Kaizen event process quality: towards a phase-based understanding of high-quality group problem-solving

José C.M. Franken, Desirée H. van Dun and Celeste P.M. Wilderom
Industrial Engineering and Business Information Systems, Faculty of Behavioural Management and Social sciences, University of Twente, Enschede, The Netherlands

Abstract

Purpose – As a problem-solving tool, the kaizen event (KE) is underutilised in practice. Assuming this is due to a lack of group process quality during those events, the authors aimed to grasp what is needed during high-quality KE meetings. Guided by the phased approach for structured problem-solving, the authors built and explored a measure for enriching future KE research.

Design/methodology/approach – Six phases were used to code all verbal contributions ($N = 5,442$) in 21 diverse, videotaped KE meetings. Resembling state space grids, the authors visualised the course of each meeting with line graphs which were shown to ten individual kaizen experts as well as to the filmed kaizen groups.

Findings – From their reactions to the graphs the authors extracted high-quality KE process characteristics. At the end of each phase, that should be enacted sequentially, explicit group consensus appeared to be crucial. Some of the groups spent too little time on a group-shared understanding of the problem and its root causes. Surprisingly, the mixed-methods data suggested that small and infrequent deviations ("jumps") to another phase might be necessary for a high-quality process. According to the newly developed quantitative process measure, when groups often jump from one phase to a distant, previous or next phase, this relates to low KE process quality.

Originality/value – A refined conceptual model and research agenda are offered for generating better solutions during KEs, and the authors urge examinations of the effects of well-crafted KE training.

Keywords Kaizen event, Structured problem-solving, Group process quality

Paper type Research paper

1. Introduction

Over the last few decades, many organisations have adopted continuous improvement strategies such as lean, agile, six sigma or total quality management (TQM). To sustain such improvements, many organisations now aim to become full "learning organisations" (Bateman, 2005; Hines *et al.*, 2004; Netland, 2016; Tortorella *et al.*, 2020) in which it is the norm to pool knowledge to solve persistent (operational) problems effectively (Argote and Hora, 2017). The aforementioned strategies all engage in multidisciplinary group effort, including kaizen (Imai, 1997; Liker, 2004); DMAIC (De Mast, 2011); and quality circles (Murray and Chapman, 2003; Rafaai *et al.*, 2018). When uniting individuals from different functional backgrounds to solve a problem, they hope to be more effective than other previously used approaches (Hackman and Morris, 1975; Murray and Chapman, 2003; Shaw, 1932). Despite how well intended or needed a group effort is, its process can be unruly and/or lacking in

effectiveness (Mathieu *et al.*, 2017); in fact, group-based problem-solving often fails (Bessant *et al.*, 2001; Jurburg *et al.*, 2017). Merely bringing organisational members with relevant knowledge together to solve a persisting problem is seldom sufficient (Bateman, 2005; Kolb *et al.*, 2008). Yet, there is some evidence that managing the group process will impact on its effects (Kolb *et al.*, 2008; LaFasto and Larson, 2001). Groups must have an adequate mechanism to solve the problem at hand in which all individual group members share their knowledge (Furlan *et al.*, 2019). Next to studies on how individuals may contribute to this mechanism (e.g. Galeazzo and Furlan, 2019), a research focus on the group process is important, as this knowledge is invaluable to becoming a learning organisation (Bessant *et al.*, 2001; Vo *et al.*, 2019). Given that many organisations wrestle with developing into a full learning organisation, there have been many calls for scholarly work on the characteristics of high-quality, group-based problem-solving (Bateman, 2005; Jadhav *et al.*, 2014; Netland, 2016). To better learn how to reap the potential benefits of taking a multidisciplinary group-based problem-solving approach, we explored the characteristics of the quality of the process.

Compared to an intuitive problem-solving approach, a structured approach is widely seen as more effective for solving a problem sustainably as a group (Mohaghegh and Furlan, 2020). A structured approach means a phased process, in which a group must go through a number of phases (Mohaghegh and Furlan, 2020). The evidence to support this basic assumption is, however, scarce and stems mainly from popularised professional literature (Imai, 1997; Kepner and Tregoe, 1965; Latzko and Saunders, 1996). Others have even challenged the need for structured problem-solving, by questioning the dictum to abide by consecutive phases (Pretz, 2008).

To contribute to the debate on how best to go about group problem-solving in organisations, we conducted exploratory mixed-methods research. We focussed on kaizen events (KEs) as an example of multidisciplinary group problem-solving (Bortolotti *et al.*, 2018) because it is a widespread – yet underused – approach in organisations and thus a good phenomenon for extending our knowledge about the quality of group problem-solving processes. Much research deals with understanding critical success factors for KEs (Aleu and van Aken, 2016). Most of these studies are rooted in Farris *et al.*'s (2009) model (Álvarez-García *et al.*, 2018), which is considered as "the only model that considers a comprehensive set of determinants, dividing them into input and process factors, through which they impact outcomes" (Bortolotti *et al.*, 2018, p. 555).

Farris *et al.*'s (2009) model, however, only briefly notes the quality of a KE group's *internal process* as a possible success factor in applying KEs. While they, and most other studies, have especially addressed input and output factors of effective KEs, none of the 98 publications that were reviewed by Aleu and van Aken (2016) studied the factor quality of tool application and KE process quality. In addition, more than 10 years after the publication of Farris *et al.*'s model, the available studies (e.g. Furlan *et al.*, 2019; Lam *et al.*, 2015; Liu *et al.*, 2015) only measured those factors through self-report. As the quality of the KE process is a critical success factor for KE, and practitioners are eager to learn how they could improve their KE outcomes, a deeper understanding of high-quality KE process quality is vital. Therefore, this study explores the question: *What are key characteristics of a high-quality kaizen event process?*

Besides obtaining detailed insight into the dynamics of a high-quality KE process from both experts' and participants' perspective, our study also contributes a new and more objective method for unlocking the features of a high-quality process in structured group problem-solving (Jones *et al.*, 2021). It entailed video-based analyses of 21 real-life KE meetings that were visualised using a state space grid type of method (Paoletti *et al.*, 2021). Subsequently, utilising model fit and process analysis theories such as six sigma, a quantitative process indicator was developed to assess the degree of a strict structure in any observable real-time KE process. We will show how this new method might add value to

future scholarly, as well as practical, efforts to raise the internal process quality of KEs and similarly structured problem-solving endeavours (Hassell and Cotton, 2017; Knapp, 2010; Powell and Coughlan, 2020).

Below, we first delineate relevant theoretical frameworks for group-based problem-solving. Then, we depict how our diverse empirical methods were employed to get the results, including how we arrived at the currently most informative indicator of the quality of multidisciplinary group meetings trying to solve a persisting organisational problem. In the Discussion, we sketch the theoretical implications of the findings, offer propositions and extend Farris *et al.*'s (2009) conceptual model in conjunction with a new research agenda.

2. Theory

2.1 Problem-solving in learning organisations

Organisational learning is defined as a behavioural renewal process that enables an organisation to achieve both change and growth simultaneously (Murray and Chapman, 2003). "Organizational learning includes processes of creating, retaining and transferring knowledge and has implications for the performance and competitiveness of organizations" (Argote and Hora, 2017, p. 579). Learning organisations aim to learn from different and especially better ways of organising their work (Örtenblad, 2001); in particular, their own members are encouraged to work together on more effective problem-solving (Ju *et al.*, 2021). Prominent examples of learning organisations are those that implement lean (Hines *et al.*, 2004; Tortorella *et al.*, 2020): working together to solve problems is an important lean principle (Liker, 2004). Hines *et al.* (2004) defined four lean implementation stages: (1) knowing; (2) understanding; (3) thinking; and (4) learning organisation. Individual- and group-level problem-solving practices must be established in the third stage (Hines *et al.*, 2004), often facilitated by external experts (Bateman, 2005; Netland, 2016). To grow to the next stage, individuals and groups need to be trained to better apply these practices and to improve understand of the continuous improvement philosophy (Hines *et al.*, 2004). In the fourth "learning organisation" stage, organisations acknowledge that the value of continuous improvement is in the people's mindset and that they are capable of selecting the best fitting solutions to solve operational problems (Hines *et al.*, 2004). Group-based structured problem-solving is essential for learning organisations as it instils a mindset of continuous improvement (Liker, 2004); applying it whenever a problem occurs should become a routine in learning organisations (Rother, 2019).

2.2 Group-based structured problem-solving methods for continuous improvement

If a group needs to solve a problem, it must decide how to approach the task. While problem-solving groups in experienced continuous improvement organisations often select, unconsciously, a structured approach (Imai, 1997; Liker, 2004), other groups rely on the experience and behavioural inclinations of individual group members or on external facilitators (Bateman, 2005; Kolb *et al.*, 2008; McFadzean, 2002). The literature on this score distinguishes between ill-structured, semi-structured and structured problem-solving approaches (Mohaghegh and Grössler, 2019). Only the latter type is seen as being suitable for solving problems sustainably (Mohaghegh and Furlan, 2020). Although many organisations implement improvement programmes and promote the use of structured group problem-solving, many groups struggle to reach satisfying outcomes over time (Bateman, 2005; Netland, 2016). It is even noted that such group problem-solving attempts can prevent organisations from learning, due to people's tendency to "fight fires" instead of finding the root cause and thus solving problems more sustainably (Tucker *et al.*, 2002).

Each popular continuous improvement strategy tends to promote an own structured problem-solving approach. Table 1 summarises a representative array: Kaizen and kaikaku within lean (Liker, 2004), DMAIC within six sigma (De Mast and Lokkerbol, 2012), quality circles within TQM (Rafaai *et al.*, 2018) and scrum within agile (Grass *et al.*, 2020; Putnik, 2012). Whereas kaizen is advised to be used for day-to-day problems for which some data is available, kaikaku is recommended for whenever a more fuzzy system-level change is needed to break through persistent organisational barriers (Imai, 1997). The goal of scrum is to deliver product innovation (Putnik, 2012), which is a slightly different focus compared to the other mentioned problem-solving strategies. The other aforementioned continuous improvement strategies focus on solving operational problems and have therefore more similarities (Andersson *et al.*, 2006) although some differences can be seen in terms of the group composition and roles. The compositions of kaizen groups are based on the perspectives required to solve a particular problem and all the group members are equal without clear roles (Imai, 1997). In DMAIC, a group is brought together for the task, with predefined roles. Agile scrum teams tend to pre-exist, with predefined roles, and the problem is brought to the group (Grass *et al.*, 2020). Despite these differences, all problem-solving groups are urged to follow a structured problem-solving process. To convince groups to apply such a structured approach (Bateman, 2005; De Mast, 2011; Jurburg *et al.*, 2017), feedback on group process performance and evidence related to the benefits of this approach are needed (Hassell and Cotton, 2017). To explore the least-studied "process" side of group-based problem-solving further, we focussed on kaizen or KE groups.

