the transfer process and a multilevel assessment across Individual, Group, and Leader levels. Beyond instrument validation, our findings provide novel empirical insights into contextual elements and mechanisms connecting training perceptions to transfer behaviors and ultimately to wellbeing outcomes. Our analysis across diverse settings demonstrates not only the IMTTAQ's psychometric properties but also its capacity to advance understanding of how OHP training interventions achieve their effects.

We found support for Hypothesis 1, confirming the IMTTAQ's psychometric validity across temporal phases. Although multilevel modelling was not feasible with our current dataset, phase-specific factor analyses consistently supported our theoretically derived structure. CFAs confirmed the IMTTAQ's factor structure across intervention phases, including pre-training (T1), during training (T2), integration (T3–T5), and post-training (T6), supporting the distinctiveness of constructs within each phase. These findings demonstrate that the IMTTAQ captures the multidimensional nature of training transfer over time while accommodating the practical realities of multi-wave data collection in work settings.

We found mixed support for Hypothesis 2 regarding convergent and discriminant validity. Pre-training (T1) context constructs showed expected relationships with workload, autonomy, and team support. At T2 and T3–T5, stronger correlations emerged between measures collected at the same time points, consistent with theoretical expectations. However, some construct pairs demonstrated notably high correlations that warrant discussion. Leadership commitment and dialogue processes were strongly correlated, likely reflecting their shared foundation in organizational support for wellbeing from the WISH assessment (López Gómez et al., 2021). These constructs function as distinct but interdependent: leadership enables dialogue processes while productive dialogue reinforces leadership commitment (Nielsen & Randall, 2012). Despite statistical overlap, their

theoretical separation remains valuable for understanding complementary pathways of intervention effectiveness (Nielsen & Noblet, 2018; Sorensen et al., 2018). Intention to transfer, training design, and acceptability of content also demonstrated substantial overlap, which align with Nielsen and Shepherd's (2022) conceptualization of these constructs as related training mechanisms that collectively influence learning. Similarly, the strong correlation between integration opportunities and training transfer echoes findings from Baldwin and Ford (1988) and Burke and Hutchins (2007) that transfer opportunity is intrinsically linked to actual transfer behavior.

These high correlations, which align with our EFA results, reflect the interconnected nature of training transfer processes. Following Messick's (1995) view of validation as ongoing rather than binary, these findings refine our understanding of construct boundaries while acknowledging their conceptual proximity. While the Fornell-Larcker criterion raises some concerns about discriminant validity for these highly correlated constructs (Cheung et al., 2024), CFA results demonstrating superior fit for our differentiated model support their treatment as distinct yet related constructs. Researchers using the IMTTAQ should consider these overlaps when interpreting results, particularly when examining relationships between constructs known to share substantial variance.

Hypothesis 3 examined predictive validity and partially supported the theoretical associations proposed in established models. Our multi-wave design captured nuanced temporal patterns in pre-training constructs. Leadership commitment (T1) was positively associated with training transfer at T3, while readiness for change (T1) consistently related to training transfer across time points (T3-T5). During training variables also showed significant temporal patterns: intention to transfer (T2) positively correlated with subsequent training transfer at T3-T4, and training design (T2) showed significant associations with transfer at T4-T5. These temporal variations expand upon previous cross-sectional research (Schoeb et

al., 2021) and existing questionnaires (Bates et al., 2012; Holton et al., 2000) that did not examine time-separated relationships. Importantly, training transfer was associated with subsequent wellbeing outcomes at T6. These results suggest a connection between knowledge application and wellbeing, extending the literature focused on performance outcomes (Cheng & Hampson, 2008). Overall, our findings align with the dynamic training transfer model (Blume et al., 2019) and highlight how different training constructs show varying relationships with transfer and outcomes across different post-training periods.

Contrary to our expectations, training transfer was not significantly associated with capacity building. This raises important theoretical questions about intervention pathways. First, in line with Nielsen et al.'s (2025) sensemaking perspective, interventions may have emphasized engagement and burnout over mental health management capabilities. Participants likely applied their learning to address immediate concerns rather than developing the broader perceived gains captured by our capacity building measure, which encompasses individual abilities, team dialogue, leadership capability, and organizational climate for wellbeing. This suggests that in organizational interventions, immediate application of specific skills often precedes the development of more comprehensive, multilevel capacities. Second, drawing on Nielsen and Miraglia's (2017) framework, our study focused on process mechanisms across intervention types but did not capture content-specific mechanisms (e.g., mindfulness techniques or leadership skills). These content-specific pathways may be crucial for linking general training transfer with the distinct domains that form capacity building. Without these mechanisms, participants may successfully transfer general training skills without developing the integrated sense of capability across individual, team, leadership, and organizational levels that characterizes capacity building. These findings help explain why achieving sustainable intervention effects in OHP contexts remains challenging (Karabinski et al., 2021; Richardson & Rothstein,

2008), suggesting that capacity building as a formative outcome requires intentional cultivation of gains across multiple complementary domains, with intervention design explicitly targeting each level of the IGLO framework.

Implications for Research and Practice

The validated IMTTAQ offers significant implications for both research and practice. By capturing the temporal dynamics of training transfer across OHP intervention phases, it addresses a crucial gap in evaluation approaches. For researchers, the IMTTAQ enables three key advancements. First, it provides a validated tool to examine how contextual factors and training mechanisms interact across time, translating realist evaluation principles (Nielsen & Shepherd, 2022) and established training transfer models (Cheng & Hampson, 2008) into empirically testable frameworks. Second, our time-separated analyses linking training perceptions, transfer behaviors, and wellbeing outcomes offer preliminary evidence for intervention pathways and clarify how transfer processes relate to occupational health outcomes such as work engagement and burnout. These features, including the temporal design and integration of contextual mechanisms and wellbeing indicators, reflect core realist evaluation principles (Nielsen & Miraglia, 2017; Pawson & Tilley, 1997). The IMTTAQ thus provides a systematic tool to explore what works, for whom, and under which circumstances, while further research is needed to refine the theoretical models underpinning it.

For practitioners, the IMTTAQ offers practical value through phase-specific assessments. While the complete instrument may be too extensive for routine use, practitioners can strategically implement specific components. Pre-training assessment of readiness and communication can identify barriers, allowing for enhanced information provision before training begins (Hughes et al., 2018). During training, data on training characteristics enables providers to optimize design elements and facilitator approaches most strongly related to subsequent transfer. At integration, measuring transfer and relevant

support mechanisms helps develop targeted activities appropriate to each intervention level. Importantly, our model should be seen as inspiration rather than prescription; evaluators should tailor their approach to the specific training delivered. For instance, mindfulness training may require less peer support for skill application than team-based interventions.

Strengths and limitations

Our study's methodological approach offers notable strengths. Our multi-wave data collection across 10 organizations in five countries allowed us to test the IMTTAQ across diverse OHP intervention contexts and timepoints. Additionally, our rigorous development process grounded in established training transfer literature and OHP-specific frameworks has resulted in an instrument with demonstrated psychometric properties.

Nevertheless, several important limitations must be acknowledged. First, although the multi-wave design could have supported tests of mediation and moderation, the present study prioritized validation of the IMTTAQ's measurement properties over examination of specific Context-Mechanism-Outcome configurations. While correlation and regression analyses for Hypotheses 2 and 3 offered preliminary insights into associations among variables, we did not systematically assess whether contextual factors moderated relationships between training components or whether mechanisms mediated outcomes. Such analyses, which would address "what works for whom and in which circumstances", represent an important avenue for future research building on the validated measure.

Second, although the IPEF encompasses interventions at the individual, group, leader, and organizational levels (Nielsen et al., 2023), the IMTTAQ was intentionally limited to the first three. This scope reflects that interventions targeting individuals, groups, and leaders typically involve training, where transfer is a central mechanism for effectiveness. In contrast, organizational-level interventions often rely on participatory approaches involving collective change processes and modifications to work practices, which require different

measurement constructs. These interventions focus on action plan implementation (Nielsen & Abildgaard, 2013; Nielsen & Randall, 2013; Randall et al., 2009) rather than training transfer (Nielsen et al., 2023). Future research could examine how organisational-level actions interact with training interventions at the individual, group, and leader levels, potentially generating synergistic effects. Such integration would align with evidence that interventions addressing work environment changes are particularly effective (Rugulies et al., 2023), especially when combined with training components across multiple organisational levels.

Third, our validation yielded mixed evidence for convergent, discriminant, and predictive validity. While pre-training measures related to working conditions as expected, some constructs (e.g., leadership commitment and dialogue process at T1; intention to transfer and training design at T2) were highly correlated, raising concerns about discriminant validity yet also reflecting the interconnected nature of training transfer processes. The data collection timeline further constrained causal inference: concurrent measurement of during-training variables limited the examination of sequential relationships, and the simultaneous assessment of outcomes such as capacity building and wellbeing hindered the establishment of temporal precedence and differentiation between proximal and distal effects. Although the time-separated design represents an advance over cross-sectional approaches, causal direction cannot be firmly determined, particularly for post-intervention outcomes where reverse effects remain plausible. To reduce participant burden, working condition measures were omitted at T2-T5, leaving the influence of workplace context in later phases unexplored. These challenges illustrate the inherent complexities of capturing multiple intervention processes effectively (Nielsen & Shepherd, 2022). Future research should consider staggered measurement designs to better trace temporal sequences in OHP interventions, incorporate contextual measures throughout the intervention period, and conduct phase-specific investigations before integrating them into comprehensive models.

