Measuring the impact of specific activities to improve innovation capacity of teams in a German insurance company

PhD Thesis

Submitted by:	Philipp Jeroen ter Haar Universitat Politècnica de València / Hochschule der angewandten Wissenschaften Hamburg phitehaa@doctor.upv.es

Supervisors:	Prof. Dr. Juan Marin-Garcia Universitat Politècnica de València jamarin@omp.upv.es	Prof. Dr. Birgit Peters Hochschule der angewandten Wissenschaften Hamburg birgitkaethe.peters@haw-hamburg.de

Date:	Dezember 2019

Abstract:

Fostering innovativeness in every team of a company is essential in order to stay competitive today. However, little is known about the real effect of measures to improve a team's innovativeness. The purpose of this dissertation is to test the impact of such measures to support human resource development. A mixed method triangulation research strategy was applied. Two approaches for measuring innovation performance were developed based on a structured literature review and qualitative expert interviews: One based on the idea of counting innovations along the steps of the innovation process (counting tool) and one which also integrates the value of the innovations into the assessment (valuing tool). A pretest-posttest control group design was used while conducting a quasi-experiment with 18 teams of a German insurance company. Findings were two-fold: On the one hand the validity of the counting tool to measure innovation performance of teams could be proven. Despite strongly correlating results this was not possible with regard to the valuing tool. On the other hand, they suggest a statistically significant effect of the measures on innovativeness, particularly implementing a pin board and providing employees with time for innovations can be recommended. The results provide a new perspective on measuring innovativeness and extends existing frameworks for future academic investigations. Additionally, it can be used by managers to invest in and select effective measures to improve the innovativeness of teams.

Medición del impacto de las medidas seleccionadas en el rendimiento de las innovaciones de equipos

Fomentar la innovación en todos los equipos de una empresa es esencial para seguir siendo competitivo hoy en día. Sin embargo, se sabe poco sobre el efecto real de las medidas para mejorar la capacidad de innovación de un equipo. El propósito de este trabajo es probar el impacto de tales medidas para apoyar el desarrollo de los recursos humanos. Se aplicó una estrategia de investigación de triangulación con métodos mixtos. Se desarrollaron dos enfoques para medir el rendimiento de la innovación basados en una revisión estructurada de la literatura y en entrevistas cualitativas con expertos: Una basada en la idea de contar el número de las innovaciones a lo largo de los pasos del proceso de innovación (herramienta de recuento) y otra basada en la integración del valor de las innovaciones en la evaluación (herramienta de valoración). Se utilizó un diseño de grupo de control antes y después de la prueba mientras se realizaba un cuasi-experimento con 18 equipos de una compañía de seguros alemana. Los resultados fueron dobles: Por un lado, se pudo comprobar la validez de la herramienta de recuento para medir el rendimiento de innovación de los equipos. A pesar de la fuerte correlación de los resultados, esto no fue posible con respecto a la herramienta de valoración. Por otra parte, sugieren un efecto estadísticamente significativo de las medidas sobre la capacidad de innovación, en particular la implantación de un tablón de anuncios y la concesión de tiempo a los empleados para innovar. Los resultados proporcionan una nueva perspectiva para medir la capacidad de innovación y amplían los marcos existentes para futuras investigaciones académicas. Además, pueden ser utilizados por los directivos para invertir y seleccionar medidas eficaces para mejorar la capacidad de innovación de los equipos.

Mesura de l'impacte de les mesures seleccionades en el rendiment de les innovacions d'equips

Fomentar la innovació en tots els equips d'una empresa és essencial per continuar sent competitiu avui en dia. No obstant això, se sap poc sobre l'efecte real de les mesures per a millorar la capacitat d'innovació d'un equip. El propòsit d'aquest treball és provar l'impacte d'aquestes mesures per donar suport al desenvolupament dels recursos humans. Es va aplicar una estratègia d'investigació de triangulació amb mètodes mixtes. Es van desenvolupar dos enfocaments per mesurar el rendiment de la innovació basats en una revisió estructurada de la literatura i en entrevistes qualitatives amb experts: Una basada en la idea de comptar el nombre de les innovacions al llarg dels passos del procés d'innovació (eina de recompte) i una altra basada en la integració del valor de les innovacions en l'avaluació (eina de valoració). Es va utilitzar un disseny de grup de control abans i després de la prova mentre es realitzava un quasi-experiment amb 18 equips d'una companyia d'assegurances alemanya. Els resultats van ser dobles: D'una banda, es va poder comprovar la validesa de l'eina de recompte per mesurar el rendiment d'innovació dels equips. Tot i la forta correlació dels resultats, això no va ser possible pel que fa a l'eina de valoració. D'altra banda, suggereixen un efecte estadísticament significatiu de les mesures sobre la capacitat d'innovació, en particular la implantació d'un tauler d'anuncis i la concessió de temps als empleats per innovar. Els resultats proporcionen una nova perspectiva per mesurar la capacitat d'innovació i amplien els marcs existents per a futures investigacions acadèmiques. A més, poden ser utilitzats pels directius per invertir i seleccionar mesures eficaces per millorar la capacitat d'innovació dels equips.

Acknowledgements

"The traveller has reached the end of the journey!" Edmund Burke

What a journey it has been. I could fill a book with the names of those to whom I am grateful for bringing me to this point in my life, for the experiences that shaped me into the person I am today. This PhD is the result of all of the experiences that led up to this point and the people in them. To each and every one of you, I thank you.

I would be amiss not to thank a number of people personally, people who have taken this journey with me. To my supervisors Birgit and Juan; it has been quite a ride, one I would have never taken were it not for you giving me the opportunity, for that and for your counsel I thank you. To Eberhard Kühn, your generosity in spending time discussing the quantitative data analysis was unrivaled. To my colleagues Susanne Grell, Stefan Markert, Frank Rother und Stefan Späth, I am not sure you will ever know how much you helped me by letting your business units take part at the experiment. I am also grateful for the commitment of you, dear team leaders, who conducted the measuring and brought the practical application of the tools forward in the discussions. To my mother, Christine ter Haar – Mama – I could never in a million years thank you enough for what you have done for me and I am sure I am well aware of how many opportunities in life I owe to your sacrifices and hard work (although I am sure I do not know the half of it). Dear Papa, Prof. Dr. Carel ter Haar, your academic career of course was a motivation and inspiration to start this research and I am sure you are proud seeing me ending this journey from heaven. Dear Mia and Levi, I thank you for your patience when you wanted to play but I was spending the time working on this thesis. Last, but absolutely in no way least, to my beautiful wife Ela, thank you for always being there. You have been by my side since the beginning of this journey and have lived the highs and lows with me. I love you with all my heart.

To those named here and those not, I thank you.

Content

List of tables

List of figures

1. Introduction

1.1. Motivation and contribution

It is broadly accepted in today’s literature that ‘it [concept of innovation] is considered one of the essential ingredients of competitive advantage given that it is an intangible component that is difficult for competitors to replicate.’ (Marin-Garcia, Perez-Peñalver, Vidal-Carreras, & Maheut, 2012, p. 920) Thus, this topic is on the one hand in the focus of nowadays research and on the other hand many companies invest (significantly) to improve their innovation performance in the next couple of years. These investments are not only allocated to hire R&D experts but also to improve the innovativeness on all levels of the organisation. In particular, the manifold resources and skills already existing in the company’s teams are regarded as the nucleus for innovation (Kurz, 2013, p. 34). Their internal activities on team level cover a broad range, e.g. Google is conducting innovation workshops (Schulz, 2014, p. 68) or 3M is accepting organisational slack (Bunduchi, 2009, p. 542) by giving employees time for being innovative (Black, 2016, p. 2). Thus, to tap the innovation potential of this nucleus should be in every company’s focus.