Structured problem-solving method	Kaizen event	Kaikaku	DMAIC	Quality circle	Scrum
Continuous improvement strategy	Lean	Lean	Six sigma	Total quality management	Agile
Goal	Incremental operational improvements	Breakthrough operational improvements	Improve existing operational problems	Operational quality improvement	Product innovation
Type of problems	Operational problems or improvement opportunities in	To overcome organisational barriers for continuous improvement	Process variances in process industry	Quality issues in production environment	Deliver product to market, including IT product development
Group composition	Multidisciplinary group brought together for the task	Multidisciplinary group brought together for the task	Multidisciplinary group brought together for the task	Sub-set of team members from existing team	Existing Sprint team
Role structure	Equality within the group	Equality within the group	Pre-defined roles	Existing roles	Pre-defined roles
Typical duration	1 h to 5 days	Depending on the challenge	Depending on the problem	Depending on the problem	Sprint time slot (ranging between 1 week and 1 month)

Note(s): Scrum is included in this table, because the method is often associated with problem-solving in operational work settings (Putnik, 2012)

Table 1.
Overview of group-based structured operational problem-solving methods

2.3 Kaizen events and their internal process

Kaizen translates to "improving for the better" and involves both the needed mindset for continuous improvement and the KE approach to group-based problem-solving (Liker, 2004). A KE is defined as "a structured project performed by a multidisciplinary group with the aim of improving a targeted work area or process in a given timeframe" (Bortolotti *et al.*, 2018, p. 555). Typically, KEs last 1 h, a half day, a full day or, occasionally, five days (Glover *et al.*, 2014). Often, a KE consists of different meetings over a period of time (Glover *et al.*, 2013). KE outcomes may be influenced by various factors, including the quality of the KE internal process and tool quality, both of which relate to the way the group executes the KE (Farris *et al.*, 2009; Glover *et al.*, 2013). Given that KE groups are composed of individual members from different departments who are brought together for the occasion, they lack the time to grow into an effective team with clear roles and responsibilities (Fisher *et al.*, 1997; Wheelan, 2009). Thus, a particular logical structure is essential for a KE process in which group members can contribute effectively (Byun *et al.*, 2014).

Even though most kaizen studies have treated kaizen as a *whole* event (e.g. Bortolotti *et al.*, 2018; Farris *et al.*, 2009; Glover *et al.*, 2013), kaizen is, in practice, a *phased*, structured approach to problem-solving. The common factor of the previously noted problem-solving instruments is that the prescribed problem-solving processes follow a phased approach (Andersson *et al.*, 2006). Table 2 compares the distinguishable phases in those and other popular group-based problem-solving instruments, based on our review of the scholarly and professional literature. Notably, each instrument uses different labels and a different number of phases. Because a KE group consists of individuals who co-construct KE process quality by solving problems through sharing accurate, timely and relevant knowledge (Galeazzo and Furlan, 2019), we also matched these approaches with the six phases that were identified by Woods (2000) in his careful, well-grounded analysis of over 150 individual-level problem-solving strategies. Individuals' problem-solving behaviours are known to affect (emerging) group processes (Kozlowski, 2015) and as individuals tend to follow a personal problem-solving strategy (Woods, 2000; Yeo and Marquardt, 2010), we decided to adopt Woods' six-phased approach that also resembles popular practice (McKinsey & Company, 2003). Below, we summarise the key aspects of those six derived KE phases.

In the first *"problem definition"* phase, all aspects of a perceived problem are explored (Choo, 2014; Liker, 2004; Shingō, 2007). It is essential that the group first develops a shared perspective on the observed issue and shares a sense of urgency to solve it (Liker, 2004). The group must therefore explore all the different aspects of the problem, leading to a description of the ideal situation and an objective target condition (De Mast and Lokkerbol, 2012; Kepner and Tregoe, 1965). The KE group then decides whether the problem is important enough to continue the KE (Shingō, 2007).

The second *"root-cause analysis"* phase explores the reasons underlying the problem (Liker, 2004; Shingō, 2007). In this phase, group members are supposed to find the problem's cause(s), for instance, by asking "why" at least five times to really get to the root of it, as well as to measure the frequency and importance of each cause vis-à-vis the problem (De Mast and Lokkerbol, 2012; Liker, 2004). After mapping all the possible root causes, the group decides which one(s) should be addressed first (De Mast and Lokkerbol, 2012; Shingō, 2007).

Thirdly, group members must *"generate ideas"* to eliminate the selected root cause(s) by creatively sharing suggestions (Shingō, 2007). Typically, the open discussion among the KE group members leads to many possible solutions to a specific root cause (Shingō, 2007). On weighing the required resources to implement them, and the expected contribution to solving the root cause, the group selects the most promising ideas (Johnson and D'Lauro, 2018; Shingō, 2007).

The fourth phase encompasses *"plan implementation"*; sustainable idea implementation requires acceptance by others in the organisation. The KE group must therefore develop a

Problem-solving approaches	Individual problem-solving (Woods, 2000)	Toyota way (Liker, 2004)	DMAIC (De Mast and Lokkerbol, 2012)	Scientific thinking (Shingō, 2007)	Quality control circle (Rafaai et al., 2018)	Kaizen circle (McKinsey & Company, 2003)
Level	Individual	Group composed to solve problem	Group composed to solve problem	Group composed to solve problem	Existing team	Group composed to solve problem
Phases	(1) Problem definition	(1) Initial problem perception (2) Clarify the problem (3) Locate area, point of cause	(1) Define	(1) Create problem awareness (2) Pursuit of purpose	(1) Identify the problem	(1) Problem definition
	(2) Root-cause analysis	(4) Five times why investigation of root cause	(2) Measure (3) Analyse	(3) Understanding the status quo (4) Better methods (5) Lock on to problems	(2) Analyse the problem	(2) Root-cause analysis
	(3) Generate ideas			(6) Idea generation (7) Judgment	(3) Generate alternative solution (4) Select best solution	(3) Brainstorm
	(4) Plan implementation	(5) Counter-measure		(8) Propose	(5) Prepare plan of action (6) Present solution to managers	(4) Implementation plan
	(5) Implement		(4) Improve	(9) Implement	(7) Implement solution	(5) Implement
	(6) Check and sustain	(6) Evaluate (7) Standardize	(5) Control			(6) Check and sustain

Table 2.
Overview of the phases of structured problem-solving approaches

plan by discussing both the technical requirements and how to create buy-in, that is, through communication and training (McKinsey & Company, 2003; Rafaai *et al.*, 2018).

The fifth phase, *"implement"*, is about realising the plan and implementing the solution(s) at the operational level and integrating them in the standard operating procedures (De Mast and Lokkerbol, 2012; Liker, 2004).

Finally, during the *"check and sustain"* phase, an assessment is made whether the original problem has been eliminated after implementing the solution(s) (Liker, 2004). Furthermore, the newly set standard needs to be embedded in (daily) performance management systems and become a routine (De Mast and Lokkerbol, 2012; Latzko and Saunders, 1996; Liker, 2004). Eliminating the whole problem might require solving more than one root cause: A KE should thus continue until the problem is truly solved and the desired situation has been reached (Imai, 1997; Liker, 2004).

All the reviewed approaches emphasise that, to be most effective, the phases should be performed in this stipulated order. KE process quality, on which we will focus next, is assumed to depend on the ability and self-discipline of the group to perform each phase in an orderly, sequential manner (Bateman, 2005).

2.4 Kaizen event process quality

Most of the studies that examined factors related to KE process quality, utilised questionnaires, interviews or expert ratings of process quality. One notable exception was Hegedus and Rasmussen (1986), who content-analysed audiotapes of KE meetings. Systematic coding of field data from real-life KEs is more likely to offer added value than recall-type data that could be fraught with many kinds of perceptual biases (Kozlowski, 2015; Mathieu *et al.*, 2019). Relying only on answers to questionnaires and participants' interviews may lead to inaccuracies due to subjectivity (Baumeister *et al.*, 2007). Thus, studying and improving internal KE processes requires measuring those processes more directly and objectively.

One of the possible methodologies that has not been explored so far in the context of KE process quality is video-based observation (e.g. Christianson, 2018). Based on the numerous calls for more video-based analysis of real-life groups (Jones *et al.*, 2021; Kozlowski, 2015; Mathieu *et al.*, 2019), the present study pioneers the coding of the verbal contributions of group members during video-recorded real-life KEs and then mapping them graphically (Paoletti *et al.*, 2021). When mapped chronologically and per phase, sequential data points occur that can be connected by lines (similar to Kepner and Tregoe, 1965). These line graphs then represent how the KE group process flows through the phases.

We looked for a way to interpret the graphs as objective as possible. Goodness-of-fit indicators typically summarise the discrepancy between observed values and the expected values based on an ideal model. Based on the previously reviewed literature, we assumed that in an "ideal" KE process, a group flows through each consecutive phase without group members making remarks that do not belong to the phase they are in. We also assumed that the more discrepancies there are between this ideal KE process and an observed one, the lower the quality of the observed process. The *model fit theory* gives a framework for judging this discrepancy and thus the level of process quality (Meijer, 1994): by counting the number of deviations between the expected ideal situation and the observed one. Comparing the number and strength of the deviations from the ideal phased process (Meijer, 1994) typically leads to an array of possible indicators.

From a statistical point of view, like six sigma process analysis (Radhakrishnan and Balamurugan, 2010), we learned that when analysing process quality, larger deviations might have a stronger impact on quality than smaller deviations. Squaring the deviation value is a way to properly weigh those deviations (Radhakrishnan and Balamurugan, 2010).

Hence, we will explore how insights stemming from both model fit theory and process analysis may aid to yield more objective and additional insights into features of a high-quality KE process.

3. Method

3.1 Research design
To understand process quality better in group problem-solving, we adopted a mixed-methods research design (Johnson *et al.*, 2007). We scrutinised 21 highly diverse, real-life KE meetings that we had videotaped. First, we coded each verbal remark made by all the group members in each meeting, according to the six KE phases, and then we visualised each KE process through 21 phase-based line graphs. Secondly, in-depth one-on-one interviews were held with ten kaizen experts who were queried about these graphs. Also based on these graphs, they rated the KE process quality of each meeting. These interviews brought out the experts' vast tacit knowledge about group members' behaviours in KE meetings. In addition, following each first KE meeting, all the KE groups were offered a standardised 1 h training in structured group problem-solving and some feedback that had been created with the newly developed measure. Their group reactions were then captured. Furthermore, after the last KE meetings, during which each group reflected once more on their own entire line graph that we had fed back to them, we analysed their reflections.

Through coding and quantifying all the group members' phase-based coded remarks ($N = 5,442$), various quantitative indicators of the process quality of each meeting were formed. Guided by the model fit theory and process analysis, these indicators were calculated based on the number and weight of the problem-solving phase transitions within each KE group. The quantitative indicators were then matched to the experts' 21 process quality ratings: to obtain the best fitting, most informative indicator for the quality of a KE process. Below, the details are provided of each research step.

3.2 Sampling and sample description
Based on convenience sampling (Barratt *et al.*, 2011), we approached ten groups and asked them to videorecord their next KE meetings. The groups were embedded within two Dutch knowledge intensive organisations, a university and a management consultancy firm. To encourage participation, we promised them a uniform 1 h KE training after their first meeting. They all agreed to participate, as well as to two group interviews, conducted after their first and last meeting. The experience with KEs varied over the groups as well as over the individuals in the groups. In some groups, participants had a lean green belt certificate, in which they learned how to apply kaizen. In other groups a KE was a synonym for a group problem-solving session while other groups were familiar with lean; they were all eager to learn how to better apply the KE approach to solve a problem. This was useful for our study since we wanted to chart a great variety of naturally occurring processes. Consequently, all the groups were observed and videotaped multiple times.

Table 3 offers insights into each group's institutional embedding; size; composition in terms of gender and member's work experience; the number of videotaped KE meetings per group; the length of these meetings; and if a group was working on the same or a different problem during the meetings. Prior to the first KE meeting, each group had self-selected a problem, with most of the groups not finishing a whole KE during the first meeting. The group performed their KE on their own, without a trained facilitator and with only a video camera.