Fourth, sample size limitations affected our measurement invariance testing across language groups. To ensure robust results, we excluded smaller samples (e.g., German) in line with CFA guidelines (Koran, 2020; MacCallum et al., 1999). While partial metric invariance was established for most constructs, limited sample sizes in some language groups may have hindered achieving full metric and scalar invariance. This common challenge in cross-cultural research (Funder & Gardiner, 2024) creates uncertainty about whether measurement differences stem from sample limitations or genuine cultural variations. Future research should focus on larger, balanced samples across language groups to better assess measurement equivalence or conduct targeted validation studies within specific language contexts to determine whether observed differences reflect methodological artifacts or substantive cross-cultural variations in training transfer mechanisms.

Finally, while merging data enabled adequate sample sizes for factor analyses and measurement invariance, this approach restricts insights into level-specific patterns. Anonymity requirements prevented the specification of some variables at their intended levels (peer, manager, and subordinate support), precluding multilevel factor analysis and between-group relationship testing. Consequently, these variables were omitted from several validity tests, and findings for group and manager support at the individual level should be interpreted with caution (Mathieu et al., 2020). Although the IMTTAQ includes items with different referents (individual, group, and leader behaviors), we did not conduct formal multilevel validation analyses to test equivalence across these referent levels. Future research should examine whether the instrument functions equivalently when assessing individual-focused, team-focused, and leadership-focused interventions, and should employ designs that allow proper nesting of responses to enable more robust multilevel validation.

Conclusion

The development and validation of the IMTTAQ represents a significant progress in

OHP training evaluation. This comprehensive tool integrates realist evaluation principles with established transfer models for theoretically-grounded assessment across multiple phases and levels. Our multinational European validation study not only demonstrated the instrument's reliability, but also provided empirical insights into how pre-training constructs, during training, and integration processes are related to wellbeing. The IMTTAQ bridges critical gaps by enabling systematic analysis of contextual factors and mechanisms related to training effectiveness. For researchers, it provides a framework to investigate “what works for whom in which circumstances”, while offering practitioners an evidence-based toolkit to design effective interventions, identify barriers, monitor delivery, and evaluate application. This contribution advances the field's capacity to implement OHP training interventions that create sustainable improvements in worker wellbeing across diverse organizational contexts.

References

- Andersen, M.F., Nielsen, K., & Ajslev, J.Z.N. (2022). The Relational Fit in Organizational Interventions – What can Organizational Research Learn from Research in Psychotherapy? *International Journal of Environmental Research and Public Health*, 18(15), 8104. <https://doi.org/10.3390/ijerph18158104>
- Ajzen, I. (2002). Perceived behavioral control, self-efficacy, locus of control, and the theory of planned behavior. *Journal of Applied Social Psychology*, 32, 1–20. <https://doi.org/10.1111/j.1559-1816.2002.tb00236.x>
- Bakker, A. B., & Demerouti, E. (2017). Job demands–resources theory: Taking stock and looking forward. *Journal of occupational health psychology*, 22(3), 273. DOI: 10.1037/ocp0000056
- Baldwin, T. T., & Ford, J. K. (1988). Transfer of training: A review and directions for future research. *Personnel Psychology*, 41(1), 63–105. <https://doi.org/10.1111/j.1744-6570.1988.tb00632.x>
- Bates, R., Holton III, E. F., & Hatala, J. P. (2012). A revised learning transfer system inventory: factorial replication and validation. *Human Resource Development International*, 15(5), 549-569. <https://doi.org/10.1080/13678868.2012.726872>
- Bernardino, G., & Curado, C. (2020). Training evaluation: a configurational analysis of success and failure of trainers and trainees. *European Journal of Training and Development*, 44(4/5), 531-546. DOI 10.1108/EJTD-10-2019-0177
- Blume, B. D., Ford, J. K., Baldwin, T. T., & Huang, J. L. (2010). Transfer of training: A meta-analytic review. *Journal of Management*, 36(4), 1065-1105. <https://doi.org/10.1177/0149206309352880>
- Blume, B. D., Ford, J. K., Surface, E. A., & Olenick, J. (2019). A dynamic model of training transfer. *Human Resource Management Review*, 29(2), 270-283.

<https://doi.org/10.1016/j.hrmr.2017.11.004>

- Boyd, R. C., Castro, F. G., Finigan-Carr, N., Okamoto, S. K., Barlow, A., Kim, B. K. E., ... & Kenney, A. (2023). Strategic directions in preventive intervention research to advance health equity. *Prevention Science*, 24(4), 577-596. <https://doi.org/10.1007/s11121-022-01462-5>
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. NY, London: The Guilford Press.
- Brown, K. G., & Sitzmann, T. (2011). Training and employee development for improved performance. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology, Vol. 2. Selecting and developing members for the organization* (pp. 469–503). American Psychological Association. <https://doi.org/10.1037/12170-016>
- Burgess, M. G., Brough, P., Biggs, A., & Hawkes, A. J. (2020). Why interventions fail: A systematic review of occupational health psychology interventions. *International Journal of Stress Management*, 27(2), 195–207. <https://doi.org/10.1037/str0000144>
- Burke, L. A., & Hutchins, H. M. (2007). Training transfer: An integrative literature review. *Human Resource Development Review*, 6(3), 263-296. <https://doi.org/10.1177/1534484307303035>
- Carolan, S., Harris, P. R., & Cavanagh, K. (2017). Improving employee well-being and effectiveness: systematic review and meta-analysis of web-based psychological interventions delivered in the workplace. *Journal of medical Internet research*, 19(7), e271. DOI: 10.2196/jmir.7583
- Cheng, E. W., & Hampson, I. (2008). Transfer of training: A review and new insights. *International Journal of Management Reviews*, 10(4), 327-341. doi: 10.1111/j.1468-2370.2007.00230.x
- Cheung, G. W., Cooper-Thomas, H. D., Lau, R. S., & Wang, L. C. (2024). Reporting

- reliability, convergent and discriminant validity with structural equation modeling: A review and best-practice recommendations. *Asia Pacific Journal of Management*, 41(2), 745-783. <https://doi.org/10.1007/s10490-023-09871-y>
- Cheng, E. W., Sanders, K., & Hampson, I. (2015). An intention-based model of transfer of training. *Management Research Review*, 38(8), 908-928. <https://doi.org/10.1108/MRR-05-2014-0107>
- Cheung, G. W., & Lau, R. S. (2012). A direct comparison approach for testing measurement invariance. *Organizational Research Methods*, 15(2), 167-198. <https://doi.org/10.1177/1094428111421987>
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural equation modeling*, 9(2), 233-255. https://doi.org/10.1207/S15328007SEM0902_5
- Chisholm, D., Sweeny, K., Sheehan, P., Rasmussen, B., Smit, F., Cuijpers, P., & Saxena, S. (2016). Scaling-up treatment of depression and anxiety: a global return on investment analysis. *The Lancet Psychiatry*, 3(5), 415-424. DOI: 10.1016/S2215-0366(16)30024-4
- Coffeng, J. K., Hendriksen, I. J., van Mechelen, W., & Boot, C. R. (2013). Process evaluation of a worksite social and physical environmental intervention. *Journal of Occupational and Environmental Medicine*, 55(12), 1409-1420. DOI: 10.1097/JOM.0b013e3182a50053
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Routledge. <https://doi.org/10.4324/9780203774441>
- Costello, A.B. & Osborne, J. (2019). Best practices in exploratory factor analysis: four recommendations for getting the most from your analysis. *Practical Assessment*,

- Research, and Evaluation*, 10(7). <https://doi.org/10.7275/jyj1-4868>
- Creon, L. E., & Schermuly, C. C. (2022). A new path to training transfer: Transformational trainers and psychologically empowered training participants. *International Journal of Training and Development*, 26(2), 228-244. <https://doi.org/10.1111/ijtd.12256>
- Day, A., & Nielsen, K. (2017). What does our organization do to help our well-being? Creating healthy workplaces and workers. In N. Chmiel, F. Fraccaroli, & M. Sverke (Eds.), *An introduction to work and organizational psychology* (pp. 295-314). Wiley Blackwell. <https://doi.org/10.1002/9781119168058.ch16>
- Deaton, A., & Cartwright, N. (2018). Understanding and misunderstanding randomized controlled trials. *Social Science & Medicine*, 210, 2-21. <https://doi.org/10.1016/j.socscimed.2017.12.005>
- DeJoy, D. M., Wilson, M. G., Vandenberg, R. J., McGrath-Higgins, A. L., & Griffin-Blake, C. S. (2010). Assessing the impact of healthy work organization intervention. *Journal of Occupational and Organizational Psychology*, 83(1), 139-165. <https://doi.org/10.1348/096317908X398773>
- Dreison, K. C., Luther, L., Bonfils, K. A., Sliter, M. T., McGrew, J. H., & Salyers, M. P. (2018). Job burnout in mental health providers: A meta-analysis of 35 years of intervention research. *Journal of Occupational Health Psychology*, 23(1), 18–30. <https://doi.org/10.1037/ocp0000047>
- Edwards, J. A., Webster, S., Van Laar, D., & Easton, S. (2008). Psychometric analysis of the UK Health and Safety Executive's Management Standards work-related stress Indicator Tool. *Work & Stress*, 22(2), 96-107. doi.org/10.1080/02678370802166599
- Eriksson, M., & Boman, E. (2018). Short Is Beautiful: Dimensionality and Measurement Invariance in Two Length of the Basic Psychological Need Satisfaction at Work Scale. *Frontiers in Psychology*, 9, 965. <https://doi.org/10.3389/fpsyg.2018.00965>