While companies are investing it is generally of high relevance to accompany these investments with frameworks which are able to measure the investment’s success. To just name a few, some companies use an adapted form of the balance score card – the innovation balance scorecard – to measure innovation (Fischer, Möller, & Schultze, 2015, p. 646), others are measuring the return on R&D (Vahs & Brem, 2015, p. 646). Alternatively, Bloomberg is offering the "Bloomberg innovation index" to measure innovation (Coy, 2015, p. 1) while other authors use the so called “Innovation Capacity Indicator” (López-Claros & Mata, 2010, p. 1) or the Innovation Competence Barometer (Butter & van Beest, 2017, p. 33) which was jointly created from universities and companies in an EU-founded project.

Generally speaking, the topic of measuring and controlling innovation is broadly discussed within the literature and by various organisations responsible for creating general innovation indexes or measuring a company's innovation. This led to a significant rise in the number of published articles on innovation, innovation measuring and innovation controlling since 2001 (for details please refer to figure 1).

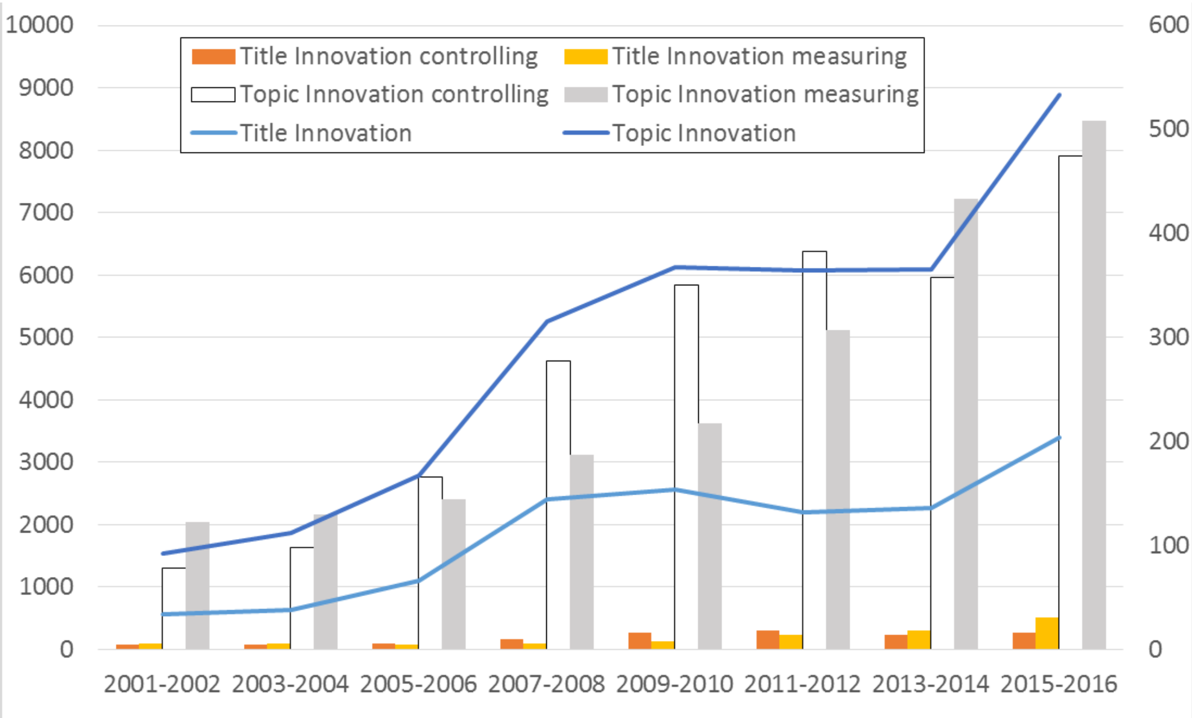

Figure 1: Growth in published papers in the field of innovation

* The literature research was undertaken in the Web of Science using *innovation* and *innovation controlling* as keywords in the subject fields business, management and business finance

While existing literature highlights a wide array of indicators and concepts on innovation measuring three aspects are essential for supporting the measuring oh innovativeness of work teams:

- Which concepts and tools for innovation measuring exist and are they exhaustive?
- How is there applicability into daily practice on organisational work team level?
- Which measure to foster innovation is the most effective?

By addressing these important questions this work contributes twofold to the field of innovation measuring. On the one hand research on innovation measuring will be strengthened. An overview over the existing literature on innovation measuring allows to compare existing concepts and their underlying assumptions to allow an assessment of the measuring approach. Additionally, it also provides the opportunity to identify potential gaps and field for future research. The application in daily business can be seen as yardstick against which theories should be assessed. Integrating the lessons learned into the theoretical concepts will move theory building forward and facilitate future dissemination. Finally, the author is not aware of any research evaluating the effect of measures to foster innovation of work teams.

On the other hand being provided with tested instruments to measure innovation is highly interesting for managers and other practitioners concerned with fostering innovation in teams. Such a tool enables to set a baseline and identify high-performing teams over the course of time. Comparing and analysing teams allows to identifying aspects with relevance to innovativeness. Being provided with an indication of the effectiveness of relevant measures from begin offers the opportunity to focus investments and the scope of own activities.

Personally, being employed as a manager of three organisational work teams, the author feels that the general understanding of the associated mechanism in his company is currently limited and an improvement would be a significant leap forward to face today's competition.

1.2. Research objective

The overall aim of the thesis is the development and testing of a tool for measuring innovation performance of organisational work teams. In differentiation to individual innovativeness the focus of this work are the innovative results of a group. Measuring individual innovation competence is often conducted by evaluating behavioural indicators such as critical thinking or networking (Zhou & George, 2001, p. 687). Innovativeness of groups of persons – ranging from small teams up to whole countries – can be assessed by indicators describing the input (e.g. per cent of revenues invested in R&D), process (e.g. per cent of ideas commercialised) or output (e.g. per cent of sales from new products) of an innovation (Goffin & Mitchell, 2017, p. 314) (please refer to chapter 4 for more details). In this setting the work intends to provide a result of high practical relevance which is based on today's research state of the art. The main objectives can be summarized as follows:

- Increase our knowledge on innovation measuring by summarising today's state of the art, identifying underlying concepts and potential gaps for future research
- Providing a tool based on today's state of the art theory and applicable by practitioners
- Identify measures to increase innovativeness and investigate their impact on work teams

As such the thesis seeks to make contributions to the field of innovation measuring and improving innovativeness in teams as described in section 1.1.

1.3. Outline of thesis

An explanatory sequential design approach is used to achieve the research objective: Developing the measuring instrument by item generating and purification and testing its usability and validity by performing a quasi-experiment. The outline of the thesis is oriented on this design:

- Chapter 2 gives an in-depth description of the research design to create the measuring instrument. In particular the approach to identify existing approaches in a structured literature review based on stages is elaborated on. Also the selection of interview experts and interview preparation and analysis is documented.
- Chapter 3 presents relevant definition in the context of innovation measuring and gives a definition of the term “innovation performance” to clarify the scope of the thesis
- The overview and assessment of existing approaches and the concept to cluster them in four levels is the content of chapter 4.
- Chapter 5 investigates in the generation an purification of measuring items to develop an instrument for innovation measuring. The two instruments based on counting innovations (Counting tool) and value the impact of an innovation (valuing tool) are introduced as well as the advantages of creating two different tools.
- Hypotheses and quasi-experimental design for evaluating the usability and quality of the tools and the impact of the measures are provided in chapter 6. This includes general design aspects as well as the discussion on potentially confounding factors for the research, such as the issues of team selection, selection of fitting measures as intervention, communication and ethical issues of covert research.
- The experimental results are described in chapter 7. Next to general observations, on the one hand the quality and the easy of handling of both tools is assessed separately. On the other

hand, the impact of the six measures selected for intervention is analysed and the hypothesis validated.

- Finally, chapter 8 summarises the findings of the research with regard to the developed instruments and the impact of the measures. It concludes with indications for further research topics.