After each of the recorded KE meetings, the videotape was transcribed and minutely coded by one of the authors: a kaizen trainer/facilitator for over 18 years. Each group

Organisation	No. of group members (male/female)	Group members' work experience in years	KE meetings ([Figure 1](#))	Duration of each KE meeting in minutes	Central problem (selected by the group themselves) "How can we…"
1. University service department	4 (2/2)	<1	Meeting 1	1: 55	Improve the knowledge transfer towards new team members?
			Meeting 2	2: 60	
2. University service department	4 (0/4)	9	Meeting 3	1: 45	Improve the assignment of appropriate meeting rooms?
			Meeting 4	2: 50	
3. University service department	4 (0/4)	23	Meeting 5	1: 65	Obtain a better balance between types of users of our meeting rooms
			Meeting 6	2: 50	
4. Consultancy	5 (3/2)	3	Meeting 7	1: 35	Meeting 7: Better onboard new team members?
			Meeting 8	2: 45	Meeting 8: Improve our project acquisition skills?
5. Consultancy	7 (5/2)	3.5	Meeting 9	1: 45	Meeting 9: Improve the quality of our weekly stand-up?
			Meeting 10	2:40	Meeting 10: Improve weekly stand-up meeting attendance?
6. Consultancy	5 (4/1)	3	Meeting 11	1: 25	Meeting 11: Better attract new employees?
			Meeting 12	2: 40	Meeting 12: Realise a smooth off-boarding process?
7. Consultancy	4 (2/2)	4	Meeting 13	1: 35	Meeting 13: Better align our projects to our team goals?
			Meeting 14	2: 50	Meeting 14: Increase the number of project leads?
8. University service department	4 (3/1)	28	Meeting 15	1: 65	Meeting 15: Improve the quality of testing new software?
			Meeting 16	2: 55	Meeting 16: Improve compliance to the IT functional management standard model?

Table 3.
Kaizen event groups' characteristics

Organisation	No. of group members (male/female)	Group members' work experience in years	KE meetings (Figure 1)	Duration of each KE meeting in minutes	Central problem (selected by the group themselves) "How can we…"
9. University service department	5 (2/3)	27	Meeting 17	1: 50	Meeting 17: Improve the English speaking skills of our staff members?
			Meeting 18	2: 55	Meeting 18: Avoid losing focus on important topics?
10. University faculty members	6 (0/6)	17	Meeting 19	1: 45	Better align our student grading among teachers of the same course?
			Meeting 20	2: 50	
			Meeting 21	3: 45	
Total	48 (27/21)				

member's verbal remark was categorised into one of the six literature-based KE phases (Table 2); the codebook is in Appendix A. For instance, all remarks related to exploring the problem were categorised into the problem definition phase. All the ideas to solve the problem were categorised into the generate ideas phase and so on. These categorised verbal remarks were then plotted chronologically, and their consecutive points were connected, leading to line graphs of each of the 21 KE meetings which resembled so-called "group state space grids". The process "allows for visualization and quantification of team states for a moment-to-moment basis by tracking how a system changes on two categorical variables" (Paoletti *et al.*, 2021, p. 14). Each graph had the six KE phases on its Y-axis, with each group member's remark placed chronologically on the X-axis (Figure 1).

Ten kaizen experts, who were unrelated to the groups, were individually interviewed about high-quality KE process characteristics. When selecting these experts, we followed a homogeneous purposive sampling method (Etikan *et al.*, 2016) using three criteria: (1) at least five years of practical experience with kaizen or structured problem-solving; (2) vast experience in multiple KE roles (e.g. participant, facilitator, trainer or sponsor) and thus knowledgeable about the great variety of group dynamics during KEs; and (3) experience with kaizen in different sectors, to be able to evaluate problem-solving process differences across contexts. We recruited the experts through the independent networks of two of the authors, who themselves are experienced lean consultants with over 25 cumulative years of lean and kaizen experience in the Netherlands: 11 experts were invited, and ten of them participated. Table 4 describes their characteristics.

3.3 Data collection
3.3.1 In-depth interviews with the kaizen experts. The purpose of the interviews with the kaizen experts was dual. We wanted to tap their expertise on high KE process quality, and we felt that their reactions to concrete visualised processes, in the form of line graphs, would elicit better responses than merely asking them abstract questions. Hence, we invited them to react to the video-based line graphs. In addition, we asked them to evaluate the process quality of each of the 21 coded meetings, the new phase-based indicator of KE process quality, which we will present shortly.
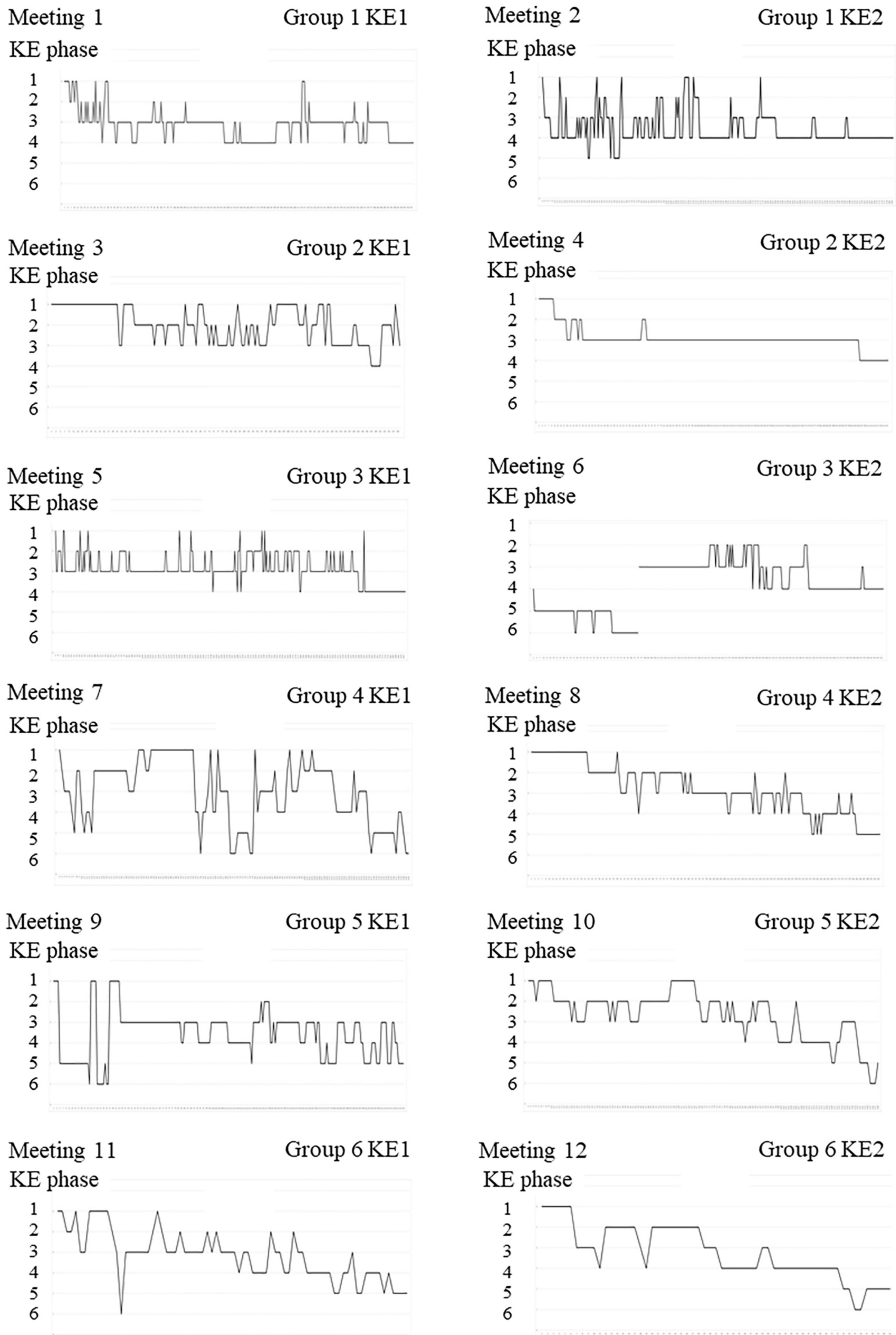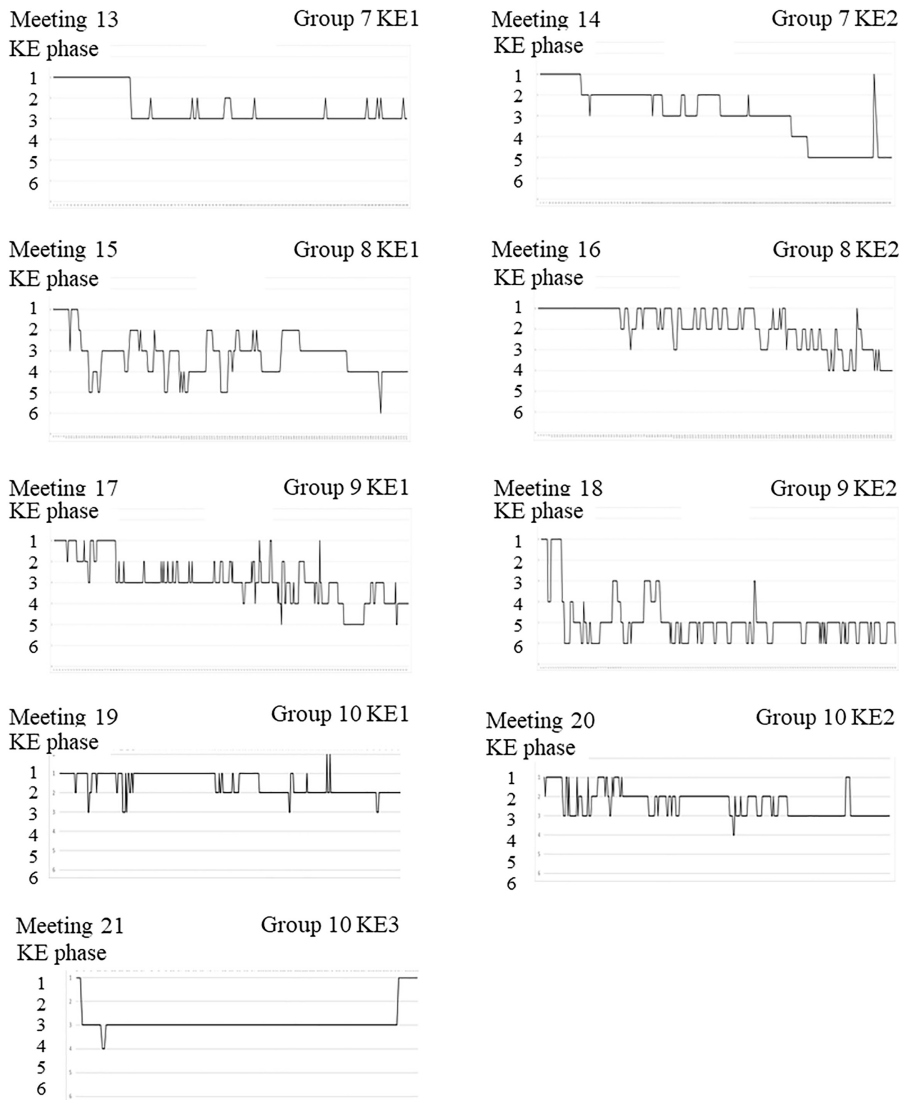
972

Meeting 1
KE phase
Group 1 KE1

Meeting 2
KE phase
Group 1 KE2

Meeting 3
KE phase
Group 2 KE1

Meeting 4
KE phase
Group 2 KE2

Meeting 5
KE phase
Group 3 KE1

Meeting 6
KE phase
Group 3 KE2

Meeting 7
KE phase
Group 4 KE1

Meeting 8
KE phase
Group 4 KE2

Meeting 9
KE phase
Group 5 KE1

Meeting 10
KE phase
Group 5 KE2

Meeting 11
KE phase
Group 6 KE1

Meeting 12
KE phase
Group 6 KE2

Figure 1.
Graphical
representation of all
videotaped Kaizen
Event meetings

Note(s): The numbers on the y-axes refer to the Kaizen Event (KE) phases. 1 = Problem
definition, 2 = Root-cause analysis; 3 = Generate ideas; 4 = Plan implementation;
5 = Implement; 6 = Check and sustain

Figure 1.