- Estevez Cores, S., Sayed, A. A., Tracy, D. K., & Kempton, M. J. (2021). Individual-focused occupational health interventions: A meta-analysis of randomized controlled trials. *Journal of Occupational Health Psychology*, 26(3), 189-203.
DOI: 10.1037/ocp0000249
- Fleuren, B. P., van Amelsvoort, L. G., Zijlstra, F. R., de Grip, A., & Kant, I. (2018). Handling the reflective-formative measurement conundrum: a practical illustration based on sustainable employability. *Journal of clinical epidemiology*, 103, 71-81.
<https://doi.org/10.1016/j.jclinepi.2018.07.007>
- Ford, J. K., Baldwin, T. T., & Prasad, J. (2018). Transfer of training: The known and the unknown. *Annual review of organizational psychology and organizational behavior*, 5(1), 201-225. <https://doi.org/10.1146/annurev-orgpsych-032117-104443>
- Fox, K.E., Johnson, S.T., Berkman, L.F., Sianoja, M., Soh, Y., Kubzansky, L.D. & Kelly, E.L. (2022). Organisational- and group-level workplace interventions and their effect on multiple domains of worker well-being: A systematic review. *Work & Stress*, 36:1, 30-59, <https://doi.org/10.1080/02678373.2021.1969476>
- Foxon, M. (1993). A process approach to the transfer of training. *Australasian Journal of Educational Technology*, 9(2). <https://doi.org/10.14742/ajet.2104>
- Frese, M., Teng, E., & Wijnen, C. J. (1999). Helping to improve suggestion systems: Predictors of making suggestions in companies. *Journal of Organizational Behavior*, 20(7), 1139-1155. [https://doi.org/10.1002/\(SICI\)1099-1379\(199912\)20:7<1139::AID-JOB946>3.0.CO;2-I](https://doi.org/10.1002/(SICI)1099-1379(199912)20:7<1139::AID-JOB946>3.0.CO;2-I)
- Funder, D. C., & Gardiner, G. (2024). Misgivings about measurement invariance. *European Journal of personality*, 0(0). <https://doi.org/10.1177/08902070241228338>
- Gegenfurtner, A. (2013). Dimensions of motivation to transfer: A longitudinal analysis of their influence on retention, transfer, and attitude change. *Vocations and*

- Learning*, 6(2), 187-205. <https://doi.org/10.1007/s12186-012-9084-y>
- Gegenfurtner, A., Festner, D., Gallenberger, W., Lehtinen, E., & Gruber, H. (2009). Predicting autonomous and controlled motivation to transfer training. *International Journal of Training and Development*, 13(2), 124-138. <https://doi.org/10.1111/j.1468-2419.2009.00322.x>
- Grohmann, A., & Kauffeld, S. (2013). Evaluating training programs: development and correlates of the Questionnaire for Professional Training Evaluation. *International Journal of Training and Development*, 17(2), 135-155. <https://doi.org/10.1111/ijtd.12005>
- Grossman, R., & Salas, E. (2011). The transfer of training: what really matters. *International journal of training and development*, 15(2), 103-120. <https://doi.org/10.1111/j.1468-2419.2011.00373.x>
- Hadžibajramović, E., Schaufeli, W., & De Witte, H. (2022). Shortening of the Burnout Assessment Tool (BAT)—from 23 to 12 items using content and Rasch analysis. *BMC Public Health*, 22(1), 560. <https://doi.org/10.1186/s12889-022-12946-y>
- Hammer, L. B., Brady, J. M., Brossoit, R. M., Mohr, C. D., Bodner, T. E., Crain, T. L., & Brockwood, K. J. (2021). Effects of a *Total Worker Health*® leadership intervention on employee well-being and functional impairment. *Journal of Occupational Health Psychology*, 26(6), 582–598. <https://doi.org/10.1037/ocp0000312>
- Hamzah, H. A., Marcinko, A. J., Stephens, B., & Weick, M. (2025). Making soft skills ‘stick’: a systematic scoping review and integrated training transfer framework grounded in behavioural science. *European Journal of Work and Organizational Psychology*, 34(2), 237-250. <https://doi.org/10.1080/1359432X.2024.2376909>
- Hinkin, T.R. (1998). A brief tutorial on the development of measures for use in survey

questionnaires. *Organizational Research Methods*, 1(1): 104-121.

<https://doi.org/10.1177/109442819800100106>

Holton III, E. F., Bates, R. A., & Ruona, W. E. (2000). Development of a generalized learning transfer system inventory. *Human Resource Development Quarterly*, 11(4), 333-360. [https://doi.org/10.1002/1532-1096\(200024\)11:4<333::AID-HRDQ2>3.0.CO;2-P](https://doi.org/10.1002/1532-1096(200024)11:4<333::AID-HRDQ2>3.0.CO;2-P)

Hu, L., & Bentler, P.M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55. <https://doi.org/10.1080/10705519909540118>

Hughes, A. M., Zajac, S., Woods, A. L., & Salas, E. (2020). The role of work environment in training sustainment: A meta-analysis. *Human Factors*, 62(1), 166-183. <https://doi.org/10.1177/0018720819845988>

Iasiello, M., Ali, K., van Agteren, J., Kolovos, E., Kyrios, M., Kashdan, T. B., & Fassnacht, D. B. (2024). What's the difference between measures of wellbeing, quality of life, resilience, and coping? An umbrella review and concept map of 155 measures of positive mental health. *International Journal of Wellbeing*, 14(2), 1-25. <https://doi.org/10.5502/ijw.v14i2.3621>

Innstrand, S. T., Melbye Langballe, E., Falkum, E., Espnes, G. A., & Aasland, O. G. (2009). Gender-specific perceptions of four dimensions of the work/family interaction. *Journal of Career Assessment*, 17(4), 402-416. <https://doi.org/10.1177/1069072709334238>

ISO (International Organization for Standardization). (2018). *BS 45002–1:2018. Occupational health and safety management systems. General guidelines for the application of ISO 45001. Guidance on managing occupational health*. International Organization for Standardization.

- Jackson, C. B., Brabson, L. A., Quetsch, L. B., & Herschell, A. D. (2019). Training transfer: A systematic review of the impact of inner setting factors. *Advances in Health Sciences Education*, 24, 167-183. <https://doi.org/10.1007/s10459-018-9837-y>
- Joyce, S., Modini, M., Christensen, H., Mykletun, A., Bryant, R., Mitchell, P. B., & Harvey, S. B. (2016). Workplace interventions for common mental disorders: a systematic meta-review. *Psychological medicine*, 46(4), 683-697. <https://doi.org/10.1017/S0033291715002408>
- Kaluza, A. J., Boer, D., Buengeler, C., & van Dick, R. (2020). Leadership behaviour and leader self-reported well-being: A review, integration and meta-analytic examination. *Work & Stress*, 34(1), 34-56. <https://doi.org/10.1080/02678373.2019.1617369>
- Karabinski, T., Haun, V. C., Nübold, A., Wendsche, J., & Wegge, J. (2021). Interventions for improving psychological detachment from work: A meta-analysis. *Journal of Occupational Health Psychology*, 26(3), 224-242. [10.1037/ocp0000280](https://doi.org/10.1037/ocp0000280)
- Kline, R.B. (2010). *Principles and Practice of Structural Equation Modeling*. NY, London: The Guilford Press.
- Koran, J. (2020). Indicators per factor in confirmatory factor analysis: More is not always better. *Structural Equation Modeling: A Multidisciplinary Journal*, 27(5), 765-772. <https://doi.org/10.1080/10705511.2019.1706527>
- Little, T. D., Preacher, K. J., Selig, J. P., & Card, N. A. (2007). New developments in latent variable panel analyses of longitudinal data. *International Journal of Behavioral Development*, 31, 357–365. <https://doi.org/10.1177/0165025407077757>
- López Gómez, M. A. L., Gundersen, D. A., Boden, L. I., Sorensen, G., Katz, J. N., Collins, J. E., ... & Williams, J. A. (2021). Validation of the Workplace Integrated Safety and

- Health (WISH) assessment in a sample of nursing homes using Item Response Theory (IRT) methods. *BMJ open*, 11(6), e045656. <https://doi.org/10.1136/bmjopen-2020-045656>
- MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4(1), 84-99. <https://doi.org/10.1037/1082-989X.4.1.84>
- Mathieu, J. E., Luciano, M. M., D’Innocenzo, L., Klock, E. A., & LePine, J. A. (2020). The development and construct validity of a team processes survey measure. *Organizational Research Methods*, 23(3), 399-431. <https://doi.org/10.1177/1094428119840801>
- Martin, A., Karanika-Murray, M., Biron, C., & Sanderson, K. (2016). The psychosocial work environment, employee mental health and organizational interventions: Improving research and practice by taking a multilevel approach. *Stress and health*, 32(3), 201-215. <https://doi.org/10.1002/smi.2593>
- Martin, A., Kilpatrick, M., Scott, J., Cocker, F., Dawkins, S., Brough, P., & Sanderson, K. (2020). Protecting the Mental Health of Small-to-Medium Enterprise Owners: A Randomized Control Trial Evaluating a Self-Administered Versus Telephone Supported Intervention. *Journal of Occupational and Environmental Medicine*, 62(7), 503-510. DOI:10.1097/JOM.0000000000001882
- Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology*, 52, 397-422. <https://doi.org/10.1146/annurev.psych.52.1.397>
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741-749. <https://doi.org/10.1037/0003-066X.50.9.741>

- Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus user's guide* (8th ed.). Muthén & Muthén.
- Miguel, C., Amarnath, A., Akhtar, A., Malik, A., Baranyi, G., Barbui, C., ... & Cuijpers, P. (2023). Universal, selective and indicated interventions for supporting mental health at the workplace: an umbrella review of meta-analyses. *Occupational and environmental medicine*, 80(4), 225-236. <https://doi.org/10.1136/oemed-2022-108698>
- Nielsen, K. (2022). Organizational interventions. In I. Johnsrude (Ed.), *Oxford research encyclopedia of psychology*. Oxford University Press.
<https://doi.org/10.1093/acrefore/9780190236557.013.109>
- Nielsen, K., & Abildgaard, J. S. (2013). Organizational interventions: A research-based framework for the evaluation of both process and effects. *Work & Stress*, 27(3), 278-297. <https://doi.org/10.1080/02678373.2013.812358>
- Nielsen, K., Abildgaard, J. S., & Daniels, K. (2014). Putting context into organizational intervention design: Using tailored questionnaires to measure initiatives for worker well-being. *Human Relations*, 67(12), 1537-1560.
<https://doi.org/10.1177/0018726714525974>
- Nielsen, K., & Daniels, K. (2012). Enhancing team leaders' well-being states and challenge experiences during organizational change: A randomized, controlled study. *Human Relations*, 65(9), 1207-1231. <https://doi.org/10.1177/0018726711433312>
- Nielsen, K., De Angelis, M., Innstrand, S. T., & Mazzetti, G. (2023). Quantitative process measures in interventions to improve employees' mental health: A systematic literature review and the IPEF framework. *Work & Stress*, 37(1), 1-26.
<https://doi.org/10.1080/02678373.2022.2080775>
- Nielsen, K., Ipsen, C., Abildgaard, J. S., & Edwards, K. (2025). What needs to improve and for whom when we implement participatory organisational interventions?. *Work &*

- Stress*, 39(1), 31-43. <https://doi.org/10.1080/02678373.2024.2435673>
- Nielsen, K., & Miraglia, M. (2017). What works for whom in which circumstances? On the need to move beyond the ‘what works?’ question in organizational intervention research. *Human Relations*, 70(1), 40-62. <https://doi.org/10.1177/0018726716670226>
- Nielsen, K., & Noblet, A. (2018). Chapter Introduction: Organizational interventions: Where we are, where we go from here?. In *Organizational Interventions for Health and Well-Being*. Taylor & Francis. DOI: 10.4324/9781315410494-1
- Nielsen, K., & Randall, R. (2012). The importance of employee participation and perceptions of changes in procedures in a teamworking intervention. *Work & stress*, 26(2), 91-111. <https://doi.org/10.1080/02678373.2012.682721>
- Nielsen, K. & Shepherd, R. (2022). Understanding the outcomes of training to improve employee mental health: A novel framework for training transfer and effectiveness evaluation. *Work & Stress*, 36(4), 377-391. <https://doi.org/10.1080/02678373.2022.2028318>
- Pawson, R., & Tilley, N. (1997). *Realistic evaluation*. London, California and New Delhi: Sage.
- Randall, R., Nielsen, K., & Tvedt, S. D. (2009). The development of five scales to measure employees’ appraisals of organizational-level stress management interventions. *Work & Stress*, 23(1), 1-23. <https://doi.org/10.1080/02678370902815277>
- Richardson, K. M., & Rothstein, H. R. (2008). Effects of occupational stress management intervention programs: a meta-analysis. *Journal of Occupational Health Psychology*, 13(1), 69-93. <https://doi.org/10.1037/1076-8998.13.1.69>
- Rönkkö, M., & Cho, E. (2022). An updated guideline for assessing discriminant validity. *Organizational research methods*, 25(1), 6-14. <https://doi.org/10.1177/1094428120968614>

- Roodbari, H., Nielsen, K., & Axtell, C. (2023). What works for whom in which circumstances? An integrated realist evaluation model for organisational interventions. *Scandinavian Journal of Work and Organizational Psychology*, 8(1). DOI: 10.16993/sjwop.171
- Rugulies, R., Aust, B., Greiner, B. A., Arensman, E., Kawakami, N., LaMontagne, A. D., & Madsen, I. E. (2023). Work-related causes of mental health conditions and interventions for their improvement in workplaces. *The Lancet*, 402(10410), 1368-1381. DOI: 10.1016/S0140-6736(23)00869-3
- Saks, A. M., & Belcourt, M. (2006). An investigation of training activities and transfer of training in organizations. *Human Resource Management*, 45(4), 629-648.
<https://doi.org/10.1002/hrm.20135>
- Salas, E., Tannenbaum, S. I., Kraiger, K., & Smith-Jentsch, K. A. (2012). The science of training and development in organizations: What matters in practice. *Psychological science in the public interest*, 13(2), 74-101.
<https://doi.org/10.1177/1529100612436661>
- Shann, C., Martin, A., Chester, A., & Ruddock, S. (2019). Effectiveness and application of an online leadership intervention to promote mental health and reduce depression-related stigma in organizations. *Journal of Occupational Health Psychology*, 24(1), 20—35. <https://doi.org/10.1037/ocp0000110>
- Schaufeli, W. B., Bakker, A. B., & Salanova, M. (2006). The measurement of work engagement with a short questionnaire: A cross-national study. *Educational and psychological measurement*, 66(4), 701-716.
<https://doi.org/10.1177/0013164405282471>
- Schaufeli, W. B., Salanova, M., González-Romá, V., & Bakker, A. B. (2002). The measurement of engagement and burnout: A two sample confirmatory factor analytic

approach. *Journal of Happiness studies*, 3, 71-92.

<https://doi.org/10.1023/A:1015630930326>

Schoeb, G., Lafrenière-Carrier, B., Lauzier, M., & Courcy, F. (2021). Measuring transfer of training: Review and implications for future research. *Canadian Journal of Administrative Sciences/Revue Canadienne des Sciences de l'Administration*, 38(1), 17-28. <https://doi.org/10.1002/cjas.1577>

Sørensen, O. H., & Holman, D. (2014). A participative intervention to improve employee well-being in knowledge work jobs: A mixed-methods evaluation study. *Work & Stress*, 28(1), 67-86. <https://doi.org/10.1080/02678373.2013.876124>

Sorensen, G., Sparer, E., Williams, J. A., Gundersen, D., Boden, L. I., Dennerlein, J. T., ... & Pronk, N. P. (2018). Measuring best practices for workplace safety, health and wellbeing: the workplace integrated safety and health assessment. *Journal of Occupational and Environmental Medicine*, 60(5), 430-439. <https://doi.org/10.1097/jom.0000000000001286>

Stansfeld, S. A., Kerry, S., Chandola, T., Russell, J., Berney, L., Hounscome, N., ... & Bhui, K. (2015). Pilot study of a cluster randomised trial of a guided e-learning health promotion intervention for managers based on management standards for the improvement of employee well-being and reduction of sickness absence: GEM Study. *BMJ open*, 5(10), e007981. <https://doi.org/10.1136/bmjopen-2015-007981>

Tafvelin, S., Nielsen, K., Abildgaard, J. S., Richter, A., von Thiele Schwarz, U., & Hasson, H. (2019). Leader-team perceptual distance affects outcomes of leadership training: Examining safety leadership and follower safety self-efficacy. *Safety Science*, 120, 25-31. <https://doi.org/10.1016/j.ssci.2019.06.019>

Van Agteren, J., Iasiello, M., Lo, L., Bartholomaeus, J., Kopsaftis, Z., Carey, M., & Kyrios, M. (2021). A systematic review and meta-analysis of psychological interventions to

- improve mental wellbeing. *Nature human behaviour*, 5(5), 631-652.
<https://doi.org/10.1038/s41562-021-01093-w>
- Van De Schoot, R., Schmidt, P., De Beuckelaer, A., Lek, K., & Zondervan-Zwijnenburg, M. (2015). Measurement invariance. *Frontiers in psychology*, 6, 1064.
<https://doi.org/10.3389/fpsyg.2015.01064>
- Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational research methods*, 3(1), 4-70.
<https://doi.org/10.1177/10944281003100>
- Vander Elst, T., De Witte, H., & De Cuyper, N. (2014). The Job Insecurity Scale: A psychometric evaluation across five European countries. *European Journal of Work and Organizational Psychology*, 23(3), 364–380.
<https://doi.org/10.1080/1359432x.2012.745989>
- Vignoli, M., Mariani, M. G., Guglielmi, D., & Violante, F. S. (2018). Leadership styles and self-efficacy in determining transfer intentions of safety training. *Journal of Workplace Learning*, 30(1), 65-76. <https://doi.org/10.1016/j.paid.2018.12.027>
- Vuori, J., Toppinen-Tanner, S., & Mutanen, P. (2012). Effects of resource-building group intervention on career management and mental health in work organizations: randomized controlled field trial. *Journal of Applied Psychology*, 97(2), 273.
<https://doi.org/10.1037/a0025584>
- von Thiele Schwarz, U., Nielsen, K. M., Stenfors-Hayes, T., & Hasson, H. (2017). Using kaizen to improve employee well-being: Results from two organizational intervention studies. *Human relations*, 70(8), 966-993. <https://doi.org/10.1177/0018726716677071>
- Watkins, M. W. (2018). Exploratory factor analysis: A guide to best practice. *Journal of black psychology*, 44(3), 219-246. <https://doi.org/10.1177/0095798418771807>

- Weiner, B. J. (2020). A theory of organizational readiness for change. In P. Nilsen & S.A. Birken (Eds), *Handbook on Implementation Science* (pp. 215–232). Northampton, MA: Edward Elgar Publishing. <https://doi.org/10.4337/9781788975995.00015>
- West, M., Markiewicz, L, & Dawson, J.F. (2004). *TPI: The Team Performance Inventory: User guide*. Birmingham: Aston Organisation Development.
- World Health Organization [WHO] (2022). *WHO guidelines on mental health at work*. Geneva: World Health Organization.
- Yelon, S., Sheppard, L., Sleight, D., & Ford, J. K. (2004). Intention to transfer: How do autonomous professionals become motivated to use new ideas? *Performance Improvement Quarterly*, 17(2), 82-103. <https://doi.org/10.1111/j.1937-8327.2004.tb00309.x>