1.4. Declaration of contribution

While this thesis is written as a monograph, parts of it are also published. They are published as single author articles and were reviewed by leading management journals. This applies for the following chapters:

- Chapter 4:

  The literature review was published in the article *Measuring innovation: a state of the science review of existing approaches*, Intangible Capital, 2018

  The evaluation of existing approaches was originally presented at DSCIM conference, Novi Sad, 2017 and published in the conference proceedings. Based on the discussions during the conference and the further course of the research the author developed the concept further.
- Chapter 5:

  The development of the instruments is also described in the article *New approaches to measure innovation performance of teams* and under review in Intangible Capital
- Chapter 6 and 7:

  The topic of research design of the quasi-experiment and the impact of measures to foster innovativeness is also addressed in the article *Fostering innovation in teams: experimental study of the effectiveness of 6 measures* which is accepted for publication in the Scopus-ranked journal “Int. J. of Business Innovation and Research”. The majority of the work has been done independently by the author of this dissertation. Dr. Eberhard Kühn provided useful insights while discussing the quantitative results of the dissertation.

Tables and figures included in the work are created by the author unless otherwise cited.

2. Research design

The research design is determined by the objective that the results fulfil managerial demands and contribute to the innovation literature. Appropriate therefore is a mixed-method design of two steps: a theoretical and expert-based development of the measuring approach followed by an experimental evaluation of its usability and validity.

2.1. Derivation of research approach

Mixed method approaches are criticized by some authors due to the fact that assumptions and preoccupation into underlying research methods are ignored (Smith & Heshusius, p. 8) and that quantitative and qualitative research is seen as incompatible paradigms not allowing combine the information produced (Morgan, 1998, p. 363). However, most researchers emphasize the strengths of the data collection and analysis techniques and believe it is possible to merge them. Therefore, 'mixed methods research has become an increasingly used and accepted approach in business research' (Bryman & Bell, 2019, p. 643). With regard to the research question of the study, this triangulation research strategy allows an investigation which is based on theory but also takes the practical applicability into account, so the use of mixed methods will strengthen the results by cross-checking the findings (Deacon, Bryman, & Fenton, 2014, p. 48).

Based on this considerations, an explanatory sequential design approach using both qualitative and quantitative methods (Bryman & Bell, 2019, pp. 646–647) was selected being the most appropriate way to fulfil the research requirements. It consisted of the following steps:

1) Theoretical and expert-based development of the measuring approach by
    a. Clarifying relevant definitions
    b. Generation of measuring items by conducting a literature review to identify existing tools and approaches for measuring innovation
    c. Purification of measuring items by expert interviews to validate and /or complement the potential tools
    d. Testing of instruments
2) Experimental evaluation of its usability and validity by
    a. Performing a quasi-experiment amongst organizational working teams
    b. Structured interviews of the participating team leaders

The method used in this study to develop measures for an organisational team's innovation performance will follow Churchill's general design involving amongst others item generation and purification, pretesting, revision, development of a preliminary instrument, ascertaining internal consistency and determination of validity (Churchill, 1979, p. 66). The procedure will be adapted for the current study. The research design to evaluate the use (step 2) is based on the results of defining tools (step 1). It will therefore be described in section Experimental goal and hypotheses6.1 after the decision which tools will be tested to measure innovation performance.

2.2. Generating of measuring indicators

A comprehensive overview over existing and relevant methods of innovation measuring not only described in academic literature but also offered as tools to customers is the base for selecting the right instrument. Considering paradigms for scale development (Gerbing & Anderson, 1988, p. 187), this stage involved the generation of indicators to measure the essence of an organizational team's innovation performance. A proven review strategy based on stages was applied (Pittaway, Robertson, Munir, Denyer, & Neely, 2004, 139f; Rousseau, Manning, & Denyer, 2008, p. 51) to identify the relevant approaches. The following steps were taken by the author (Spender, Corvello, Grimaldi, & Rippa, 2017, p. 5):

i) Based on a preliminary review of the literature on the topic of innovation, the author identified a set of key words by using the mindmapping technique. Exemplary key words were "innovation", "measuring", "assessment" and "controlling".

ii) Search strings were used to find the relevant contributions, e.g. the string [innovat* AND measure* OR control* OR assess*] was used at the beginning.

iii) This string was used for an initial search in Google Scholar to determine additional key words. For example, additional words such as "performance", "capability", "R&D", "indicator" and "index" were found to be relevant and added to the analysis.

iv) As a base a search using the string ["measur* innovat*" OR "control* innovat*"] was undertaken in three search engines: Elsevier's Scopus (576 contributions), Web of Science (372) and EconBiz (240).

v) Due to the highest number of resulting contributions, Scopus was selected to apply the strings identified in step ii) and iii). These strings were continuously refined up to the most complex. 321 contributions were gained using the final string [(TITLE-ABS-KEY (“measur* innovate*” OR “control* innovat*” OR “measure* R&D” OR “innovation index” OR “measure* innovat* performance”)) AND (performance) AND (LIMIT-TO (SUBJAREA, “BUSI”) OR LIMIT-TO (SUBJAREA, “ENGI”) OR LIMIT-TO (SUBJAREA, “ECON”))]. The author is glad to provide the full list to anyone interested

vi) Inclusion and exclusion criteria were defined to further refine the results (see Table 1). The reasoning was to include all methods and tools which describe a way to measure innovation independent from the measured object (e.g. single product, individual or country). Due to the fact that organizations or specialized companies offer to evaluate innovativeness, this aspect was included as well in the list of criteria as long as the measuring method could be obtained.

Inclusion Critieria	
Criteria	**Reason for inclusion**
Theoretical papers	Provide the working assumptions used in this study and the study on existing tools
Qualitative and quantitative empirical studies	To include all empirical evidence
Study books / lecture documents	To describe tools for measuring innovation relevant for students and practitioners in a concise manner
Tools offered by companies or organizations	To describe tools and prove their applicability in reality for measuring innovation. These were only included if the methodology of the approach was described.

Exclusion criteria	
Criteria	**Reason for exclusion**
Studies on activities / methods to increase innovation	The literature review aims to obtain indepth knowledge of existing measuring techniques. Activities / methods to improve innovation performance or capabilities have a different focus and will not support this research.
Influencing factors on innovativeness	The study shows an overview over the existing tools / methods for measuring innovation and does not intend to develop an own approach based on existing influencing factors.
General measures of team effectiveness not focusing on innovation	The focus of this research is the ample field of measuring innovation. To include adjacent research areas, such as measuring team effectiveness, inhibit the risk of loosing focus. So the broad discussion on measures of team effectiveness was comprehensively described by Delgado et al (Delgado Piña, María Romero Martínez, & Gómez Martínez, 2008, p. 7-21) and was excluded unless a reference to innovation was recognizable.

Table 1: Inclusion and exclusion criteria

vii) These criteria were used to assess the contributions gained from Scopus in two steps: to begin with, the titles and keywords of the articles were evaluated based on the inclusion and exclusion criteria. Secondly, the abstracts of supposedly relevant articles were analyzed accordingly. It was found out, that the vast majority of literature is addressing the need for being innovative and activities or methods to increase the level of innovation and were excluded. So finally 30 approaches fulfilled the focus of this review.

viii) To include tools and methods offered by companies or organisations even though they might not have been referenced by Scopus, the author also used the search string from step iv) in Google to ensure the completeness of the search results. While applying the inclusion and exclusion criteria on the 69 results and there linked references, 6 additional approaches from the business sector and country indices were identified.

ix) Finally, the reference sections of the total of 36 approaches was evaluated according to these criteria to assess the search strategy and search for additional approaches. However, additional advances could not be identified.

x) Each approach was assigned to one of the levels described above and evaluated based on the aspects measuring technique and indicators used. During this process themes common to the approaches were identified.

Collected ideas and open aspects with regard to measuring innovation performance were than discussed in structured interviews with 5 experts to ensure completeness of research and applicability in real business life.