Each audiotaped and transcribed interview lasted about 1.5 h and was structured as
follows: After an introduction, we asked for answers to two open questions: *"How would you
describe a kaizen event?"* and *"If you teach or train people in kaizen, what do you emphasise?"*
Then, we solicited the kaizen experts' reactions to the four pre-selected example line graphs,
to obtain their inferences about the characteristics and quality of each KE process.
Together, these four graphs maximised the meetings' process variety; two of the four

graphs showed the phases clearly (see, Figure 1, meeting 8 and 20) while the other two graphs fluctuated between phases or skipped some of them (see, Figure 1, meeting 7 and 18). Whilst looking at each of the four graphs, we asked the experts to think aloud (Van Someren *et al.*, 1994) about the likely KE process that each of the graphs reflected. Next, we asked the experts to react to these four process graphs, plus we asked them: *"To what extent do you expect this group will end up with a valuable solution to their problem?"* At the very end of each interview, we asked the experts to judge the KE process quality of each of the 21 graphs on a scale from 1 = very bad to 10 = excellent and then queried them about the potential value of the line graphs.

3.3.2 An observational measure of KE process quality. To build a non-perceptual measure of high KE process quality, we turned to the model fit theory (Meijer, 1994) and statistical process analysis (Radhakrishnan and Balamurugan, 2010) to compare the actual versus the ideal KE process. We calculated various possible process quality indicators based on the observed deviations from the ideal six-phased structure during the 21 KE meetings (Meijer, 1994). Each observed deviation from the ideal process sequence is called here a "jump" between phases. Hence, the number of jumps in a videotaped KE meeting can be counted. Also, for each jump, the jump value can then be calculated, being the number of phase transitions that are made in each jump. The "total jump-value" of a KE meeting is the sum of all single jump values. Since higher jump values (e.g. making a remark that jumps back or forth by more than one phase) were deemed to be more disruptive, the squared jump value was also calculated to give those jumps a higher weight in the calculated total squared jump value (Radhakrishnan and Balamurugan, 2010). The total squared jump value is the sum of all squared jump values. We also calculated the same indicators vis-à-vis the total number of

Current position	Gender	Kaizen experience (years)	Current sector	Sectors in which an expert had applied kaizen
1. Continuous improvement process manager	F	5	Higher education	Higher education, educational logistics
2. Lean programme manager	M	9	Healthcare	Education, healthcare, IT
3. Lean and agile consultant	F	14	Consultancy	Production, sales and marketing, research and development, finance
4. Operational excellence leader	M	14	Production	Energy, production
5. Circular economy consultant	F	15	Consultancy	Production (food, fast-moving consumer goods, technical components), energy, services
6. Lean and agile team coach	M	15	Consultancy	Services, finance
7. Director of supply chain and operations manager	M	20	Production	Production, fast-moving consumer goods, research and development, sales, and marketing
8. Change advisor	M	20	Consultancy	Production, energy, services
9. Departmental programme lead infrastructure	M	25	Semi-government	Production, services
10. Continuous improvement programme manager	F	30	Higher education	Chemical engineering, production, higher education

remarks in each KE meeting (Cohen *et al.*, 2011), leading to the following six possible process quality indicators:

(1) total number of jumps between phases within a KE meeting

$$Count(jump)$$

(2) total jump values of a KE meeting

$$\sum abs(jump - value)$$

(3) total squared jump value of a KE meeting

$$\sum (abs(jump - value))^2$$

(4) total number of jumps of a KE meeting divided by the total number of remarks

$$\frac{total\ number\ of\ jumps}{total\ number\ of\ verbal\ contributions}$$

(5) total jump value of a KE meeting divided by the number of remarks

$$\frac{total\ jumps - value}{total\ number\ of\ verbal\ contributions}$$

(6) total squared jump value of a KE meeting divided by the number of remarks

$$\frac{total\ squared\ jumps - value}{total\ number\ of\ verbal\ contributions}$$

3.3.3 KE group members' perceptions. To collect the KE group members' perceptions of the quality of their group's recent meeting, three sources were used. First, we held group interviews after each group's first meeting, during which they had received a 1 h KE process training (the six KE phases of structured group problem-solving were explained), after which members could react to the line graphs of their own first KE meeting. We made field notes during those interviews (to capture the group members' highly diverse reactions). Secondly, we made transcripts of the second KE meetings in which the group members made remarks regarding the problem-solving phases they were trained in. Thirdly, during 1-h group feedback sessions (Shute, 2008), held after the last recorded meeting, the first author presented each group with their KE meetings' line graphs, after which the group could reflect freely on their problem-solving process. Field notes were taken by the first author.

For explorative purposes, we asked all the members of one of the ten KE groups to rate their impression of the effectiveness of the KE: on a seven-point Likert scale, based on four Van den Bossche *et al.* (2006) items (e.g. "*I am satisfied with the performance of this group*"). Although this survey measure was not a systematic part of our original data collection, it was used alongside the qualitative group member perceptions to triangulate the expert-based and quantitative indicators of KE process quality.

3.4 Data analysis
All the expert interview transcriptions were content-analysed using an inductive approach (Paoletti *et al.*, 2021). First, process-quality-related quotes were selected. Secondly, all the experts' remarks were categorised according to the following themes (Grodal *et al.*, 2020): the

KE's goal; the KE's structure; KE phase specific; and members' behaviours in the KEs. Then, process quality characteristics were determined from the experts' feedback.

After the experts had rated the process quality of each of the 21 meetings, their inter-rater reliability was calculated. The r_{WG} was 0.94, which means that the experts were almost unanimous in their judgement of the process quality of the presented graphs (Lindell *et al.*, 1999). Then, we calculated the six possible KE process quality indicators for each meeting. Second, the correlation between the average expert ratings and all the possible KE process quality indicators was computed. Finally, the expert process quality ratings of the 21 meetings were compared against the quantitative indicator with the highest correlation.

The available transcripts of the KE meetings and the field notes of the group interviews were content-analysed in terms of their evaluations of the quality of their own problem-solving process. Illustrative quotes were selected and used to offer insight into group members' views on their recent KE process quality.

4. Results
Below we first report the characteristics of high-quality KEs, based on the views of the kaizen experts. We then present and explore the possible quantitative indicators of KE process quality. Moreover, we report our analysis of the KE group members' diverse reactions to their own KE process quality.

4.1 Experts' evaluations of KE process quality
The experts converged on various characteristics of a high-quality KE (Table 5). Regarding the goals of a KE, the experts noted that kaizen must contribute to (organisational) learning by facilitating a dialogue aimed at finding and implementing sustainable solutions for a persistent problem. As highlighted by one of them: *"The worst thing that can happen (. . .) is that group members have not learnt anything from the process"*. All the experts also stressed that KE process quality hinges on adherence to the six literature-prescribed phases. One expert reacted: *"I can see a lot of jumping between all phases. This might make it difficult to follow, which reduces effectiveness"* and *"When they skip the entire root-cause phase, I do not*

Category	Characteristics
Goal	Kaizen is about creating an environment to learn
	Kaizen facilitates the needed dialogue to achieve a working solution
	Kaizen supports sustainable solutions to prevent the problem from reoccurring
Structure	Following the phases in an consecutive manner will result in an effective solution
	A sequentially-phased approach should be recognisable
	The KE group should have consensus about the result of a phase, before moving to the next phase
	Iterations between phases occur and will enrich the shared understanding, as long as the group members are aware as to which phase they are currently in
Phase specific	Without a shared problem definition, the KE will not end up with effective solutions
	Without a clear root-cause analysis, the KE will not end up with effective solutions
	Most effective KE groups spend over 60% of their effort in the first two phases, to really come to a clear understanding of the problem and the root causes
	Implementation is about learning, the faster a solution is implemented, the faster you will learn about its impact
Members' behaviour	Group members need discussion to arrive at a shared perspective
	Group members need to dare to share problems
	Group members need to be able to discuss task-focussed conflicts

Table 5.
Characteristics of high-quality KE processes according to the experts

expect them to come to a sustainable solution". However, nine experts explicitly mentioned the need for exploration, sharing perspectives, as long as "*you still know which phase you are in*". They noted that to share and explore perspectives, some degree of jumping between phases is unavoidable in practice and can even be valuable. They stressed the importance of the group knowing, at any point in time, in which problem-solving phase they are in, and the need for any KE group to reflect on the completeness of the discussion and group consensus on the results of each phase, before moving to the next phase. For example, one expert noted: "*Iterations can occur during the phases; that is no problem. The interesting thing is how they deal with it*". Another expert mentioned "*sometimes some 'wandering' between phases is needed, as long as the group rediscovers the right track*". For instance, checking at the end of each phase whether the group is still focussing on the right problem is a strength in a KE. Reaching group consensus in each phase before proceeding to the next sequential phase of problem-solving was commended. To quote one of the experts: "*I am actually looking for some kind of stairs*". Moreover, the quality of the result of each phase depends on the quality of the previous phases, for example, "*With limited problem definition, and limited root-cause analysis, you can only expect limited ideas*". Thus, it is no wonder that all the experts also stated the importance of the first two phases, that is, problem definition and root-cause analysis. They argued that the most effective KE groups spent over "60%" of their effort on those two phases: "*If the root-cause phase is executed thoroughly, solutions will be found easily*". Finally, the consensus among the experts was that an open discussion about the problem must take place during a high-quality KE, whereby participants should feel free to share their perspectives, even if they conflict with others' ideas. The experts felt that it is crucial to create the conditions for such discussions.

The experts' responses to the four exemplary graphs were quite unanimous. For instance, all of them thought that meeting 7 (Table 6) was chaotic, whereas meetings 8 and 11 followed an orderly path. In meeting 18, the group skipped many KE phases; four experts thought that this group process could not be regarded as being a KE. Those responses are also reflected in the experts' expectations whether the meetings would lead to a valuable solution (see Table 6): meeting 7 was rated 2.67 (Sd = 0.47) whereas meeting 8 was rated 6.00 (SD = 0.82). Together with the two other means of Table 6, we interpreted this as the anecdotal evidence of a link between high-quality process and better outcomes.

In addition, some experts noted that discussing the graphs made them more aware of their own views on high-quality KE processes: it "*gives me the opportunity to enrich my kaizen trainings*". Those comments sparked the idea to start using the graphs for educational purposes and because each of the ten KE groups was very eager to see their own graphs.

4.2 Comparison of the experts' perceived KE process quality and quantitative indicators
Table 7 presents the correlations between the expert rating of perceived KE process quality and the six possible quantitative quality indicators. A high correlation was expected between all the possible indicators and the average expert rating, as all the indicators reflect the KE process quality to some extent, but the highest appeared with the "total squared jump-value" indicator ($r = 0.62$; $p < 0.01$). In other words, KE meetings in which participants' verbal remarks make fewer jumps between the six KE phases may be of higher quality than meetings in which group members' verbal contributions *often* jump from one phase to a distant past or distant next phase.