Table 1. Participants' demographic information for each time point

	T1	T2	T3	T4	T5	T6
Full sample (employees and leaders)						
N questionnaires distributed / returned	2730 / 1966	1331 / 569	981 / 324	1259 / 508	1331 / 470	2730 / 874
Response rate	72.01%	42.75%	33.03%	40.35%	35.21%	32.01%
Gender						
Male	42.1%	45.7%	50.6%	43.1%	53.9%	48.8%
Female	56.7%	54.3%	49.4.2%	56.9%	46.1%	51.2%
Age						
Under 25 years	3%	1.5%	1.3%	1.3%	2.6%	1.7%
25–34 years	29.5%	41.5%	36.6%	38.5%	40.0%	30.5%
35–44 years	26.1%	31.6%	34.1%	29.4%	27.5%	30.0%
45–54 years	24.7%	16.3%	17.7%	20.5%	17.7%	20.7%
55–64 years	15.7%	7.9%	9.5%	8.9%	10.7%	15.1%
65+ years	1.2%	1.2%	.9%	1.3%	1.4%	2%
Educational level						
Primary	0.4%	0.2%	0.4%	0%	0%	0.6%
Secondary	7.9%	5.2%	6.9%	5.9%	5.8%	8.3%
Post-secondary	33.7%	26.5%	27.6%	32.1%	28.4%	28.5%
Bachelor	29.2%	23.5%	19.4%	20.8%	20%	27%
Master/Doctoral	27.1%	43.6%	44.8%	39.9%	44.3%	34.5%
Other	1.7%	1.0%	0.9%	1.3%	1.5%	1.1%
Country						
Italy	50.6%	26.0%	22.5%	39.6%	32.1%	40.7%
Netherlands	18.9%	10.4%	6.5%	5.5%	5.5%	16.0%
Spain	13.4%	18.6%	21.9%	18.3%	14.5%	15.3%
Czech Republic	12.3%	41.3%	42.9%	33.9%	43.8%	21.9%
Germany	4.8%	3.7%	6.2%	2.8%	4.0%	6.1%
Leaders						
N questionnaires returned	541	75	37	91	66	275
Gender						
Male	42.7%	69.2%	82.9%	77.4%	65.4%	60.4%
Female	57.3%	30.8%	17.1%	22.6%	34.6%	39.7%
Age						
Under 25 years	29.5%	0%	0%	1.2%	0%	0.4%
25–34 years	26.1%	41.3%	42.9%	56%	40.4%	21.8%
35–44 years	24.7%	27.7%	25.7%	25%	36.5%	29.8%
45–54 years	15.7%	15.4%	17.1%	11.9%	13.5%	25.5%
55–64 years	1%	6.2%	8.6%	1.2%	9.6%	18.5%
65+ years	0.2%	3.1%	5.7%	4.8%	0%	4%
Educational level						
Primary	0.6%	0%	0%	0%	0%	1.1%
Secondary	5.6%	1.5%	0%	1.2%	3.8%	6.2%
Post-secondary	17.9%	9.3%	20%	6%	34.6.9%	19.6%
Bachelor	39.2%	18.5%	17.1%	14.3%	30.8%	29.8%
Master/Doctoral	35%	70.8%	62.8%	78.6%	30.7%	42.9.7%
Other	1.8%	0%	0%	0%	0%	0.4%

Table 2. Description of the measures included in the IMTTAQ and validity analyses

Name	Item	Response scale	Measure Origin	Response Type	Role in Validity Testing
IMTTAQ scales					
Pre-training					
Leadership commitment	4	1 (Not all the time) to 4 (All of the time)	Adapted from López Gómez et al., 2023; IRT-validated (López Gómez et al., 2021); one item removed.	Employee – Rating of organization	Convergent / Discriminant / Predictive (Training transfer)
Dialogue processes	3	1 (Not all the time) to 4 (All of the time)	Adapted from López Gómez et al., 2023	Employee – Rating of organization	Convergent / Discriminant / Predictive (Training transfer)
Communication	3	1 (Strongly disagree) to 5 (Strongly agree)	Adapted from Nielsen et al., 2023; one item added; validated using EFA, CFA, and predictive regression analyses.	Employee – Rating of intervention	Convergent / Discriminant / Predictive (Training transfer)
Readiness for change	4	1 (Strongly disagree) to 5 (Strongly agree)	Adapted from Randall et al., 2009; validated using EFA, CFA, and predictive regression analyses.	Employee – Self-rating	Convergent / Discriminant / Predictive (Training transfer)
During training					
Training design	4	1 (Strongly disagree) to 5 (Strongly agree)	Adapted from Integrity of the Intervention (Vuori et al., 2012)	Employee – Rating of training	Convergent / Discriminant / Predictive (Training transfer)
Acceptability of content	3	1 (To a very low extent) to 4 (To a high extent)	Adapted from Acceptability (Martin, et al., 2020)	Employee – Self-rating	Convergent / Discriminant / Predictive (Training transfer)
Perception of facilitators	4	1 (Not all the time) to 5 (All of the time)	Adapted from Integrity of the intervention (Vuori, et al., 2012).	Employee – Rating of facilitators	Convergent / Discriminant / Predictive (Training transfer)
Intention to transfer	3	1 (Strongly disagree) to 5 (Strongly agree)	Developed for this study; based on Yelon et al., 2004.	Employee – Self-rating	Convergent / Discriminant / Predictive (Training transfer)
Training integration					
Training transfer	3	1 (Strongly disagree) to 5 (Strongly agree)	Adapted version of the Application to Practice subscale of the Questionnaire for Professional Training Evaluation, which was validated by Grohmann & Kauffeld (2013) using EFA and CFA. One item added.	Employee – Self-rating	Convergent / Discriminant / Predictive (Outcomes)
Opportunities to integrate	3	1 (Strongly disagree) to 5 (Strongly agree)	Adapted from Learning Transfer System Inventory (Holton III, et., 2000). Validated EFA (Holton et al., 2006)	Employee – Self-rating	Convergent / Discriminant
Manager support	3	1 (Strongly disagree) to 5 (Strongly agree)	Adapted from Social Stressors (Frese et al., 1999)	Employee – Rating of leader	Convergent / Discriminant
Individual peer support	3	1 (Strongly disagree) to 5 (Strongly agree)	Adapted from Social Stressors (Frese et al., 1999)	Employee – Rating of peers	Convergent / Discriminant
Group peer support	3	1 (Strongly disagree) to 5 (Strongly agree)	Adapted from Social Stressors (Frese et al., 1999)	Teams – Self-rating	Convergent / Discriminant

Subordinate support	3	1 (Strongly disagree) to 5 (Strongly agree)	Adapted from Social Stressors (Frese et al., 1999)	Leaders – Rating of Subordinates	Convergent / Discriminant
Training Outcomes					
Capacity building	4	1 (Has become better) to 3 (Has become worse)	Developed for this study; based on von Thiele Schwarz et al., 2017	Employee – Rating of self, team, leader, organization	Predictive
Additional measures included in the validation process					
Workload+	3	1 (Never) to 5 (Always)	HSE Indicator Tool (Edwards et al., 2008)	Employee – Self-rating	Convergent / Discriminant
Job insecurity+	4	1 (Totally disagree) to 5 (Totally agree)	Job Insecurity Scale (Vander Elst et al., 2014)	Employee – Self-rating	Convergent / Discriminant
Work-home conflict+	4	1 (Strongly disagree) to 5 (Strongly agree)	Work-life conflict (Innstrand et al., 2009)	Employee – Self-rating	Convergent / Discriminant
Autonomy+	4	1 (Not true at all) - 7 (Very true)	Short version of the Basic Psychological Need Satisfaction at Work (BPNS-W) (Eriksson & Boman, 2018)	Employee – Self-rating	Convergent / Discriminant
Team support+	3	1 (Strongly disagree) to 5 (Strongly agree)	Team Performance Inventory (West et al., 2004)	Team members – Rating of peers	Convergent / Discriminant
Burnout*	12	1 (Never) to 5 (Always)	Burnout Assessment Tool (BAT-12) (Hadžibajramović et al., 2022)	Employee – Self-rating	Predictive
Work engagement*	9	0 (Never) to 6 (Always)	Utrecht Work Engagement Scale (UWES-9) (Schaufeli, et al., 2006)	Employee – Self-rating	Predictive

+ included in convergent and discriminant validity, *included in analyses for predictive validity

Note. *Response Type* indicates who provided the data and the referent (e.g., self, leader, peer, subordinate, etc). *Role in Validity Testing* indicates the purpose of the measure in the validation study: Convergent = expected to correlate with related constructs at the same time point; Discriminant = expected to remain distinct from unrelated constructs; Predictive (Training transfer or outcomes) = expected to predict later outcomes, such as training transfer or outcome variables (e.g., capacity building).