2.3. Expert interviews - approach

Reviewing the literature aimed to find a tool appropriate to measure a team's innovation performance. To ensure an optimal measurement during the experiment, the researcher sought additional input from experts. Factual expert interviews were the most suitable approach to obtain the relevant information. Information can be gathered in a structured way but also allow flexibility for additional questions and open discussion (Lamnek, 2002, 173ff). Thereby it is possible to get in-depth insight into the experts' knowledge and particular experiences how innovation performance can be measured (Bogner & Menz, 2009, 7ff) so the literature results will be strengthened by this triangulation research strategy (Scandura & Williams, 2000, p. 1249) before the use of the tool during the quasi-experiment. Thus, an explorative expert interview is the right mean in this setting (Kvale & Brinkmann, 2015, p. 132).

2.3.1. Selection of interviewees

Before approaching potential experts for the interviews, criteria for their selection were defined. To ensure interview results supporting the research, the experts were selected based on the following criteria:

- Expert knowledge in or responsibility for innovation and / or innovation measuring
  To support the specific topic of measuring innovation performance of teams beyond the means of a literature review requires in-depth knowledge of the topic. Interviewees therefore should have had intense contact with the topic of assessing or measuring innovation, for example by their actual role or responsibility, by their participation in projects related to innovation measuring or by being part of the academic discussion on innovation.
- The group of experts should represent business and academic background
  The pursue for completeness of research and the applicability in real business life can best be fulfilled by a combination of experts with thorough business and scientific background.
- Willingness to cooperate

By applying these criteria, five persons could be won to support the research:

- Dr. Matthias Wiedenfels, CEO of Stada AG by the time of the interview
  Being innovative is highly relevant for Stada AG, as it is for all companies in the pharmaceutical industry. In addition, researchers in Germany are entitled to participate on an innovations financial success by law ("Arbeitnehmererfindungsgesetz"). Next to generally pushing the company to increase innovativeness, Dr. Wiedenfels managed the project to reorganize the process of innovation controlling and linking the financial success to single persons or teams

during the time of the research. Thus, he qualified as a business expert by his role and the intense project work.

- Prof. Dr. Juan Marin-Garcia, Professor - Universitat Politècnica de València - Research group: ROGLE

  Next to moderating workshops on innovation in the industry, Prof. Marin-Garcia is deeply involved in researching innovation measuring in the EU-founded Fincoda project. This project developed a psychometric tool that measures individuals' capacity for innovation (www.fincoda.eu), so his insights and experiences will support this work from the academic perspective. In addition he published articles on innovation in various journals.

- Dr. Rene Butter, Expert at University of Applied Science Utrecht & Consultant

  Next to his participation on the Fincoda project, Dr. Butter is an expert in optimizing and innovating assessment programmes and conducts lectures on research and innovation. He also published articles on innovation measuring. Interviewing him gave the opportunity to support the research academically, in particular on new ways of assessing innovation.

- Marja Salenius-Ranki, SVP Human Resources, Elomatic

  Elomatic, a leading supplier of worldwide industrial engineering and consulting services needs to be innovative to compete in the globalised world. Ms. Salenius-Ranki is responsible for fostering innovation amongst employees. She was also part of the Fincoda project which gives her additional credits as an expert on innovation.

- Matthias Heutger, SVP DHL Innovation Centres

  To identify innovative logistics solutions and test robotics and automation of processes, the world's biggest logistic provided DHL founded three innovation centres in Bonn (GER), Singapore and Chicago (US). Mr. Heutger is the head of this initiative. Next to his focus on continuously improving his teams to identify trends and provide new solutions, he also had the responsibility for the cost-benefit-ration of these centres.

In total five experts could be won to support the research: three with a strong business and two with an extensive scientific background. Business as well as academic background is represented by them, so the criteria mentioned above are fulfilled. In addition, this number is within the recommended sample range for expert interviews (Kvale & Brinkmann, 2015, p. 140). Furthermore, the seniority of the experts and the fact that the interview results were not the single source within this study ensures that this sample size achieved the goal to strengthen the results from literature research.

2.3.2. Interview preparation, framework and analysis

Conducting the expert interviews aims to gain information and insights in addition to the expertise from literature review. Allowing new ideas to be brought up in an interview and discussing them can be best achieved by the method of a semi-structured interview (Ritchie, Lewis, McNaughton Nicholls, & Ormston, 2013, p. 183; Bryman & Bell, 2019, p. 434). A rigorous set of questions as practiced by a structured interview guideline will not allow adding additional content. Due to the clear focus on innovation measuring a guideline for the semi-structured interviews (Bryman & Bell, 2019, p. 437) was developed.

The guideline consisted of three parts: personal background, measuring innovation (what to measure and how) and questions with regard to applying a measuring tool. The last two reflected the structure of the present study and ensured to obtain relevant knowledge with regard to developing the tool and conducting the experiment. It also prepared prepared the structure of the planned thematic analysis (Bryman & Bell, 2019, p. 519). Within these sections results from literature review are presented, challenged and improved by the interviewee. The structure and the questions of the interview framework are described in detail in section 5.2.2

To inform the interviewees about the planned contents and allow them time for preparation, the author sent information about the research and the short version of the guideline consisting of 17 questions at least 1 week upfront the interview. 18 additional questions of various types (see Appendix B), were not disclosed to the interviewees to allow flexibility in the course of the talk and to allow the possibility to document spontaneous answers (Bryman & Bell, 2019, p. 448)..

The interviews represent the personal impressions and experiences of the interviewee’s profound experience in innovation measuring. The interviews took between 60min – 90min and were recorded as well as notes were taken. Due to the fact that the results from literature review should be purified by the interviews, it was defined that the analysis of the documented information focused on meaning (Kvale & Brinkmann, 2015, pp. 231–232). By using a thematic analysis approach six categories (see Table 2) were identified and the interview results were documented within these. The categories followed the main structure of the interview framework by addressing the KPIs for measuring innovation performance, their application on the given examples, the questions of weighting, evaluation of innovation value and applicability. The aspect “measures relevant in praxis” was added to allow documentation of general information / insights on practical innovation measuring independently from the framework’s structure. This approach allows a structured and comparable analysis of the collected information (Bryman & Bell, 2019, p. 519).

Topic	Interviewee 1	Interviewee…
KPIs/important aspects to measure: - Output - Process		
Experiences from rating example		
Weighting indicators (need / method)		
Evaluation of innovation value		
Important aspects of applicability		
Measures relevant in praxis		

Table 2: Thematic analysis matrix

The outcome of the thematic analysis matrix was used for purification of the literature review and built the fundament for the development of tools for measuring the teams' innovation performance.

2.4. Testing

As the main objective was to develop an instrument to measure an organisational team's innovation performance, the instrument was tested with leaders' of organisational teams in a quasi-experiment. This is consistent with the piloting approach that the subjects selected should be those who are likely to use the instrument (Bryman & Bell, 2019, p. 289). The test details build upon the result of developing the measurement and will be described in section 6.2.

3. Definitions

What is innovation? Relevant definitions have to be clarified to set the scope for the study. It is said that the term innovation was firstly used by Saint Augustin (around 400 AD) referring to something newly-created (Quadbeck-Seeger, 1998, p. 101). Schumpeter introduced the term in modern economical theory in 1939. He pinpointed that innovation not only includes invention but also the realisation of the ideas (Schumpeter, 1939). Thus, Innovation is a process through which a new product, technique or useful service is obtained from the generation of new ideas and their development, which in time provides new solutions to problems and becomes useful for people, companies or society. Based on the perceived level of change, the innovation can be incremental (improvement of existing products), radical (generation of something new), or a transformation (Marin-Garcia, Aznar-Mas, & Ladrón de Guevara, 2011, p. 26). 'Based on what is changed, the innovation can involve products, services, or processes' (Mudrak, van Wagenberg, & Wubben, 2004, p. 291). 'Therefore, innovation starts with the proposal and generation of new ideas and finishes with the use and commercial exploitation of the outcomes (Tonnessen, 2005)' (Marin-Garcia et al., 2011, p. 25).