Table 8 shows that the total squared jump value either declined or remained almost the same for six out of the ten groups (no. 3–8), while the average expert-based KE process quality rating increased. For groups 1, 9 and 10, the total squared jump value increased, while the average expert-based KE process quality rating declined. This finding aligns with the idea that during a high-quality KE meeting, a clear sequential use of the phases is desirable.

Table 6.
Experts' reactions to the four example graphs, based on the think aloud method

Example graphs	Experts' quotes	Mean experts' rating of expected valuable solution (SD)
Meeting 20 KE phase PD RA GI IP I CS	"They took their time to get to the PD, spent time on root causes. That is good." "They start with PD, and then jump about, which is good although I would have expected them to take more time. It seems a bit restless." "I think they are doing quite well. You see them jumping but I think this is their way of exploring the problem." "It seems they get a bit stuck in the PD phase. They're doing quite well on the phases." "In the first phase, they mixed PD and GI up, a bit like playing ping pong, which might be a way of exploring the problem."	5.22 (1.03)
Meeting 7 KE phase PD RA GI IP I CS	"They discuss more phases, but over time it becomes chaotic, looks like lack of consensus and decision making, so probably a confusing meeting." "Very interesting, as if they are going nowhere. I don't expect good outcomes." "Seems very chaotic, goes everywhere, and they discuss the PD rather late" "They know the solution before the problem; a complex group, seems they have some collaboration issues, and no effective goal." "It seems they have a solution and now they are looking for a problem, you can call it iterative, but I think it is just going everywhere."	2.67 (0.47)
Meeting 8 KE phase PD RA GI IP I CS	"They finish something before they continue, PD is fast, RA takes longer, that's what you want to see, the problem seems clear." "They don't go back to PD: it seems there is consensus on that, that is positive. Looks structured." "They really start with the PD discussion, rather step by step, looks good." "Rather structured, they seem to recap before moving to the next phase, that is good." "Seems structured, not wobbly. If they continue this way, I expect they will learn how to attain a working countermeasure."	6.00 (0.82)
Meeting 18 KE phase PD RA GI IP I CS	"They start with experiments and this will lead, at some point, to a working solution, but it is not a kaizen process." "It seems like they are firefighting, this is not about doing kaizen." "This looks like non-structured problem solving; will they be lucky and find a solution?" "Oops, interesting group dynamics but not kaizen." "Low expectation; they jump to solutions. It looks like the kata approach. I really can't see the root cause analysis and the 'stairs'."	1.89 (0.74)

Note(s): The abbreviations in the left column stand for: KE = Kaizen event; PD = Problem Definition; RA = Root-cause Analysis; GI = Generate Ideas; PI = Plan Implementation; I = implement; CS = Check and Sustain. SD = Standard deviation. The experts rated the example graphs on a 7-point Likert scale, 1 = very poor, 7 = very strong

However, a very strict use of the sequential phases was not found either. For example, while group 2's total squared jump value starkly declined after their first meeting, their average expert-based KE process quality rating also declined. We reported earlier that some iterations among the phases might enrich the problem-solving discussion: after all, problem-solving often requires learning and some creativity as well, which does not necessarily augur well with enforcing the sequential order of the phases very strictly.

4.3 Group self-evaluations and the new observational measure
The KE meeting transcripts, training and feedback pointed to how the groups reflected differentially on the quality of their process. Some groups and/or members took the feedback for granted, while others showed an eagerness to use it in their next meeting. For instance, a member of KE group 7 noted: *"In the second meeting we really came up with some other ideas because we talked about root causes, which we had never discussed before"* (Figure 1, meeting 14). A member of group 8 mentioned: "*This feedback helped us to first explore if we really have a problem, before jumping to solutions. We found out that having a problem only in one specific area makes it a lot easier to solve now*" (Figure 1, meeting 16). At the end of the second KE meeting, both groups 7 and 8 acknowledged that the method was valuable for them (Group 7: *"This was useful, thanks trainer"* and group 8: *"Having such a structured approach makes it much easier"*). Multiple members of group 10 tried to force the group to stick to the KE phases; one member reflected: "*In the second KE meeting we started arguing again about the problem definition, and we did not continue before we had consensus. It felt uncomfortable, but I realise*

Phase-based quantitative indicators	Mean expert ratings of perceived KEPQ	p
1. Number of jumps	-0.41^*	0.066
2. Total jump value	-0.52^*	0.015
3. Total squared jump value	-0.62^{**}	0.003
4. Ratio number of jumps	-0.13	0.567
5. Ratio total jump value	-0.38	0.286
6. Ratio total squared jump value	$-0.37^{\dagger}$	0.095
Note(s): $\dagger p < 0.10$; $^*p < 0.05$; $^{**}p < 0.01$		

Table 7.
Paired samples
correlation

Group no	Total squared jump value			Average expert KE process quality rating (SD)		
	Meeting 1	Meeting 2	Δ	Meeting 1	Meeting 2	Δ
1	85	185	100	4.56 (0.96)	3.89 (0.87)	−0.66
2	80	9	−81	7.33 (1.05)	6.83 (0.67)	−0.50
3	110	41	−69	5.61 (1.20)	6.39 (1.56)	0.78
4	167	44	−123	4.89 (1.10)	7.67 (1.05)	2.78
5	152	40	−112	4.22 (0.79)	7.28 (0.79)	3.06
6	66	22	−44	6.39 (1.20)	7.78 (1.00)	1.39
7	24	30	6	4.67 (0.82)	7.50 (1.00)	2.87
8	77	77	0	5.00 (1.33)	5.78 (1.13)	0.78
9	53	96	43	5.72 (1.13)	3.44 (1.50)	−2.28
10	65	86	21	8.00 (1.08)	7.39 (1.15)	−0.61

Note(s): Average expert KE process quality rating scale 1 = very bad to 10 = excellent. SD = Standard deviation. Δ = difference between meeting 1 and meeting 2. For comparison reasons video 21, being the third meeting of the group 10, was left out

Table 8.
Differences in KE
meeting process
quality before and after
training

that, after we agreed, the meeting became more effective as it was clear to everyone for which cause ideas had to be shared" (Figure 1, meeting 20). Paying attention to each of the KE phases, without skipping phases but allowing some jumping when required, was thus felt to be related to a higher-quality KE process. This lends support to the key assumptions underlying the total squared jump value as a new, quantified measure of KE process quality.

As an additional robustness check, we used one of the recorded KEs to illustrate the utility of the identified total squared jump-value indicator of KE process quality vis-à-vis group self-evaluation: meeting 21 in Figure 1. The total squared jump value (10) was very low in comparison to all the other videos, which matched the experts' rating of the process quality of this meeting (7.61 on a ten-point scale). We had only asked the members of this group to rate their meeting's effectiveness, and it was high as well (5.38 on a seven-point scale). Indeed, the group members mentioned the relative quietness of the meeting and their non-hectic discussions. This anecdotal evidence shows that the phase-based total squared jump-value measure of process quality is commensurate with both the group's self-rated meeting effectiveness and the experts' judgement of this group's process quality.

5. Discussion
Effective group problem-solving is part and parcel of becoming a learning organisation (Bateman, 2005; Hines *et al.*, 2004; Netland, 2016). Despite the abundant literature on learning, for group problem-solving in general, and specifically KEs, hardly any other study has disentangled what temporary multidisciplinary groups must do during KE meetings to effectuate high tool quality (Aleu and van Aken, 2016). Using mixed methods, including a new video-observation method in which real-life KE data was coded, we explored the characteristics of high-quality KE meetings. Based on the literature, we first identified six prototypical phases of a KE: (1) problem definition; (2) root-cause analysis; (3) generate ideas; (4) plan implementation; (5) implement; and (6) check and sustain. Every single verbal contribution by each group member during 21 diverse KE meetings was then coded into one of these phases. Next, ten kaizen experts and all the 48 KE group members reflected on the resulting visual, phase-based state space grids. According to the experts, KE phases should be enacted sequentially with explicit group consensus at the end of each phase. Yet, they suggested that small deviations of the phased dialogue might be necessary for a high-quality process which was corroborated by our quantitative data as well as by the participants' own reflections. The various findings reported herein have a number of theoretical implications which could be used to refine the 2009 edition of Farris *et al.*'s model of critical KE success factors (see, Figure 2).

Firstly, our study challenges the assumption that effective problem-solving needs groups to adhere very strictly to a phased approach (Mohaghegh and Furlan, 2020; Woods, 2000). KE scholars have long assumed that desirable outcomes only result from group members' close adherence to the tool's sequential phases (Choo, 2014; De Mast, 2011; Imai, 1997; Liker, 2004; Tucker *et al.*, 2002). This dictum is based on the idea that structured group problem-solving is most effective when facing ill-structured problems (Mohaghegh and Furlan, 2020). When a KE is set up to solve a persistent problem, the outcome should contribute to solving that problem (Farris *et al.*, 2009). But a KE is also a place for employees to share knowledge and learn (Bateman, 2005; Bessant *et al.*, 2001; Galeazzo and Furlan, 2019), leading to social outcomes such as developing problem-solving skills and growing a more thorough understanding of continuous improvement (Farris *et al.*, 2009). Having such a "place to learn" is essential for employees' acceptance of continuous improvement (Bateman, 2005; Netland, 2016). Collective learning is enabled by reflective talk among participants, on the task itself as well as on the group problem-solving process (Dittrich *et al.*, 2016; Furlan *et al.*, 2019). Given that a KE group consists of individuals who are united temporarily to solve a persisting
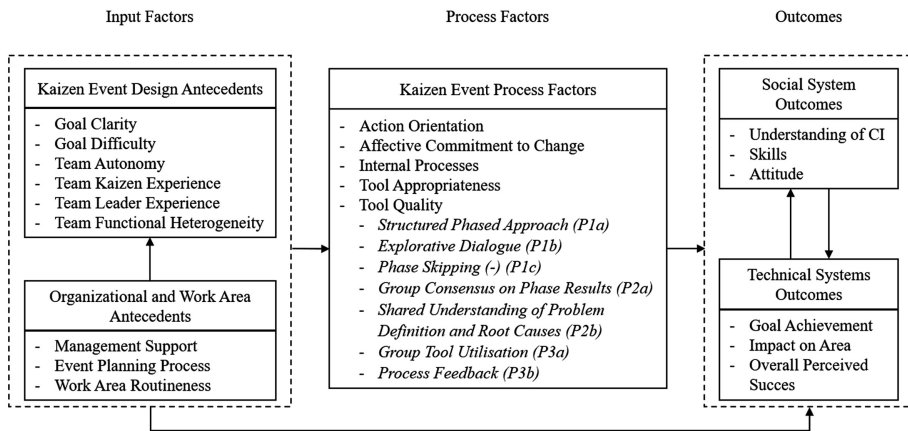
Note(s): The text in italics has been added to the original model of Farris *et al.* (2009, p. 46). *P* = proposition

problem they are experiencing themselves (Bortolotti *et al.*, 2018), the importance of an open group dialogue is crucial (Dittrich *et al.*, 2016; Hagemann and Kluge, 2017; Kolb *et al.*, 2008; LaFasto and Larson, 2001). This implies that, from time to time, a KE group must allow members to breakthrough an overly rigid use of the problem-solving phases. However, we found that when the participants break out of their current phase too often, or even skip phases, KE process quality declines. This outcome is reflected in our new measure where a high value for the total squared jump indicator resembles low KE process quality. Hence, based on our results, we want to nuance any overly strict adherence to the sequential order of the phases during meetings, even though sticking to its long-prescribed sequential order is found here to be part of high-quality KEs. The process quality of KE meetings may thus be higher if the members are permitted occasionally to deviate slightly from the prescribed sequential order, expressly to enrich the explorative quality of their dialogue. Hence:

Proposition 1. KE process quality is likely to be high when KE groups follow the prototypical sequential phases (a), whereby explorative dialogues to adjacent phases may occur (b), whereas KE process quality is likely to be low when KE groups skip phases (c).