Table 3. Time 1 - Pre-training EFA and CFA Results

	EFA			CFA	
	Factors			Factor	SE
	1	2	3	loadings	
Leadership commitment					
The company’s leadership, such as senior leaders and middle managers, communicate their commitment to a work environment that supports employee mental health and wellbeing.	.82	-.02	-.03	.82	.01
The organization allocates enough resources (such as enough workers and money) to implement policies or programs to protect and promote worker mental health and wellbeing.	.81	-.03	.02	.81	.01
Worker mental health and wellbeing is part of the organization’s mission, vision or business objectives	.85	.00	.01	.86	.01
The importance of mental health and wellbeing is consistently reflected in actions across all levels of the organization, both formally and informally.	.78	.03	.06	.79	.01
Dialogue processes					
Managers and employees work together in planning, implementing, and evaluating comprehensive mental health and wellbeing programs, policies, and practices for employees.	.82	.02	.04	.82	.01
In this organization, managers across all levels consistently seek employee involvement and feedback in decision making.	.75	.03	-.05	.80	.01
Employees are encouraged to voice concerns about mental health and wellbeing without fear of retaliation.	.75	-.03	-.12	.82	.01
Communication					
To what extent is the (NAME) intervention relevant for solving important problems in your organization?	.02	.02	-.91	.93	.01
I have received information about the goals of the (NAME) intervention	.01	-.01	-.97	.96	.01
It has been clearly communicated what is expected of me in the (NAME) intervention	.02	.03	-.90	.93	.01
Readiness for change					
I feel confident that I can use the (NAME) intervention to improve my mental health and wellbeing	.05	.85	.02	.88	.01
I have high expectations that the (NAME) intervention will improve my mental health	.04	.85	.08	.85	.01
I look forward to the changes brought about by the (NAME) intervention	-.05	.81	-.05	.78	.01
I am ready to accept the changes brought about by the (NAME) intervention	-.03	.68	-.08	.69	.02

Note. N = 1510. Extraction Method: Principal Axis Factoring. Rotation Method: Oblimin with Kaiser Normalization. Standardized factor values, SE = Standard errors. NAME = training name

Table 4. Time 2 – During training EFA and CFA Results

	EFA		CFA	
	Factors		Factor loadings	SE
	1	2		
Training design				
The activities and exercises the consultant(s) used helped me know how to apply my learning on the job	.82	-.07	.80	.02
It is clear to me that the consultant(s) conducting the training understand how I will use what I learn	.55	.20	.71	.02
The consultant(s) used good examples that showed me how I could use my learning on the job	.63	.13	.77	.02
The way the consultant(s) taught the material made me feel more confident I could apply it	.72	.06	.79	.02
Acceptability of content				
I feel the training met my expectations	.67	.19	.87	.01
I would recommend the training to others in a similar situation	.67	.21	.89	.01
The format of the training and the materials were appropriate for my needs	.66	.19	.87	.01
Perception of facilitators				
Did the consultant(s) make you feel like your participation was valued?	-.04	.86	.81	.02
Did the consultant(s) do or say something that made you think they understand your situation?	.23	.51	.68	.03
Did the consultant(s) encourage you to participate in the assignments they gave?	.10	.67	.75	.02
Intention to transfer				
I believe what I learned on the training can help me at work	.85	-.02	.88	.01
The skills I developed during the training will help me at work	.90	-.13	.89	.01
I developed new skills for my work that I didn't have before	.82	-.14	.75	.02
Did you find the atmosphere friendly and encouraging?	-.01	.77	.75	.02

Note. N = 569. Extraction Method: Principal Axis Factoring. Rotation Method: Oblimin with Kaiser Normalization. β = Standardized values, SE = Standard errors

Table 5. Time 3-5 – Training integration EFA results

	Solution 1		Solution 2			Solution 3			Solution 4		
	Factor		Factor			Factor			Factor		
	1	2	1	2	3	1	2	3	1	2	3
Training transfer											
In my everyday work, I often use the knowledge I gained in the training	.76	-.08	.08	-.07	.79	.67	-.10	.17	.70	.23	-.14
I successfully manage to apply the training contents in my everyday work	.78	.00	-.05	.04	.97	.66	-.01	.16	.71	.29	-.06
I am able to transfer the skills learned in training courses back to my actual job	.78	.05	.08	.05	.78	.61	.07	.22	.72	.08	.07
Opportunities to integrate											
I have been given the tasks necessary to apply the skills and knowledge I learned on the training	.68	.01	.74	.04	.04	.08	.02	.67	.73	-.07	.00
I have the necessary resources to use what I learned in training	.74	.04	.78	.05	.06	.04	.04	.79	.77	-.13	.12
I have the information necessary to apply the skills and knowledge I learned on the training course	.80	.00	.92	-.07	.04	-.03	.02	.93	.84	-.18	.16
Manager support											
My immediate manager asks me if I have problems or trouble at work	-.03	.85	-.12	.83	.03	.00	.88	.02	-.06	.14	.79
My immediate manager helps to make my work easier	.02	.92	-.09	.88	.10	-.02	.92	.03	.02	.07	.75
My immediate manager listens to me when I have problems	.02	.86	-.02	.81	-.01	.03	.86	-.01	.15	-.06	.77
Peer support											
My colleagues ask if I have problems or trouble at work			-.02	.57	.04						
My colleagues help to make my work easier			.17	.62	-.05						
My colleagues listen to me when I have problems			.15	.61	-.01						
Group support											
We appreciate each other's efforts to use new skills we have learned in training						.81	.00	-.01			
We encourage each other to use the skills we have learned in training						.89	.10	-.14			
At work, we expect each other to use what we have learned in training						.79	-.01	-.03			
Subordinate support											
My subordinates are understanding of my problems or trouble at work									-.09	.62	.18
My subordinates help to make my work easier									.08	.87	.03
My subordinates listen to me when I have problems									.09	.88	.00

Note. Solution 1 (N = 1189); Solution 2 (N = 343); Solution 3 (N = 692); Solution 4 (N = 139). Extraction Method: Principal Axis Factoring. Rotation Method: Oblimin with Kaiser Normalization.

Table 6. Time 3-5 – Training integration CFA results

	Solution 1		Solution 2		Solution 3		Solution 4	
	<i>Factor loadings</i>	<i>SE</i>	<i>Factor loadings</i>	<i>SE</i>	<i>Factor loadings</i>	<i>SE</i>	<i>Factor loadings</i>	<i>SE</i>
Training transfer								
In my everyday work, I often use the knowledge I gained in the training.	.81	.01	.81	.01	.82	.01	.81	.01
I successfully manage to apply the training contents in my everyday work.	.91	.01	.91	.01	.90	.01	.91	.01
I am able to transfer the skills learned in training courses back to my actual job.	.87	.01	.87	.01	.87	.01	.87	.01
Opportunities to integrate								
I have been given the tasks necessary to apply the skills and knowledge I learned on the training.	.76	.01	.76	.01	.76	.01	.76	.01
I have the necessary resources to use what I learned in training.	.85	.01	.85	.01	.85	.01	.85	.01
I have the information necessary to apply the skills and knowledge I learned on the training course.	.92	.01	.92	.01	.91	.01	.92	.01
Manager support								
My immediate manager asks me if I have problems or trouble at work.	.84	.01	.85	.01	.84	.01	.84	.01
My immediate manager helps to make my work easier.	.92	.01	.92	.01	.92	.01	.92	.01
My immediate manager listens to me when I have problems.	.87	.01	.87	.01	.87	.01	.87	.01
Peer support								
My colleagues ask if I have problems or trouble at work.			.55	.04				
My colleagues help to make my work easier.			.86	.03				
My colleagues listen to me when I have problems.			.79	.03				
Group support								
We appreciate each other's efforts to use new skills we have learned in training.					.86	.01		
We encourage each other to use the skills we have learned in training.					.87	.01		
At work, we expect each other to use what we have learned in training.					.80	.02		
Subordinate support								
My subordinates are understanding of my problems or trouble at work.							.64	.05
My subordinates help to make my work easier.							.95	.02
My subordinates listen to me when I have problems.							.93	.02

Note. Solution 1 ($N = 1189$); Solution 2 ($N = 343$); Solution 3 ($N = 692$); Solution 4 ($N = 139$). Standardized Factor Values, SE = Standard Errors.

Table 7. Time 6 – Training outcomes EFA and CFA results

	EFA	CFA	
	Factor 1	Factor loadings	SE
Capacity building			
My ability to manage my mental health and wellbeing	.61	.58	.03
In my team, the dialogue and honest discussions about mental health and wellbeing	.77	.73	.02
My leader's ability to manage mental health and wellbeing issues	.78	.79	.02
In my organization, the ability to create a good working environment and ensure workers' mental health and wellbeing	.83	.81	.02

N = 787. Extraction Method: Principal Axis Factoring. Rotation Method: Oblimin with Kaiser Normalization. CB = Capacity building. Standardized Factor Values, SE = Standard errors.

Table 8. Language measurement invariance – all time points

	$\chi^2(df)$	CFI	RMSEA	SRMR	Model comp	$\Delta\chi^2 (\Delta df)$	ΔCFI	$\Delta RMSEA$	$\Delta SRMR$	Decision
Time Point 1: Pre-training										
M1: Configural	763.583 (284)	.967	.068	.040	-	-	-	-	-	-
M2: Metric	837.415 (314)	.964	.068	.054	M1	73.832 (30)**	.003	0	-.014	Reject
M2': Partial Metric	785.145 (302)	.967	.066	.043	M1	21.562 (18)	0	.002	-.003	Accept
M3: Scalar	1277.975 (344)	.936	.087	.071	M2'	492.83 (42)**	.031	-.019	-.031	Reject
Time Point 2: During training										
M1: Configural	413.645 (213)	.954	.076	.049	-	-	-	-	-	-
M2: Metric	433.576 (233)	.954	.073	.068	M1	19.931 (20)	0	.003	-.019	Accept
M3: Scalar	538.166 (253)	.935	.083	.079	M2	104.59 (20)**	.019	-.01	-.011	Reject
Time Points 3-5: Training integration										
M1: Configural	171.179 (72)	.986	.059	.031	-	-	-	-	-	-
M2: Metric	196.244 (84)	.984	.058	.057	M1	25.065 (12)**	.002	.001	-.026	Reject
M2': Partial Metric	182.471 (82)	.985	.057	.042	M1	11.292 (2)	.001	.002	-.011	Accept
M3: Scalar invariance	296.605 (96)	.971	.073	.067	M2'	114.134 (14)**	.015	-.014	-.036	Reject
Time Point 6: Training outcomes										
M1: Configural	23.105 (8)	.983	.101	.027	-	-	-	-	-	-
M2: Metric invariance	41.31 (17)	.972	.088	.071	M1	18.205 (9)**	.011	.013	-.044	Reject
M2': Partial Metric	29.333 (14)	.983	.077	.049	M1	6.228 (6)	0	.024	-.022	Accept
M3: Scalar	94.621 (26)	.922	.119	.089	M2'	65.288 (12)**	.061	-.018	-.062	Reject

Note. ** $p < .01$. Time 1 $N = 1448$ (Italian $n = 701$, Spanish $n = 236$, Czech $n = 202$, Dutch $n = 309$, German $n = 67^a$); Time 2 $N = 489$ (Italian $n = 148$, Spanish $n = 106$, Czech $n = 235$, Dutch $n = 59^a$, German $n = 21^a$); Time 3-5 $N = 1174$ (Italian $n = 425$, Spanish $n = 232$, Czech $n = 517$, Dutch $n = 75^a$, German $n = 53^a$); Time 6 $N = 740$ (Italian $n = 317$, Spanish $n = 122$, Czech $n = 179$, Dutch $n = 122$, German $n = 47^a$).^a Subsamples with $n < 100$ are presented for transparency but were excluded from the invariance analyses due to insufficient sample size for confirmatory factor analysis (CFA).