This definition implies a process from idea to implementation. While all authors are following this process logic, the number of stages of innovative behaviour differs from two up to five stages (see Table 3). Most authors follow Scott & Bruce (1994, p. 581-582) supporting the idea that innovative behaviour is complex and basically consists of three different stages: idea generation, idea promotion and idea implementation.

Another group of authors identified four stages. They also agree on the elements idea promotion and idea implementation. In contrary to the three stages process they differ on the view of the stage idea generation by defining problem recognition as a stage on its own. Innovative behaviour is assumed to already start with the concrete and active search for possibilities to improve existing products, services or processes (Jong & den Hartog, 2010, p. 24). Only one study splits the innovation process into five stages (Kleysen & Street, 2001, p. 285).

Within this work, the concept of a three stages process consisting of idea generation, idea promotion and idea realization will be followed. The innovation process therefore begins with the identification of problems and the generation of ideas and solutions (Kanter, 1996, p. 96). Then it is necessary to convince other people of the idea's value. One has to build coalitions and to find sponsors on the one hand to ensure the supply of resources necessary for implementation and on the other hand to overcome potential resistance against the new product, process or service. Finally, one has to put the idea into practice to fulfil to distinct an invention from an innovation (Marin-Garcia et al., 2011, p. 25).

Author	**Stages**
2 Stages	
Bunce & West, 1995	• Introduction of something that is new to the unit of adoption • Application
Basu & Green, 1997	• Creation • Implementation of concepts and products
Axtell et al., 2000	• Idea generation • Idea implementation

Author	**Stages**
3 Stages	
Scott & Bruce, 1994	• Problem recognition and generation of ideas and solutions • Seek sponsorship for an idea • Complete the idea by producing a prototype or model of the innovation
Janssen, 2001, Ramamoorthy, Flood, Slattery, & Sardessai, 2005, Reuvers, van Engen, Marloes L., Vinkenburg, & Wilson-Evered, 2008	• Idea generation • Idea promotion • Idea realization
Stashevsky, Carmeli, Meitar, & Weisberg, 2006	• Idea generation • Promote solutions or ideas • Realize ideas
Carmeli & Spreitzer, 2009	• Recognize a problem • Promote solutions or ideas • Produce a prototype or model of the innovation
4 Stages	
Jong & den Hartog, 2010	• Idea exploration • Idea generation • Idea championing • Idea implementation
Pieterse, van Knippenberg, Schippers, & Stam, 2010	• Problem recognition • Generation of ideas and solutions • Building support for ideas • Idea implementation
Dorenbosch, van Engen, Marloes L., & Verhagen, 2005	• Problem recognition • Idea generation • Idea promotion • Idea realization
5 Stages	
Kleysen & Street, 2001	• Opportunity exploration • Generativity • Formative investigation • Championing • Application

Table 3: Stages of the innovation process (Kurz, 2013, p. 22)

When innovative individuals, teams, companies and countries are mastering this process effectively, they are assigned attributes in manifold terms:

- Innovative ability
- Innovation competence
- Innovation skills
- Innovative capacity
- Innovation capability
- Innovation performance

This work studies the innovation of organisational teams. While the general definitions indicate a potential relevance in this context by describing a talent / skill, an expertise or the power to do something or a successful performed task (see Table 4), the specific use with regard to innovation has to be evaluated. The objective is to identify the term which fits best to the setting of the study.

Term	**Definition (Oxford dictionary, accessed Nov 15th, 2018)**
Ability	Talent, skill or proficiency in a particular area
Competence	Ability to do something successfully or efficiently
Skill	Ability to do something well, expertise
Capacity	The amount that something can produce
Capability	Power or ability to do something
Performance	A task or operation seen in terms of how successfully it is performed

Table 4: Definition of terms

‘*Innovative ability* is the ability of an enterprise’s employees to generate ideas and to work with these ideas to develop new or improved products, services, technologies, work processes or markets’ (Jong, Kemp, & Snel, 2001, p. 13). On the one hand this definition is very narrow by describing single employees being able to handle the innovation process as defined above. On the other hand it is very wide by covering employees of different kinds and in different settings, from doctors (Brolmann, Vervest, & Heineman, 2001, p. 743) to facility managers (Mudrak et al., 2004, p.290). Some authors also see it as a prerequisite to be innovative and describe people’s characteristics which can affect innovation ability, such as willingness to take risks and commitment (Jong et al., 2001, pp. 13–14).

The use of the wording innovation competence is strongly linked with the EU-founded Fincoda project (www.fincoda.eu), which focussed on rating innovativeness of persons, and its predecessor, the Incode project. Within these projects ‘innovation competence can be defined as those capabilities, which are needed for a successful innovation (Forsman, 2009).[...] Organizational competence identifies an environment where innovations can easily be developed, identified and encouraged. Individual competence identifies an individual person's capabilities of being involved in the different innovation processes of the organization’ (Kairisto-Mertanen & Mertanen, 2011, p. 27). Thus, innovation competence in general terms can be defined as ‘ability to create, introduce, adapt and/or apply beneficial novelty at any organizational level (Marin-Garcia et al., 2016, p. 121). Based on this definition there is also a relation with the term innovation skills. ‚Competency can also be defined as complex knowledge resulting from the integration and adaptation of capacities and skills to situations that share common characteristics. Capacity is moderately complex expertise, which incorporates skills that require procedural and conditional knowledge. A skill, on the other hand, is straightforward know-how (Bessant et al. 2001; Drejer 2001)‘ (Marin-Garcia et al., 2012, p. 921). The Fincoda project has defined three capacities innovation competence is comprised of: creativity (ability to think beyond traditional ideas), critical thinking (ability to analyse and deconstruct issues) and initiative (ability to carry out actions) (Marin-Garcia et al., 2016, p. 121). These elements are linked to personal abilities and are the base for measuring individual innovative behaviour by the newly developed tool Innovation Competence Barometer (Butter & van Beest, 2017, p. 1), so, even though it is also used for describing organizations, the term innovation competence has a clear focus on the employee's input to create innovations.

The term innovation / innovative capacity has a different notion. This phrase is used particularly in a context while describing the innovative strength of countries or geographical areas. *Innovation capacity* has been mentioned by several authors with regard to a country's institutional potential to sustain innovation (Hu & Mathews, 2005, p. 1328). K. Pavitt, M.E. Porter and L. Suarez –Villa, all of them claim rights of a term invention (Lukjanska, 2010, p. 43). It is defined as 'a country's potential – as both a political and economic entity – to produce a stream of commercially relevant innovations. L. Suarez -Villa defined a similar concept of innovation capacity, but named it innovative capacity, measuring the level of invention and the potential for innovation in any nation, geographical area or economic activity (Villa, 1990, p. 295). This capacity is not simply the realized level of innovation but also reflects the fundamental conditions, investments, and policy choices that create the environment for innovation in a particular location or nation.' (Stern, Porter, & Furman, 2000, p. 5). In discrepancy to the terms innovation ability and innovation capacity addresses innovation capacity therefore not only the conditions facilitating innovation (input) but also includes the implemented innovations (output) and has a focus on entities instead of individuals.

'*Innovation capability* is seen more as the ability to exploit new technological combinations; it embraces the notion of incremental innovation and 'innovation without research'' (Cornell University, INSEAD, & WIPO, 2014, p. 68). It is used to define the ranking of countries within the

Global Innovation Index. Input (such as infrastructure) and output indicators (such as knowledge creation) are used (Cornell University et al., 2014, p. 73). It is also defined as ‘the sum of underlying processes that enable innovation, from organizational culture and strategy, to ideas, their selection and their implementation’ (Goffin & Mitchell, 2017, p. 347). Mentioning the underlying process links this definition to the stages of innovation (idea generation, idea promotion and idea realisation). Martinez et al (2011) deduce from an analysis of the literature that the term is used to describe an organisation’s internal ability to innovate consisting of the three dimensions knowledge, organization and the human factor (Martínez-Román, Gamero, & Tamayo, 2011, p. 460). Similarly to innovation capacity the term is used in the context of input and output elements of entities.