Secondly, KE groups must reach consensus about the result of each phase before continuing to the next phase (Andersson *et al.*, 2006). Likewise, Taggar and Brown (2001) found that focussing on the task at hand, by reducing the number of off-topic discussions, may contribute to group problem-solving. Group decision-making literature already pointed to the importance of achieving group consensus to be effective (Pérez *et al.*, 2018; Van Ginkel and Van Knippenberg, 2008). Group consensus at the end of each KE phase is achieved by openly sharing perspectives (Galeazzo and Furlan, 2019; Pérez *et al.*, 2018). Thus, a KE group should, based on an open dialogue, reach phase results everyone can support and avoid jumping overly among the phases, for instance, by questioning the problem definition again when the group has progressed into the brainstorming phase. Such jumping may be perceived by some group members as setbacks, which are seen to affect group dynamics negatively, including creativity (Cohen *et al.*, 2011) and motivation (Marks *et al.*, 2001). Group consensus seems especially important in the problem definition and root-cause analysis phases. Previous studies showed that a thorough problem analysis and root-cause analysis should be performed to avoid solutions that solve only symptoms instead of the underlying problem

(Choo, 2014; Kepner and Tregoe, 1965; Liker, 2004; Mohaghegh and Furlan, 2020; Tucker *et al.*, 2002). Similarly, Taggar and Brown (2001) showed that group member participation in setting team goals is essential for achieving them. So far, evidence is lacking for the relative weight of each phase, including the idea voiced in the popular kaizen literature that effective problem definition and root-cause analysis may take up at least 60% of the total time of a KE. Thus, we propose:

Proposition 2. A high-quality KE process is more likely to occur when group members decide consensually on the result of each phase before the group moves explicitly to its next phase (a), especially during the first two phases of KE: problem definition and root-cause analysis (b).

Becoming a learning organisation entails the practice of effective problem-solving (Argote and Hora, 2017; Furlan *et al.*, 2019; Hines *et al.*, 2004; Yeo and Marquardt, 2010). Since failing to solve a problem might demotivate employees from active participation in further improvements (Bessant *et al.*, 2001; Jurburg *et al.*, 2017), an adequate level of group problem-solving skills is needed to avoid such failure (Bessant *et al.*, 2001; Ju *et al.*, 2021; Vo *et al.*, 2019). Even though organisations often invest in external facilitators in the early stages of adopting continuous improvement, to safeguard an effective group problem-solving process (Bateman, 2005), the group members will still need to develop their own skills (Netland, 2016). Hence, before starting to solve a problem in a KE, the group should have an adequate level of problem-solving skills (Kolb *et al.*, 2008; LaFasto and Larson, 2001). Related to Farris *et al.*'s model (2009), the better groups are able to apply the KE tool during a KE meeting (i.e. "tool quality"), the higher the quality of their process. Applying the KE tool well requires showing group problem-solving skills (see Figure 2).

Acquiring these skills is thus especially important for employees of organisations that aim to operate as full learning organisations (Jadhav *et al.*, 2014; Lizarelli and Alliprandini, 2020; Netland, 2016; Powell and Coughlan, 2020). Other than the imparting of various abstract group problem-solving approaches, employees are more likely to learn by doing (Lizarelli and Alliprandini, 2020; Van Dun and Wilderom, 2021), particularly if feedback is available (Hassell and Cotton, 2017; Knapp, 2010). Deep employee reflections on past (group) processes and their outcomes can often be enriched through visualisations (Aoki, 2019; Hassell and Cotton, 2017). While a KE provides people with a natural setting for knowledge sharing and learning, reflecting on the event's group dynamics has not been a standard item on kaizen's agendas. Yet, according to Cronin *et al.* (2011), it is important to give a group opportunities to reflect on their recent dynamics and thus on their past process quality. Many members of the ten groups we studied were highly interested in the visualised feedback we offered with the line graphs depicting the course of their recent meetings, although not all the groups were equally motivated and/or capable to put this learning immediately into action during their next meeting. Visual process feedback may contribute to honing the needed sophistication in KEs. Hence:

Proposition 3. A high-quality KE process is more likely to occur when groups are able to utilise their KE tool skills (a), which can be improved by visually displayed process feedback (b).

The observational method we used can be seen as an invaluable by-product, both for future research on group problem-solving and for individual and group KE training. In fact, with the observational data as input, we developed a new non-subjective measure of the process quality of KE meetings that have often been noted to underperform, despite the wealth of favourable input factors laid out particularly well by Farris *et al.* (2009). In our efforts to explicate factors that may enrich the process of KEs, one of the derived possible indicators of KE meeting quality was found to be very informative, also in the eyes of both the KE experts and the KE group members. This new indicator, entitled "total squared jump-value", uniquely

combines people's content and process types of remarks during KE meetings which makes it not only practically relevant but, at the same time, a baseline for further refinement as well as other academic KE research.

6. Practical implications
First and foremost, this study has unearthed various indicators of high-quality KEs. Based on our analysis of various data sources, KE groups are advised to follow a structured, phased approach of problem-solving; allow for an explorative dialogue; and ensure group consensus before moving to a next phase, especially in terms of problem definition and root causes. Surprisingly, our informants converged on the idea that the quality of the dialogue is sometimes more important than sticking strictly to the prescribed order of phases. Therefore, some degree of switching back and forth between phases must be allowed, for example, to quickly double-check whether a potential solution a group member has in mind may solve the root cause.

Secondly, the video-based, graphically visualised feedback can be offered to kaizen groups. The line graphs will not only assist the group members themselves but even KE facilitators to reflect on the features of an experienced high-quality KE process. Hence, the developed measure of KE process quality affords group members the chance to learn from previous experiences: by reflecting on how well they have behaved in a particular KE meeting. This is because the most prominent question we got from all the participating KE groups was: "*How did we do?*" Although our data could not provide them with the requested performance score yet, the graphs we provided gave the members practical insights into the process quality level of their just completed KE meeting.

People's motivation to participate in future KEs might be especially relevant for group members' self-efficacious behaviour during kaizen events. Many scholars have called for the building of effective methods for training groups to solve problems effectively (see, e.g. Jadhav *et al.*, 2014; Jurburg *et al.*, 2017; Netland, 2016). The graphical display of people's phase-based coded remarks can offer group members a mirror of, as well as a benchmark for, their KE experience. Our training mainly influenced group members who were motivated to learn about KEs. Hence, there is room for a solid design and test of group-training interventions to enrich KE processes and their outcomes. The "total squared jump-value" measure developed herein can then serve as a future benchmark.

7. Conclusion
To understand the process features of high-quality KEs better, an observational measure was built that visualised and quantified the course of 21 highly diverse KE meetings. Soon after these meetings, the ten KE groups reflected on these visualisations. Moreover, ten individual kaizen experts were queried about the same visualisations; their quantified reactions were a first robustness check of the new measure's most informative indicator total squared jump value. Some of the experts' other reactions corresponded to the insights derived from the KE groups.

Given the exploratory nature of this research, several limitations must be noted. First, only ten kaizen experts were questioned, all from the Netherlands, and we phase-coded a relatively small number of videotaped KE meetings. And, although the KE groups we observed were recruited from knowledge-intense organisations, a wider range of organisations need to be studied. Analysing more KE meetings, in conjunction with various KE input and output indicators, will undoubtedly validate and refine both the process measure and findings further. To reduce potential single-observer effects, the phase coding could be carried out by multiple coders (Van Dun *et al.*, 2017; Van Dun and Wilderom, 2021). Furthermore, allowing some small deviations from the strict sequential process might help to refine a subsequent version of the quantitative indicator. To improve this phase-based measure further, for use in

practice as a KE benchmark, future studies should tape *all* KE meetings within whole KEs and consider the maturity level of KE groups. Another limitation might be that we did not ask *non*-kaizen group problem-solving experts to assess the graphs as well. Despite these imperfections, data saturation and robustness occurred, and we strongly encourage new field research that examines the effects of phase-based feedback to KE groups.

Now that we have begun to substantiate the process factors of Farris *et al.*'s (2009) input–output model, new pressing research questions have arisen such as: How do the process characteristics relate to the degree of sustainability of the problems being solved through KE? Previously, related questions were only addressed based on self-reports that are known to contain biases such as memory loss or lack of interest. The new observational measure enables a much closer and more objective examination of the KE process factors in conjunction with the potentially controllable input factors. In addition, we still know little about the relative importance of each KE phase for the desirable social and technical outcomes, or whether all phases are truly needed for effective problem-solving (Pretz, 2008). Furthermore, it would be interesting to explore how the use of standardised instruments such as 5-times-Why or Ishikawa impacts KE group members discipline to adhere to the phases. The relations between Farris *et al.*'s (2009) input factors and the relative quality of the KE process and outcome must also be studied; for instance, what is the impact of goal unclarity on a group's actual problem-solving process quality and its outcome?

The phase-coding method presented herein is based on individual contributions, so it may help to shed light on the optimal group composition (Kozlowski, 2015). Given that group problem-solving skills do not come naturally to all members (Bateman, 2005; Jadhav *et al.*, 2014; Netland, 2016; Vuculescu *et al.*, 2021) and that a 1 h KE-skill training was not found to greatly improve the process, it is time to craft and test different modes of sophisticated KE training on real-life groups (Oliva, 2019). Such interventions can benefit from visual representations for KE participants' learning and performance (Aoki, 2019) and may even be extended to similar studies of structured problem-solving approaches such as DMAIC and Quality Circles (De Mast and Lokkerbol, 2012; Rafaai *et al.*, 2018). KEs can go through the consecutive phases multiple times (Imai, 1997; Liker, 2004), and the groups might learn along the way. By acknowledging that previous group learnings can be input for next KEs (Ilgen *et al.*, 2005), existing models of KE effectiveness must be refined further as well.

In this study, we focussed on KEs as a practically and highly relevant example of multidisciplinary group problem-solving (Bortolotti *et al.*, 2018). Operations Management scholars have long depended only on perceptions of the quality of KEs. The herein presented KE process measure, consisting of a KE process visualisation and a quantitative indicator, may also function as a springboard towards examining more sophisticated group-based problem-solving.

References

Aleu, F.G. and van Aken, E.M. (2016), "Systematic literature review of critical success factors for continuous improvement projects", *International Journal of Lean Six Sigma*, Vol. 7 No. 3, pp. 214-232, doi: 10.1108/IJLSS-06-2015-0025.

Álvarez-García, J., Durán-Sánchez, A. and Del Río-Rama, M. (2018), "Systematic bibliometric analysis on kaizen in scientific journals", *The TQM Journal*, Vol. 30 No. 4, pp. 356-370, doi: 10.1108/TQM-12-2017-0171.

Andersson, R., Eriksson, H. and Torstensson, H. (2006), "Similarities and differences between TQM, six sigma and lean", *The TQM Magazine*, Vol. 18 No. 3, pp. 282-296, doi: 10.1108/09544780610660004.