Table 9. Time 3-5 – Temporal invariance

Model	$\chi^2 (df)$	CFI	RMSEA	SRMR	Model comp	$\Delta\chi^2 (\Delta df)$	ΔCFI	$\Delta RMSEA$	$\Delta SRMR$	Decision
M1: Configural invariance	196.94 (72)	.984	.063	.032	-	-	-	-	-	-
M2: Metric invariance	203.53 (84)	.985	.057	.037	M1	6.589 (12)	.001	.006	.005	Accept
M3: Scalar invariance	207.37 (96)	.986	.052	.037	M2	3.842 (12)	.001	.005	.000	Accept

Note. $N=1302$; Time 3 $n=324$; Time 4 $n=508$; Time 5 $n=470$.

Table 10. Reliabilities (Cronbach alpha)

	Cronbach alpha					
	full sample	Italy	Netherlands	Spain	Czech Republic	Germany
Time Point 1: Pre-training						
Leadership commitment	.89	.88	.85	.92	.83	.85
Dialogue processes	.85	.89	.78	.87	.81	.73
Communication	.96	.97	.91	.92	.92	.96
Readiness for change	.88	.88	.83	.90	.86	.92
Time Point 2: During training						
Training design	.85	.85	.83	.82	.85	.85
Acceptability of content	.91	.94	.90	.81	.88	.82
Perception of facilitators	.83	.83	.80	.82	.82	.47
Intention to transfer	.87	.90	.87	.78	.86	.84
Time Points 3-5: Training integration						
Training transfer	.89	.90	.90	.81	.92	.85
Opportunities to integrate	.87	.86	.85	.83	.90	.57
Manager support	.91	.93	.92	.86	.88	.86
Peer support	.75	.75	.78	.79	.70	.63
Group Support	.88	.88	.84	.77	.91	-. ^a
Subordinate support	.86	.93	-. ^a	.43	.95	.75
Time Point 6: Training outcomes						
Capacity building	.81	.79	.63	.84	.77	.63

Note. Time 1 N= 1486 (Italy n=689, Spain n= 226, Czech Republic n= 198, Netherlands n=306, Germany n=67); Time 2 N=568 (Italy n=147, Spain n=106, Czech Republic n=235, Netherlands n=59, Germany n=21); Time 3-5 N=1273 (Italy n=418, Spain n=223, Czech Republic n=508, Netherlands n=75, Germany n=50); Time 6 N=777 (Italy n= 313, Spain n=122 , Czech Republic n=177, Netherlands n=121, Germany n=44). T3–5 values represent variables with data at all levels. ^a Not enough cases to estimate reliability.

Table 11. Descriptive statistics, zero-order correlations for pre-training, during training, integration, and outcomes variables

	Mean	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1. WL (T1) ^a	2.81	.77																								
2. JI (T1) ^a	2.17	.88	.20**																							
3. WFC (T1) ^a	3.26	1.00	.49**	.13**																						
4. AT (T1) ^a	4.59	1.23	-.17**	-.04	-.32**																					
5. TS (T1) ^a	3.99	.86	-.09	-.01	-.17**	.48**																				
6. LC (T1)	2.22	.76	-.22**	.00	-.32**	.38**	.19**																			
7. DP (T1)	2.25	.74	-.21**	-.06	-.31**	.46**	.26**	.80**																		
8. CM (T1)	3.30	1.14	-.09	.02	-.20**	.40**	.29**	.47**	.48**																	
9. RC (T1)	3.48	.73	-.11*	-.02	-.16**	.27**	.18**	.35**	.35**	.54**																
10. TD (T2)	3.83	.68	-.06	-.15*	-.11	.11	.20**	.28**	.30**	.26**	.34**															
11. AC (T2)	3.77	.88	-.02	-.20**	-.09	.20**	.11	.17*	.24**	.18**	.35**	.70**														
12. PF (T2)	4.19	.61	-.02	-.13	-.04	.26**	.30**	.20**	.23**	.25**	.31**	.64**	.67**													
13. IT (T2)	3.72	.77	-.04	-.18**	-.09	.13*	.12	.28**	.27**	.21**	.40**	.75**	.69**	.57**												
14. TT (T3)	3.38	.78	-.04	-.12	-.15	.19*	.07	.33**	.23**	.24**	.40**	.54**	.52**	.49**	.60**											
15. OI (T3)	3.85	.64	-.17	-.21*	-.06	.19*	.25**	.27**	.25**	.32**	.32**	.57**	.52**	.48**	.53**	.58**										
16. MS (T3)	3.71	.98	-.24**	-.12	-.28**	.40**	.29**	.36**	.38**	.25**	.16	.17**	.12*	.16**	.20**	.20**	.27**									
17. TT (T4)	3.29	.77	.00	-.10	-.04	.01	-.09	.19**	.23**	.19*	.26**	.49**	.47**	.33**	.53**	.54**	.30**	.32**								
18. OI (T4)	3.73	.75	-.03	-.19**	-.14	.11	.06	.31**	.36**	.20**	.28**	.61**	.58**	.47**	.57**	.42**	.49**	.39**	.60**							
19. MS (T4)	3.66	.95	-.12	.00	-.18*	.32**	.19**	.27**	.28**	.30**	.07	.27**	.13*	.19**	.22**	.23**	.26**	.74**	.28**	.20**						
20. TT (T5)	3.33	.84	.00	-.11	.01	.06	-.03	.13	.13	.25**	.33**	.37**	.32**	.30**	.33**	.53**	.23*	.11	.55**	.36**	.05					
21. OI (T5)	3.77	.73	.05	-.17*	-.15	.14	.10	.15	.15	.25**	.30**	.44**	.49**	.43**	.39**	.33**	.30**	.37**	.33**	.43**	.22**	.65**				
22. MS (T5)	3.75	.89	-.05	.06	-.18*	.28**	.39**	.23**	.29**	.29**	.11	.29**	.14	.18*	.16*	.17	.23*	.57**	.24**	.24**	.59**	.25**	.29**			
23. CB (T6)	1.94	.53	.09	.15*	.10	-.03	.03	.12	.08	-.01	.00	-.18*	-.33**	-.28**	-.14	-.12	-.17	-.14	.06	-.11	-.06	-.05	-.24**	.06		
24. WE (T6) ^a	3.89	1.11	-.03	-.07	-.25**	.21**	.10	.23**	.23**	.09	.13	.10	-.03	-.02	.00	.33**	.25*	.25*	.23**	.12	.17	.17*	.15	.20*	.31**	
25. BO (T6) ^a	2.30	.62	.24**	.10	.43**	-.27**	-.24**	-.23**	-.26**	-.19**	-.18**	-.04	.01	.01	.08	-.13	-.16	-.07	-.16	-.05	-.05	-.22**	-.12	-.24**	-.23**	-.64**

Note. N = 321-691. ** p < .01, * p < .05. WL = Workload; JI = Job insecurity; WFC = Work-home conflict; AT = Autonomy; TS = Team support; LC = Leadership commitment; DP = Dialogue process; CM = Communication; RC = Readiness for change; TD = Training design; AC = Acceptability of content; PF = Perception of facilitators; IT = Intention to transfer; TT = Training transfer; OI = Opportunities to integrate; MS = Manager support; TT = Training transfer; OI = Opportunities to integrate; MS = Manager support; CB = Capacity building; WE = Work engagement; BO = Burnout. ^a = Work environment and wellbeing variables used only as part of the validation process, not part of IMTTAQ

Table 12. Regression Analyses of Pre- and During-Training Variables Predicting Training Transfer

Pre-training	Time 3	Training transfer	
		Time 4	Time 5
<i>Intercept</i>	1.70 (.32)**	1.96 (.30)**	1.81 (.38)**
<i>Direct effects</i>			
Leadership commitment	.37 (.13)**	.01 (.13)	.03 (.15)
Dialogue processes	-.18 (.14)	.17 (.13)	-.00 (.15)
Communication	-.03 (.07)	.01 (.07)	.10 (.08)
Readiness for change	.36 (.10)**	.22 (.10)*	.32 (.11)**
<i>F (df1, df2)</i>	8.20 (4, 126)	4.30 (4, 177)	4.57 (4, 135)
<i>R</i> ² Model	.20	.09	.12
During-training	Time 3	Training transfer	
		Time 4	Time 5
<i>Intercept</i>	.12 (.27)	1.50 (.28)**	.92 (.47)
<i>Direct effects</i>			
Training design	.12 (.09)	.20 (.10)*	.29 (.13)*
Acceptability of content	.07 (.08)	.14 (.07)	.10 (.10)
Perception of facilitators	.22 (.08)*	-.13 (.09)	.06 (.14)
Intention to transfer	.42 (.08)**	.29 (.09)**	.14 (.12)
<i>F (df1, df2)</i>	49.73 (4, 289)	28.04 (4, 264)	10.32 (4, 189)
<i>R</i> ² Model	.41	.30	.18

Note. Pre-training: Time 3 n=131; Time 4 n=182; Time 5 n=140. During training: Time 3 n=294; Time 4 n=260; Time 5 n=194. Unstandardized estimates. * $p < .05$. ** $p < .01$.