Finally, innovation performance is used to describe companies innovative output (Chen, 2009, p. 109; Tidd & Bessant, 2014, p. 289; Laursen & Foss, 2003, p. 249) and defined as ‘the output of new products, services, processes and business models and the financial impact of these.’ (Goffin & Mitchell, 2017, p. 346). The clear company focus is accentuated by the wording “commercialization”. In addition, it can be assumed that the term “new products” is a placeholder for new services or processes as well. To find the right wording for the specific setting of this study it is necessary to evaluate these phrases and identify the most suitable term. Criteria are defined to support this evaluation. With regard to the objective of this work – measuring the impact of activities to improve innovation in teams – it is feasible to challenge the terms by ensuring their:

- Applicability on organizational teams

  Due to the objective of this work the definition must be connected to teams / organizational units

- Output orientation

  While measuring innovation one has to distinct between innovative results ("output") and conditions / circumstances to create innovations ("input"). The conditions could be given by a highly skilled work force, the fundamental conditions for investments, the company strategy or an innovation friendly environment within the company. Within this work, interventions to foster innovativeness are varied with the goal to identify their output. Therefore, the selected term should have an output orientation.

The terms innovation ability and innovation skills are purely used to characterize individuals. In addition, they are both also seen as a prerequisite to create output, so they fail to fulfil both set criteria. Nearly all authors use the wording "innovation competence" to describe individuals as well. In particular the state of the art definition of its elements creativity, critical thinking and initiative developed by some of the participants of the Fincoda project (Marin-Garcia et al., 2016, p. 121) emphasises this view. An application on teams had to be verified. Due to the fact, that innovation competence describes a prerequisite for innovation, this term will not be used in the study. The other three terms are used to describe innovation in entities. The phrase "innovation performance" is already used in the company context. Even it is uncommon, it should be also possible to transfer "innovation capability" and "innovation capacity" from describing countries or regions to companies. But innovation capability still does not fulfil the criterion output orientation due to its focus on the underlying process. The other two wordings are used to describe the innovative output of a company, thus are conform with the criteria. Referring to the definitions of both, the notion on "its financial impact" gives the term innovation performance a clear output focus. Due to the fact that it is also already used in the company context, it will be used in this work.

	Applicability on organizational teams	Output orientation
Innovative ability	Term used to describe individuals, transfer to organizational units by accumulating the single results might be possible	Prerequisite to create an output due to its narrow focus
Innovation competence	Term is used by most authors to characterize individuals, few authors also use it to describe organisations, so a transfer to teams has to be verified	Prerequisite to create an output due to its focus on the environment
Innovation skills	Term describes straightforward know-how. This is rarely used in the context of organisations	Prerequisite to create an output due to its focus on individual's aspects
Innovative capacity	Transferability from the country context on organizational teams possible but not yet common	Next to the focus on input / conditions the term also includes elements to describe output
Innovation capability	Applied on nations and companies, a scaling down to organizational teams should be possible	Prerequisite to create an output by taking environment for innovations / the underlying process into account
Innovation performance	The phrase is applied to describe companies	By using the words "output" and "impact" there is a strong result orientation

Table 5: Analysis of terms describing innovation

To sum up, most of the terms indicate a focus on input or the environment for innovation. While innovative capacity also includes elements to describe output, the phrase "performance" has a stronger focus on produced innovation. Thus, the wording *innovation performance* is the most suitable to be used in this work. For this purpose it is defined as 'The total innovation produced by a single organisational team, which includes the value created with regard to the three stages of an innovation (idea generation, promotion and realization) and their underlying process' (ter Haar, 2017, p. 50). This definition partly integrates the idea of the term innovation capabilities. This is recommendable because to value a team's innovativeness one should assess the results on the one hand, on the other hand it is also necessary to take the way the results are achieved into account (Goffin & Mitchell, 2017, p. 313), due to the fact that it can be assumed that a higher level of innovation performance will lead to reduced resource needs while implementing the innovation. Hence, whereas the word "produced" describes a strong output orientation, the reference to the underlying process includes the second aspect sufficiently.

4. Overview over existing approaches

Searching the term “measuring innovation” in Google shows that manifold ways to measure innovation exists (> 100m hits). To give an overview over the 36 approaches identified by the structured literature review (see section 2.2) the following aspects were discussed:

- Thus, firstly a framework for clustering these approaches into the four levels country, company, team and individual was developed (section 4.1)
- Secondly, a quantitative analysis of the identified approaches was conducted (section 4.2)
- Thirdly, a descriptive review of the approaches was provided (section 4.3 to section 4.6)
- Finally, the gained insights were summarized (section 4.7 and 4.8) and limitations described (section 4.9).

These steps enabled to identify whether these approaches are exhaustive and applicable to this setting

4.1. Four Levels of Analysis framework

The manifold approaches resulting from the structured literature review (see section 2.2) were clustered to facilitate the analysis and improve clarity. In a similar context Anderson et al. used a framework of 4 levels to evaluate innovation and creativity in organizations: the individual, the work team, organizational, and multi-level approaches (Anderson, Potocnik, & Zhou, 2014, 1302f). Within this study, this idea was transferred on innovation measuring to cluster the identified approaches based on the object which is measured:

- The assessment of particular skills which are supposed to lead to innovativeness on an individual level are highly relevant e.g. for universities to adapt their curriculum or HR in the hiring process. Employee level innovation metrics also help on deciding about rewards and recognition (Goffin & Mitchell, 2017, p. 302). This level of analysis should therefore be taken into account within the present review.

- It is also highly relevant for companies to be able to judge on the innovation performance of a single organisational team or project team. Knowing the “as is” status of a team`s innovation performances enables managers to on the one hand decide which teams need specific measures to increase their innovativeness and on the other hand to control the success of an implemented measure, in case the level of innovation has increased.
- Today, most innovations are not based on a single person or teams but on the effort of the whole company or at least teams working across business units (Fischer et al., 2015, p. 642). To mirror this situation, approaches to measure innovation are also required on organizational level. Even though there is a grey area with regard to the organisational team level in which exact definitions are difficult to apply, within this review advances to evaluate multi-project teams and business units consisting of more than one organisational or project team respectively will be evaluated within this category.
- However, ‘Every country wants to foster a culture of innovation’ (Coy, 2015, p. 1). To do so, one must know which countries succeed in this task to learn from their success. Because it is necessary to be able to evaluate the innovation level of a country or region, relevant advances exist which cannot be categorized to one of the other levels. Thus, this level of measuring should be added to the framework used by Anderson et al.
- Anderson et al. are evaluating studies on creativity and innovation and were also including multi-level research into their review (Anderson et al., 2014). Within the context of this paper taking multilevel analysis’ of innovation into account will not be reasonable. Even though

work teams and projects might also be defined as small subordinated units of organizations, the individual, the organizational and the country level differ too much from each other. So, a specific category for measuring innovation based on multilevel analysis is not required.

To sum up, on the one hand Anderson et al. already proved that this framework is working to cluster and analyse innovation approaches in organisations. On the other hand, it allows a plausible grouping in a mutual exclusive but collectively exhaustive structure of advances measuring innovation while only minor adjustments are required, in particular the layer of multi-level analysis has to be replaced by the country / region level (see Figure 1). Other clustering criteria, such as source of approach or measuring technique do not support the research question for measuring team performance in an equally good manner. In the following sections this framework will therefore be used within this study to cluster the identified measuring approaches in more detail. The identified advances of each level will be described and their main characteristics evaluated. By reviewing representative academic approaches and business tools clustered in the four levels individual, work team, company and country this study pursues two objectives: assessing open fields of today's research and supporting decision makers to identify the tool fitting to their specific need.