Aoki, K. (2019), "The roles of material artifacts in managing the learning-performance paradox: the kaizen case", *Academy of Management Journal*, Vol. 63 No. 4, pp. 1266-1299, doi: 10.5465/amj.2017.0967.

Argote, L. and Hora, M. (2017), "Organizational learning and management of technology", *Production and Operations Management*, Vol. 26 No. 4, pp. 579-590, doi: 10.1111/poms.12667.

Barratt, M., Choi, T.Y. and Li, M. (2011), "Qualitative case studies in operations management: trends, research outcomes, and future research implications", *Journal of Operations Management*, Vol. 29 No. 4, pp. 329-342, doi: 10.1016/j.jom.2010.06.002.

Bateman, N. (2005), "Sustainability: the elusive element of process improvement", *International Journal of Operations & Production Management*, Vol. 25 No. 3, pp. 261-276, doi: 10.1108/01443570510581862.

Baumeister, R.F., Vohs, K.D. and Funder, D.C. (2007), "Psychology as the science of self-reports and finger movements: whatever happened to actual behavior?", *Perspectives on Psychological Science*, Vol. 2 No. 4, pp. 396-403, doi: 10.1111/j.1745-6916.2007.00051.x.

Bessant, J., Caffyn, S. and Gallagher, M. (2001), "An evolutionary model of continuous improvement behaviour", *Technovation*, Vol. 21 No. 2, pp. 67-77, doi: 10.1016/S0166-4972(00)00023-7.

Bortolotti, T., Boscari, S., Danese, P., Medina Suni, H.A., Rich, N. and Romano, P. (2018), "The social benefits of kaizen initiatives in healthcare: an empirical study", *International Journal of Operations & Production Management*, Vol. 38 No. 2, pp. 554-578, doi: 10.1108/IJOPM-02-2017-0085.

Byun, H., Lee, J. and Cerreto, F.A. (2014), "Relative effects of three questioning strategies in ill-structured, small group problem solving", *Instructional Science*, Vol. 42 No. 2, pp. 229-250, doi: 10.1007/s11251-013-9278-1.

Choo, A.S. (2014), "Defining problems fast and slow: the U-shaped effect of problem definition time on project duration", *Production and Operations Management*, Vol. 23 No. 8, pp. 1462-1479, doi: 10.1111/poms.12219.

Christianson, M.K. (2018), "Mapping the terrain: the use of video-based research in top-tier organizational journals", *Organizational Research Methods*, Vol. 21 No. 2, pp. 261-287, doi: 10.1177/1094428116663636.

Cohen, M.A., Rogelberg, S.G., Allen, J.A. and Luong, A. (2011), "Meeting design characteristics and attendee perceptions of staff/team meeting quality", *Group Dynamics: Theory, Research and Practice*, Vol. 15 No. 1, pp. 90-104, doi: 10.1037/a0021549.

Cronin, M.A., Weingart, L.R. and Todorova, G. (2011), "Dynamics in groups: are we there yet?", *Academy of Management Annals*, Vol. 5 No. 1, pp. 571-612, doi: 10.1080/19416520.2011.590297.

De Mast, J. (2011), "The tactical use of constraints and structure in diagnostic problem solving", *Omega*, Vol. 39 No. 6, pp. 702-709, doi: 10.1016/j.omega.2011.02.002.

De Mast, J. and Lokkerbol, J. (2012), "An analysis of the Six Sigma DMAIC method from the perspective of problem solving", *International Journal of Production Economics*, Vol. 139 No. 2, pp. 604-614, doi: 10.1016/j.ijpe.2012.05.035.

Dittrich, K., Guérard, S. and Seidl, D. (2016), "Talking about routines: the role of reflective talk in routine change", *Organization Science*, Vol. 27 No. 3, pp. 678-697, doi: 10.1287/orsc.2015.1024.

Etikan, I., Musa, S.A. and Alkassim, R.S. (2016), "Comparison of convenience sampling and purposive sampling", *American Journal of Theoretical and Applied Statistics*, Vol. 5 No. 1, pp. 1-4, doi: 10.11648/j.ajtas.20160501.11.

Farris, J.A., van Aken, E.M., Doolen, T.L. and Worley, J. (2009), "Critical success factors for human resource outcomes in kaizen events: an empirical study", *International Journal of Production Economics*, Vol. 117 No. 1, pp. 42-65, doi: 10.1016/j.ijpe.2008.08.051.

Fisher, S.G., Hunter, T.A. and Macrosson, W.D.K. (1997), "Team or group? Managers' perceptions of the differences", *Journal of Managerial Psychology*, Vol. 12 No. 4, pp. 232-242, doi: 10.1108/02683949710174838.

Furlan, A., Galeazzo, A. and Paggiaro, A. (2019), "Organizational and perceived learning in the workplace: a multilevel perspective on employees' problem solving", *Organization Science*, Vol. 30 No. 2, pp. 280-297, doi: 10.1287/orsc.2018.1274.

Galeazzo, A. and Furlan, A. (2019), "Good problem solvers? Leveraging knowledge sharing mechanisms and management support", *Journal of Knowledge Management*, Vol. 23 No. 6, pp. 1017-1038, doi: 10.1108/JKM-05-2018-0290.

Glover, W.J., Van Aken, E.M., Farris, J.A. and Liu, W.H. (2013), "Characteristics of established kaizen event programs: an empirical study", *International Journal of Operations and Production Management*, Vol. 33 No. 9, pp. 1166-1201, doi: 10.1108/IJOPM-03-2011-0119.

Glover, W.J., Farris, J.A. and Van Aken, E.M. (2014), "Kaizen events: assessing the existing literature and convergence of practices", *Engineering Management Journal*, Vol. 26 No. 1, pp. 39-61, doi: 10.1080/10429247.2014.11432003.

Grass, A., Backmann, J. and Hoegl, M. (2020), "From empowerment dynamics to team adaptability: exploring and conceptualizing the continuous agile team innovation process", *Journal of Product Innovation Management*, Vol. 37 No. 4, pp. 324-351, doi: 10.1111/jpim.12525.

Grodal, S., Anteby, M. and Holm, A.L. (2020), "Achieving rigor in qualitative analysis: the role of active categorization in theory building", *Academy of Management Review*. doi: 10.5465/amr.2018.0482.

Hackman, J.R. and Morris, C.G. (1975), "Group tasks, group interaction process, and group performance effectiveness", in Berkowitz, L. (Ed.), *Advances in Experimental Social Psychology*, Academic Press, London, pp. 45-99, doi: 10.1016/S0065-2601(08)60248-8.

Hagemann, V. and Kluge, A. (2017), "Complex problem solving in teams: the impact of collective orientation on team process demands", *Frontiers in PsychologySEP*, Vol. 8, doi: 10.3389/fpsyg.2017.01730.

Hassell, M.D. and Cotton, J.L. (2017), "Some things are better left unseen: toward more effective communication and team performance in video-mediated interactions", *Computers in Human Behavior*, Vol. 73, pp. 200-208, doi: 10.1016/j.chb.2017.03.039.

Hegedus, D.M. and Rasmussen, R.V. (1986), "Task effectiveness and interaction process of a modified nominal group technique in solving an evaluation problem", *Journal of Management*, Vol. 12 No. 4, pp. 545-560, doi: 10.1177/014920638601200409.

Hines, P., Holweg, M. and Rich, N. (2004), "Learning to evolve: a review of contemporary lean thinking", *International Journal of Operations and Production Management*, Vol. 24 No. 10, doi: 10.1108/01443570410558049.

Ilgen, D.R., Hollenbeck, J.R., Johnson, M. and Jundt, D. (2005), "Teams in organizations: from input-process-output models to IMOI models", *Annual Review of Psychology*, Vol. 56 No. 1, pp. 517-543, doi: 10.1146/annurev.psych.56.091103.070250.

Imai, M. (1997), *Gemba Kaizen: A Commonsense, Low-Cost Approach to Management*, The Kaizen Institute, Seattle.

Jadhav, J.R., Mantha, S.S. and Rane, S.B. (2014), "Exploring barriers in lean implementation", *International Journal of Lean Six Sigma*, Vol. 5 No. 2, pp. 122-148, doi: 10.1108/IJLSS-12-2012-0014.

Johnson, B.R. and D'Lauro, C.J. (2018), "After brainstorming, groups select an early generated idea as their best idea", *Small Group Research*, Vol. 49 No. 2, pp. 177-194, doi: 10.1177/1046496417720285.

Johnson, R.B., Onwuegbuzie, A.J. and Turner, L.A. (2007), "Toward a definition of mixed methods research", *Journal of Mixed Methods Research*, Vol. 1 No. 2, pp. 112-133, doi: 10.1177/1558689806298224.

Jones, C., Volet, S. and Pino-Pasternak, D. (2021), "Observational research in face-to-face small groupwork: capturing affect as socio-dynamic interpersonal phenomena", *Small Group Research*, Vol. 52 No. 3, pp. 341-376, doi: 10.1177/1046496420985920.

Ju, B., Lee, Y., Park, S. and Yoon, S.W. (2021), "A meta-analytic review of the relationship between learning organization and organizational performance and employee attitudes: using the dimensions of learning organization questionnaire", *Human Resource Development Review*, pp. 1-45, doi: 10.1177/1534484320987363.

Jurburg, D., Viles, E., Tanco, M. and Mateo, R. (2017), "What motivates employees to participate in continuous improvement activities?", *Total Quality Management and Business Excellence*, Vol. 28 Nos 13-14, pp. 1469-1488, doi: 10.1080/14783363.2016.1150170.

Kepner, C.H. and Tregoe, B.B. (1965), *The Rational Manager*, McGraw-Hill, New York, NY.

Knapp, R. (2010), "Collective (team) learning process models: a conceptual review", *Human Resource Development Review*, Vol. 9 No. 3, pp. 285-299, doi: 10.1177/1534484310371449.

Kolb, J.A., Jin, S. and Song, J.H. (2008), "A model of small group facilitator competencies", *Performance Improvement Quarterly*, Vol. 21 No. 2, pp. 119-133, doi: 10.1002/piq.20026.

Kozlowski, S.W.J. (2015), "Advancing research on team process dynamics: theoretical, methodological, and measurement considerations", *Organizational Psychology Review*, Vol. 5 No. 4, pp. 270-299, doi: 10.1177/2041386614533586.

LaFasto, F.M.J. and Larson, C.E. (2001), *When Teams Work Best: 6,000 Team Members and Leaders Tell what it Takes to Succeed*, Sage, Thousand Oaks.

Lam, M., O'Donnell, M. and Robertson, D. (2015), "Achieving employee commitment for continuous improvement initiatives", *International Journal of Operations and Production Management*, Vol. 35 No. 2, pp. 201-215, doi: 10.1108/IJOPM-03-2013-0134.

Latzko, W.J. and Saunders, D.M. (1996), *Four Days with Dr. Deming: A Strategy for Modern Methods of Management*, Addison-Wesley, Reading.

Liker, J.K. (2004), *The Toyota Way*, McGraw-Hill, New York, NY.

Lindell, M.K., Brandt, C.J. and Whitney, D.J. (1999), "A revised index of interrater agreement for multi-item ratings of a single target", *Applied Psychological Measurement*, Vol. 23 No. 2, pp. 127-135, doi: 10.1177/01466219922031257.

Liu, W.H., Asio, S., Cross, J., Glover, W.J. and Van Aken, E. (2015), "Understanding team mental models affecting kaizen event success", *Team Performance Management: An International Journal*, Vol. 21 Nos 7/8, pp. 361-385, doi: 10.1108/TPM-03-2015-0012.