Table 13. Regression analysis for predictive validity of training transfer on trainees' capacity building and being outcomes

	Capacity building	Work engagement	Burnout
Time 3			
<i>Intercept</i>	2.26 (.34)**	2.23 (.51)**	2.71 (.34)**
Training transfer (T3)	-.11 (.10)	.47 (.15)*	-.11 (.10)
F (<i>df1</i> , <i>df2</i>)	.37 (1, 79)	7.68 (1, 84)	.44 (1, 84)
<i>R</i> ² Model	.02	.11	.02
Time 4			
<i>Intercept</i>	1.81 (.24)	2.84 (.45)**	2.68 (.26)**
Training transfer (T4)	.05 (.07)	.36 (.14)*	-.14 (.08)
F (<i>df1</i> , <i>df2</i>)	.11 (1, 118)	7.22 (1, 126)	1.15 (1, 126)
<i>R</i> ² Model	.01	.05	.03
Time 5			
<i>Intercept</i>	2.07 (.20)**	3.30 (.38)**	2.79 (.21)**
Training transfer (T5)	-.03 (.06)	.22 (.11)*	-.15 (.06)*
F (<i>df1</i> , <i>df2</i>)	.08 (1, 125)	4.24 (1, 127)	2.05 (1, 127)
<i>R</i> ² Model	.01	.03	.05

Note. Time 3 n=57; Time 4 n=100; Time 5 n=107. Unstandardized estimates. * $p < .05$. ** $p < .01$.

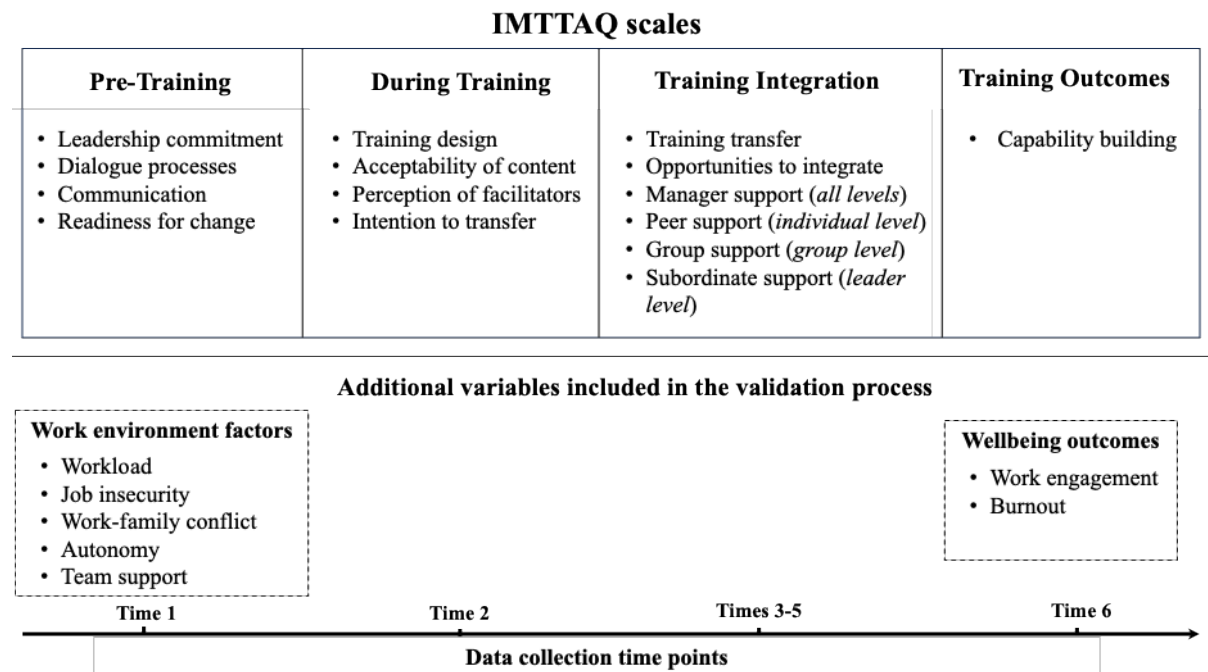


Figure 1. IMTTAQ measurement framework and validation variables. The upper section presents the core IMTTAQ scales organized across four temporal phases: Pre-Training (T1), During Training (T2), Training Integration (T3-5), and Training Outcomes (T6). The lower section displays the additional validation measures included separately from the IMTTAQ instrument to establish its psychometric properties. Data collection occurred at six time points as indicated by the timeline.

Appendix A

Table A1. Time 1 – Pre- training CFA Comparison of Measurement Models

Model	Factors	df	χ^2	$\Delta\chi^2$	CFI	TLI	RMSEA	SRMR
Model 1	Four factors (LC, DIAL, COM, READ)	71	554.870		.969	.961	.067	.028
Model 2	Three factors (COM, READ, SM)	74	686.016	131.146**	.961	.953	.074	.032
Model 3	One factor (Harman test)	77	7282.398	6727.528**	.546	.463	.249	.159

Note: N = 1510 employees. ** $p < .01$. LC = Leadership commitment, DIAL = Dialogue processes, COM = Communication, READ = Readiness for change. SM = Combined factor including leadership commitment and dialogue processes

Table A2. Time 2 – During training CFA. Comparison of Measurement Models

Model	Factors	df	χ^2	$\Delta\chi^2$	CFI	TLI	RMSEA	SRMR
Model 1	Four factors (INTT, TDES, ACCP, FACI)	71	204.320		.974	.967	.057	.035
Model 2	Two factors (INTA, FACI)	76	652.880	448.560**	.888	.866	.115	.052
Model 3	One factor (Harman test)	77	1001.452	797.132**	.821	.788	.145	.072

Note: N = 569 employees. ** $p < .01$. INTT = Intention to transfer, TDES = Training design, ACCP = Acceptability of content, FACI = Perception of facilitators, INTA = Combined factor including intention to transfer, training design and acceptability of content

Table A3. Time 3-5 – Training integration CFA. Comparison of Measurement Models

Model	Factors	df	χ^2	$\Delta\chi^2$	CFI	TLI	RMSEA	SRMR
Solution 1								
Model 1	Three factors (OPP, TRANS, MAN)	24	125.714		.987	.980	.057	.025
Model 2	Two factors (TRA, MAN)	26	1146.371	1020.66**	.855	.799	.182	.065
Model 3	One factor (Harman test)	27	3450.791	3325.08**	.556	.408	.312	.169
Solution 2								
Model 1	Four factors (OPP, TRANS, MAN, PEER)	48	247.762		.976	.967	.057	.040
Model 2	Three factors (OPP, TRANS, SUP)	51	409.766	162.004**	.957	.944	.074	.068
Model 3	One factor (Harman test)	54	4017.062	3769.30**	.524	.419	.237	.191
Solution 3								
Model 1	Four factors (OPP, TRANS, MAN, GPEER)	48	217.765		.982	.975	.065	.028
Model 2	Three factors (TRA, MAN, GPEER)	51	565.739	347.974**	.945	.929	.088	.046
Model 3	One factor (Harman test)	54	3841.147	3623.38**	.598	.509	.232	.137
Solution 4								
Model 1	Four factors (OPP, TRANS, MAN, LSUB)	48	173.972		.984	.978	.045	.045
Model 2	Three factors (OPP, TRANS, SSUP)	51	426.034	252.026**	.953	.940	.075	.120
Model 3	One factor (Harman test)	54	3750.693	3324.66**	.540	.438	.229	.181

Note: N = 1302 employees. ** $p < .01$. OPP = Opportunities to integrate, TRANS = Training transfer, MAN = Manager support, PEER = Peer support, GPEER = Peer support, LSUB = Leader subordinate support, TRA = Combined factor including opportunities to integrate and training transfer, SUP = Combined factor including manager support and peer support, SSUP = Combined factor including manager support and subordinate support

Table A4. Time 3-5 – Training integration CFA results. Standardized factor loadings for each time point

	T3		T4		T5		T3 – T5	
	CFA		CFA		CFA		CFA	
	<i>Factor loadings</i>	<i>SE</i>	<i>Factor loadings</i>	<i>SE</i>	<i>Factor loadings</i>	<i>SE</i>	<i>Factor loadings</i>	<i>SE</i>
Training transfer								
In my everyday work, I often use the knowledge I gained in the training	.78	.03	.79	.02	.85	.02	.81	.01
I successfully manage to apply the training contents in my everyday work	.93	.02	.89	.02	.91	.01	.91	.01
I am able to transfer the skills learned in training courses back to my actual job	.87	.02	.86	.02	.88	.02	.87	.01
Opportunities to integrate								
I have been given the tasks necessary to apply the skills and knowledge I learned on the training	.71	.03	.78	.02	.77	.02	.76	.01
I have the necessary resources to use what I learned in training	.83	.03	.86	.02	.86	.02	.85	.01
I have the information necessary to apply the skills and knowledge I learned in the training course	.87	.02	.93	.01	.92	.01	.92	.01
Manager support								
My immediate manager asks me if I have problems or trouble at work	.84	.02	.85	.02	.84	.02	.84	.01
My immediate manager helps to make my work easier	.91	.02	.94	.01	.91	.01	.92	.01
My immediate manager listens to me when I have problems	.87	.02	.86	.02	.87	.02	.87	.01

Note. N=1302; Time 3 n=324; Time 4 n=508; Time 5 n=470. Standardized Factor Values, SE = Standard errors.