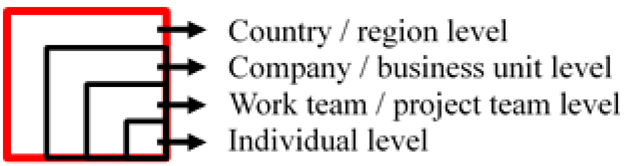

Figure 1: Levels of analysis of measuring innovation

Thus, this work will give an overview over 36 approaches to measure innovation which are grouped into the four levels described above. Then, it will be discussed whether the existing approaches describe the topic exhaustively.

4.2. Quantitative analysis of approaches

A data analysis of the 36 identified approaches (see Table 9) was conducted. Generally speaking, descriptions of approaches can be found in three different sources: study books or lecture documents, journal articles and as service from organisations offering to assess the level of innovation. Naturally, some advances are described in more than one source. To examine the typology of identified approaches and trace the major sources the content analysis method was adopted (Leonidou, Katsikeas, & Coudounaris, 2010, p. 80); (Furrer, Thomas, & Goussevskaia, 2008, p. 5) with regard to the aspects characteristics of approach description and the approaches' measuring technique. The author also considered conducting an analysis of key words. Due to the high percentage of tools described in study books or being offered by companies or organizations – sources which lack specific key words regarding the approach – a meaningful result could not be expected and the idea was dismissed.

The advances were described in three ways characterized by the frequency of applying them: either generally / theoretically (how one should do it but not applied yet), in a study / experiment (one time application to validate the concept) or the applicability as a tool was proved by showing the results of their regular application. Analysing these characteristics in relation to the categories of the four levels of analysis framework (as described in section 4.1) all ways can be found in nearly equal shares with general descriptions as the most common one (see Table 6). Within the levels of analysis major differences exist. By far the most advances are focussing on the company / business unit level (50%) of the 4-levels of analysis framework while the other levels have an equal share. Overall, general descriptions are dominating (42%), studies and tools can be found in an nearly equal share (31% vs. 28%).

All approaches on country level are applied regularly and provide actual rankings (e.g. Bloomberg Indicator or the Global Innovation Index). Innovation level of countries as a major competitive advantage is regarded of high importance for economical and political decision making. While this shows that these tools are working, of course it does not prove the validity of the results or the meaningfulness of the chosen criteria.

Some of the tools on company level are described by various authors, e.g. the IPOO-model (Fischer et al, 2015, p.646; Vahs & Brehm, 2015, p.644; Möller et al. 2011, p.30) or the innovation balanced scorecard (Fischer et al, 2015, p.647; Goffin & Mitchell, 2015, p.351) an example for their actual application in daily business could not be found. This might be due to various reasons. Some of them are very complex requiring a lot of data gathering effort (e.g. IPOO model) which might make the dissemination difficult. Alternatively, their might be a discrepancy between the importance of innovation as stated by the companies and their willingness to invest resources in actually establishing a controlling system. The high number of studies on company level emphasize their importance as a frequent study objects. Some of the advances seem to be very complex and unlikely to be implemented into daily business (e.g. Laursen & Foss, 2003, p. 510, who use a complex empirical analysis model to analyse company data). Others use simple systems, e.g. Michie & Sheen (2003). Their approach consists of simply two questions: Firms were asked: 'during the past three years, has the company introduced any product innovations?' and 'during the past three years, has the company introduced any process innovations?' (Michie & Sheehan, 2003, p. 129). Still, the percentage of tools actually applied is very low with only 11%.

All approaches on project level are generally described. The tools mentioned are often (variances of) tools which are currently applied in companies (e.g. ROI, NPV, milestone controlling). This high grade of acceptance causes that further studies do not seem to be necessary. Thus, while general descriptions on work / project team level also include commonly used tools, on company level this mostly includes general concepts. The intensity of research on employee innovativeness is mirrored by the fact that 50% of advances could be found in studies. The companies' interested in assessing the employees' innovation competence is supported by the high implementation rate of fitting approaches. In addition, the EU founding for projects researching approaches on measuring individual innovation competence like Fincoda show the high importance of the topic.

	Total	**Tool**	**General description**	**Study**
Country	6	100%	0%	0%
Company / BU	18	11%	44%	44%
Work / project team	6	0%	100%	0%
Employee	6	33%	17%	50%
Total	36	28%	42%	31%

Table 6: Characteristics of approach description

Different techniques are used to measure the approaches indicators' (see Table 7): quantitative, semi-quantitative and / or qualitative. Semi-quantitative techniques are basically qualitative judgements that are converted to numbers. They differ from quantitative technique in that no attempt is made to use a sophisticated formula to complete the data. Qualitative techniques are intuitive judgements (Pappas & Remer, 1985, p. 15).

It can be stated that nearly all approaches try to quantify the level of innovation, 16 out of the total of 36 advances (44%) even using a purely quantitative technique to achieve the most objective result in a comparison and be able to provide a ranking on clear facts. Using semi-quantitative measuring or the combination of both attempts to achieve the same goal to strive for objectivity. For example, by using semi-quantitative measuring techniques the advances strive to provide a complete picture of the innovation within a company, also taking cultural and strategic aspects into account which are transformed into data by using a scale such as the 4-point Likert scale (e.g. Innovation Balanced Scorecard, Persaud (2005,p. 414). This is also required to enable the tool to also include behavioural aspects (e.g. knowledge sharing, creativity) into the assessment of employees, e.g. Keys Scale (Amabile, Conti, Coon, Lazenby, & Herron, 1996, p. 1166) or the Innovation Competence Barometer (Butter & van Beest, 2017, p. 33). In particular employee metrics only covering output-related aspects (e.g. ideas created) might falsify the created picture. In total, the vast majority (91%) of advances measures innovation with quantitative and semi-quantitative indicators, while avoiding qualitative ones. Decision makers obviously prefer the easy comparability of the results. Qualitative assessments are far more difficult to compare.

On country level only quantitative data is used for evaluation. While semi-quantitative evaluations are difficult and time-consuming to conduct for companies, the higher complexity and amount of information required for countries is taken into account be applying purely quantitative ratings. Even though the majority of advances to measure innovation projects is also focusing on quantitative or semi-quantitative methods, also qualitative methods are used. While describing the project's feasibility or sustainability qualitative information is required. Assessing individual innovative behaviour in a purely qualitative way is not proposed, all approaches use quantitative (17%) or semi-quantitative (67%) scales or the combination of both (17%). Managers get the opportunity to compare employees and assign task fitting to their strengths.

	Total	Qualitative	Quantitative	Quantitative & qualitative	Semi-quantitative	Semi-quantitative & quantitative
Country	6	0%	100%	0%	0%	0%
Company / BU	18	6%	39%	0%	39%	17%
Work / project	6	17%	33%	17%	0%	33%
Employee	6	0%	17%	0%	67%	17%
Total	36	6%	44%	3%	31%	17%

Table 7: Measuring technique per level of analysis

Assessing which measuring techniques were applied in the different ways it can be noticed that the identified tools did not use qualitative indicators, but they occurred in advances depicted in general descriptions (14%, e.g. integrated evaluation method Vahs & Brem, 2015, p. 346) and studies (7%, e.g. Inno-framework Kauffeld, Jonas, Grote, Frey, & Frieling, 2004, p. 156). The majority of tools use purely quantitative indicators (60%), while amongst studies an equal split of quantitative and semi-quantitative indicators could be found (45% each). It can be noticed that tools and studies do not measure by combinations of indicators (e.g. like semi-quantitative and quantitative) while general descriptions provide the most heterogeneous picture by collecting data with all kinds of indicators. This confirms that qualitative indicators are difficult to use in practical business life. The more likely the application of an approach becomes the more specific were the indicators suggested by the authors.