Lizarelli, F.L. and Alliprandini, D.H. (2020), "Comparative analysis of Lean and Six Sigma improvement projects: performance, changes, investment, time and complexity", *Total Quality Management and Business Excellence*, Vol. 31 Nos 3-4, pp. 407-428, doi: 10.1080/14783363.2018.1428087.

Marks, M.A., Mathieu, J.E. and Zaccaro, S.J. (2001), "A temporally based framework and taxonomy of team processes", *Academy of Management Review*, Vol. 26 No. 3, p. 356, doi: 10.2307/259182.

Mathieu, J.E., Hollenbeck, J.R., van Knippenberg, D. and Ilgen, D.R. (2017), "A century of work teams in the Journal of Applied Psychology", *Journal of Applied Psychology*, Vol. 102 No. 3, pp. 452-467, doi: 10.1037/apl0000128.

Mathieu, J.E., Gallagher, P.T., Domingo, M.A. and Klock, E.A. (2019), "Embracing complexity: reviewing the past decade of team effectiveness research", *Annual Review of Organizational Psychology and Organizational Behavior*, Vol. 6 No. 1, pp. 17-46, doi: 10.1146/annurev-orgpsych-012218-015106.

McFadzean, E. (2002), "Developing and supporting creative problem-solving teams: Part 1 - a conceptual model", *Management Decision*, Vol. 40 No. 5, pp. 463-475, doi: 10.1108/00251740210430443.

McKinsey & Company (2003), "Kaizen", unpublished internal training material, McKinsey & Company, Amsterdam.

Meijer, R.R. (1994), "The number of Guttman errors as a simple and powerful person-fit statistic", *Applied Psychological Measurement*, Vol. 18 No. 4, pp. 311-314, doi: 10.1177/014662169401800402.

Mohaghegh, M. and Furlan, A. (2020), "Systematic problem-solving and its antecedents: a synthesis of the literature", *Management Research Review*, Vol. 43 No. 9, pp. 1033-1062, doi: 10.1108/MRR-06-2019-0284.

Murray, P. and Chapman, R. (2003), "From continuous improvement to organisational learning: developmental theory", *The Learning Organization*, Vol. 10 No. 5, pp. 272-282, doi: 10.1108/09696470310486629.

Mohaghegh, M. and Grössler, A. (2019), "Developing problem-solving capabilities for sustainable solutions: a dynamic capability approach", *paper presented at the 26th EurOMA Conference: Operations Adding Value to Society*, 17–19 June, Helsinki.

Netland, T.H. (2016), "Critical success factors for implementing lean production: the effect of contingencies", *International Journal of Production Research*, Vol. 54 No. 8, pp. 2433-2448, doi: 10.1080/00207543.2015.1096976.

Oliva, R. (2019), "Intervention as a research strategy", *Journal of Operations Management*, Vol. 65 No. 7, pp. 710-724, doi: 10.1002/joom.1065.

Örtenblad, A. (2001), "On differences between organizational learning and learning organization", *The Learning Organization*, Vol. 8 No. 3, pp. 125-133, doi: 10.1108/09696470110391211.

Paoletti, J., Bisbey, T.M., Zajac, S., Waller, M.J. and Salas, E. (2021), "Looking to the middle of the qualitative-quantitative spectrum for integrated mixed methods", *Small Group Research*. doi: 10.1177/1046496421992433.

Pérez, I.J., Cabrerizo, F.J., Alonso, S., Dong, Y.C., Chiclana, F. and Herrera-Viedma, E. (2018), "On dynamic consensus processes in group decision making problems", *Information Sciences*, Vol. 459, pp. 20-35, doi: 10.1016/j.ins.2018.05.017.

Powell, D.J. and Coughlan, P. (2020), "Rethinking lean supplier development as a learning system", *International Journal of Operations and Production Management*, Vol. 40 Nos 7-8, pp. 921-943, doi: 10.1108/Ijopm-06-2019-0486.

Pretz, J.E. (2008), "Intuition versus analysis: strategy and experience in complex everyday problem solving", *Memory and Cognition*, Vol. 36 No. 3, pp. 554-566, doi: 10.3758/MC.36.3.554.

Putnik, G.D. (2012), "Lean vs agile from an organizational sustainability, complexity and learning perspective", *The Learning Organization*, Vol. 19 No. 3, pp. 248-266, doi: 10.1108/09696471211220046.

Radhakrishnan, R. and Balamurugan, P. (2010), "Six sigma based exponentially weighted moving average control chart", *Indian Journal of Science and Technology*, Vol. 3 No. 10, pp. 1052-1055.

Rafaai, Z.F.B.M., Rasib, A.H.B.A. and Ishak, Y.B. (2018), "Short review of QCC (Quality Control Circle) implementation toward productivity improvement: case study", *International Journal of Engineering Research and Technology*, Vol. 7 No. 11, pp. 68-72.

Rother, M. (2019), *Toyota Kata: Managing People for Improvement, Adaptiveness and Superior Results*, MGH, New York, NY.

Shaw, M.E. (1932), "A comparison of individuals and small groups in the rational solution of complex problems", *The American Journal of Psychology*, Vol. 44 No. 3, pp. 491-504, doi: 10.2307/1415351.

Shingō, S. (2007), *Kaizen and the Art of Creative Thinking: The Scientific Thinking Mechanism*, Enna Products Corporation, Bellingham.

Shute, V.J. (2008), "Focus on formative feedback", *Review of Educational Research*, Vol. 78 No. 1, pp. 153-189, doi: 10.3102/0034654307313795.

Taggar, S. and Brown, T.C. (2001), "Problem-solving team behaviors: development and validation of BOS and a hierarchical factor structure", *Small Group Research*, Vol. 32 No. 6, pp. 698-726.

Tortorella, G.L., Fettermann, D., Cauchick Miguel, P.A. and Sawhney, R. (2020), "Learning organisation and lean production: an empirical research on their relationship", *International Journal of Production Research*, Vol. 58 No. 12, pp. 3650-3666, doi: 10.1080/00207543.2019.1633028.

Tucker, A.L., Edmondson, A.C. and Spear, S. (2002), "When problem solving prevents organizational learning", *Journal of Organizational Change Management*, Vol. 15 No. 2, pp. 122-137, doi: 10.1108/09534810210423008.

Van den Bossche, P., Gijselaers, W., Segers, M. and Kirschner, P. (2006), "Social and cognitive factors driving teamwork in collaborative learning environments", *Small Group Research*, Vol. 37 No. 5, pp. 490-521, doi: 10.1177/1046496406292938.

Van Dun, D.H. and Wilderom, C.P.M. (2021), "Improving high lean team performance through aligned behaviour-value patterns and coactive vicarious learning-by-doing", *International Journal of Operations and Production Management*. doi: 10.1108/IJOPM-11-2020-0809.

Van Dun, D.H., Hicks, J.N. and Wilderom, C.P.M. (2017), "Values and behaviors of effective lean managers: mixed-methods exploratory research", *European Management Journal*, Vol. 35 No. 2, pp. 174-186, doi: 10.1016/j.emj.2016.05.001.

Van Ginkel, W.P. and Van Knippenberg, D. (2008), "Group information elaboration and group decision making: the role of shared task representations", *Organizational Behavior and Human Decision Processes*, Vol. 105 No. 1, pp. 82-97, doi: 10.1016/j.obhdp.2007.08.005.

Van Someren, M., Barnard, Y. and Sandberg, J. (1994), *The Think Aloud Method: A Practical Approach to Modelling Cognitive Processes*, Academic Press, London.

Vo, B., Kongar, E. and Suárez Barraza, M.F. (2019), "Kaizen event approach: a case study in the packaging industry", *International Journal of Productivity and Performance Management*, Vol. 68 No. 7, pp. 1343-1372, doi: 10.1108/IJPPM-07-2018-0282.

Vuculescu, O., Beretta, M. and Bergenholtz, C. (2021), "The IKEA effect in collective problem-solving: when individuals prioritize their own solutions", *Creativity and Innovation Management*, Vol. 30 No. 1, pp. 116-128, doi: 10.1111/caim.12416.

Wheelan, S.A. (2009), "Group size, group development, and group productivity", *Small Group Research*, Vol. 40 No. 2, pp. 247-262, doi: 10.1177/1046496408328703.

Woods, D.R. (2000), "An evidence-based strategy for problem solving", *Journal of Engineering Education*, Vol. 89 No. 4, pp. 443-459, doi: 10.1002/j.2168-9830.2000.tb00551.x.

Yeo, R.K. and Marquardt, M.J. (2010), "Problems as building blocks for organizational learning: a roadmap for experiential inquiry", *Group & Organization Management*, Vol. 35 No. 3, pp. 243-275, doi: 10.1177/1059601110369729.

Appendix.
Coding scheme of verbal contributions in the videotaped KE meetings

Kaizen event phase	Goal of the phase	Orientation and core question to be answered in this phase	Coding of verbal remarks based on the intention of a participant's verbal contribution
1. Problem definition	Understand the objective(s), situation and future impact	What is the expected impact of this problem?	Participant explores the context and content of the problem
2. Root-cause analysis	Explore the root causes	What causes this problem?	Participant discusses and weighs the potential causes
3. Generate ideas	Creatively develop possible solutions	Which ideas may help to remove the root cause(s)?	Participant elaborates on ideas (e.g. possible countermeasures)
4. Plan implementation	Develop a plan of what is needed to realise a change	What should be done to create the desired change?	Participant explores the commitment to change and the actions required to create the desired change
5. Implement	Execute the plan	Can you accept the change and adopt the new standard operating procedure?	Participant discusses the observations and own and other's responses to the change
6. Check and sustain	Set the new standard	Did the new standard operating procedure solve the problem which started it all?	Participant discusses the extent to which the problem is solved and the solution is embedded in the daily operations

About the authors
José C. M. Franken (MEng.) is currently working as a continuous improvement programme manager at the University of Twente, the Netherlands. With a background in computer science, she worked for over 15 years as IT project leader and ICT manager. From 2003, she has been working as a lean advisor (lean master black belt) in many different multinational production as well as service organisations. She started her PhD research in 2018 from a personal interest in improving the dynamics and outcomes of teams and individuals in kaizen events. Her 2020 EurOMA conference paper was nominated for the Harry Boer Best Doctoral Paper Award. José C.M. Franken is the corresponding author and can be contacted at: j.c.m.franken@utwente.nl

Dr. Desirée H. van Dun is an assistant professor in Change Management and Organisational Behaviour at the University of Twente, the Netherlands. Apart from 10 years of lean management consulting experience, she has a cum laude PhD degree in Organisational Behaviour and Operations Management (2015); the thesis won various local and international awards. Her research has appeared in journals such as the International Journal of Operations and Production Management, European Management Journal, Journal of Manufacturing Technology Management, International Review of Industrial and Organizational Psychology, BMC Medical Research Methodology, International Journal of Lean Six Sigma and Business Ethics: A European Review. Van Dun is currently studying effective leadership and team dynamics to achieve sustainable operational excellence.

Prof. dr. Celeste P.M. Wilderom holds the chair in Change Management and Organisational Behaviour at the University of Twente, Enschede, the Netherlands. In 1987, she obtained her PhD in Psychology from the State University of New York, Buffalo (USA) and then worked in the Economics and Business Administration Faculties at both Tilburg University and the Free University in the Netherlands. She has been an associate editor of various international management journals such as the British Journal of Management, the Academy of Management Executive/Perspectives and the Journal of Service Management; she is currently a member of various editorial boards. Her research pivots on effective leader- and followership as well as teamwork in various profit-making and non-profit work settings.