	Total	Qualitative	Quantitative	Quantitative & qualitative	Semi-quantitative	Semi-quantitative & quantitative
General description	15	7%	33%	7%	13%	40%
Study	11	9%	45%	0%	45%	0%
Tool	10	0%	60%	0%	40%	0%

Table 8: Measuring technique per type of description

To sum up, the quantitative analysis identified a great variety of ways and measuring approaches. With regard to the selection of an appropriate tool for measuring innovation performance of teams it has to be taken into account, that:

- Even though the company level is in focus of today's literature there is a lack of tools currently applied
- Quantitative or semi-quantitative indicators were proposed by the authors of tools actually applied. This indicates that they are preferable for the planned quasi-experiment as well

With regard to the descriptive analysis the evaluated advances show a great variety in methodology and objective while measuring innovation. Based on the definition given above (see section 4.1), the identified approaches for measuring innovation were reviewed to find an approach with fit to the planned research. To facilitate the analysis, they were clustered in the four levels focussing on country level, company level, organisational team / project team level or the level of the individual.

4.3. Descriptive review – country level

One level of approaches were focussing on country level. They all had in common that they were already applied as tools using quantitative indicators to be able to support policy makers and business leaders to decide on their investment or innovation strategies (The Economist, 2014, p.2, Cornell University, INSEAD, & WIPO, 2017, p. ix). Some of them, such as the Innovation Capacity Index (López-Claros & Mata, 2010) were using the same focus and methodology year for year. Others changed their focus, e.g. the Global Innovation Index was focussing on the human factor of innovation in the report 2014 (Cornell University et al., 2014), in the report 2016 the theme 'innovation is feeding the world' was covered (Cornell University et al., 2017) or the European Innovation Scorecard changed the framework significantly for its 2017 report compared to the report from 2016 (Hollanders & Es-Sadki, 2017, p. 8). Changing or adapting indicators allowed the authors to react to actual developments, e.g. changing politic priorities, technical advances such as digitisation or taking newly identified structural differences into account. Depending on the themes the focus of the analysis and of the stated comments is changing independently from the measured variables which might stayed nearly the same (e.g. Global Innovation Index) or underwent major changes (e.g. European Innovation Scorecard). While adapting indicators allowed reacting to actual topics, sticking to the same method increased the comparability of the development over the years. Both alternatives enhance evidence for policy making purposes.

Looking at the indicators the indices are based on one can recognise two different types:

1. Approaches covering a broad range of relevant input and output factors of innovation
2. Approaches using only key indicators

Approaches of the first type developed significantly in recent years. Early approaches measured less indicators – such as the National Innovative Capacity using 27 indicators with a strong focus on education, R&D and the use of innovative goods (Stern et al., 2000, p. 4), such as the effectiveness of intellectual property protection, the sophistication and pressure to innovate from domestic buyers, availability of venture capital (Stern et al., 2000, pp. 8–10). Recently published indices were based on far more indicators and attempt to show a broad picture of aspects describing an innovative country. The actual Global Innovation Index for example was based on a total of 81 indicators (Cornell University et al., 2017), the Index on Innovation Capacity 61 (López-Claros & Mata, 2010, p. 19) and the Economist indicator on 91 criteria (The Economist, 2014, p. 9). Even though the concrete indicators of this kind of approach were not identical, the reasoning beyond the selected aspects is similar. The criteria can be allocated to the following clusters:

- Stability and quality of the public institutions with regard to fostering business

  e.g. political stability and absence of violence/terrorism, depth level, foreign investment gross inflows
- Existence and development of human capital

  e.g. adult literacy rate, gender equity, expenditure on education, researchers
- Easiness of doing business within the existing regulatory & legal framework

  e.g. ease of starting a business, ease of doing a business, ease of employing workers
- Good and ecologically friendly infrastructure in particular with regard to information & communication technology

  e.g. mobile subscribers per 100 inhabitants, internet user per 100 inhabitants, electricity output, GDP per unit of energy use
- R&D infrastructure

  e.g. university/industry research collaboration, R&D worker density, R&D expenditure (as % of GDP)
- Outputs of innovation activities

  e.g. patents granted to citizens, industrial design by origin, citable documents H index

Due to a lack of data availability in some countries, all approaches struggled to base their innovation index exclusively on hard data. Thus, they also used composite indicators or survey results to find the appropriate trade-off of quality of the variables and achieving a good country coverage (Cornell University et al., 2017, 48f).

The second type of approaches was particularly represented by the Bloomberg innovation indicator. Even though it followed similar reasoning like the first group in selecting important elements to describe a country's innovation level it was using only 6 indicators: R&D expenditure as percentage of GDP, manufacturing, Hi-tech companies, education, research personnel and patents (patents then split in 2 groups – filed and granted) (Coy, 2015, p. 8). Overall, it had a stronger output focus by leaving enablers such as regulatory framework or infrastructure aside. Comparing the approaches results made noticeable differences obvious: For example, the Bloomberg indicator identified 2016 South Korea, Germany and Sweden as top 3 countries (Coy, 2015, p. 7), the Global Innovation Index claims Switzerland, Sweden and Netherlands to be the top 3 (Cornell University et al., 2017, xviii) and the Economist Singapore, Switzerland and Hong Kong (The Economist, 2014, p. 2).

To conclude, while all advances on country level were actually applied tools using quantitative measuring technique, major differences exist. Recent frameworks either used broad variety of indicators while attempting to provide a concise picture of the situation or reduced the evaluation on core outputs of innovation, enablers such as the regulatory framework or innovation infrastructure were not measured. In addition, the authors followed different philosophies in selecting the indicators. Some focussed on the development of the countries over the years by always employing the same indicators, others regularly revised their frameworks and survey themes to react to political or technological development. With regard to the applicability of indicators on this work's setting it can be stated that it is unlikely that variables describing enabling elements (such as aspects of political situation or infrastructure) can be transferred. On the contrary, output indicators should be kept in mind for the further assessment in section 4.8.

4.4. Descriptive review – company or business unit level

The variety of approaches to measure innovation on company or business unit level is high. Generally speaking they can be clustered into the following groups based on their focus:

1. Being linked to the innovation process
2. Indicators clustered into dimensions
3. Other approaches, e.g. assessing innovation climate or ability

Models of the first group use the idea of measuring input – process – output of an innovation (Goffin & Mitchell, 2017; Tidd & Bessant, 2014, p. 289). Some authors, in particular German ones, propose to separate output (short term effect) from outcome (long term result) (Möller, Menninger, & Robers, 2011, p. 30-32) or even differentiate between output, receiving system (marketing and sales efforts) and outcome (Vahs & Brem, 2015, p. 644). The Input-Process-Output-Outcome model (IPOO) for example uses particular indicators for each step of the company's innovation process, e.g. training cost per employee (input stage) or number of new products (output stage) (Fischer et al., 2015, p. 644) (Fischeretal.,É. Despite this differentiation the indicators used are quite similar, e.g. the indicator "number of patents received" is proposed by Pappas & Remer (1985, p.18), Tidd & Bessant (2014, p.289), the IPOO approach and Goffin & Mitchell (2017, p.316) or the indicator "implemented improvement ideas by Tidd & Bessant (2014, p. 289) and Fuchs (2014, p.40). The detailed overview over the indicators on company level is documented in Appendix C).

Including outcome criteria into the approach (such as market share gained by innovations, etc. (Goffin & Mitchell, 2017, p. 314) implies two consequences:

- A serious delay in being able to obtain valid measuring results. It may take years before an innovation is implemented and positive effects can be identified.
- Unless a product innovation is evaluated it will be difficult to separate the effects of the innovation clearly from the other factors which might have led to the measured increase in market share.

Next to these holistic approaches one can also find ones focussing on specific aspects of the company. To be able to specifically value the efficiency of the R&D efforts only, the R&D Return framework was created (Vahs & Brem, 2015, p. 648). In this approach the R&D productivity was valued by the potential productivity and the technology development efficiency and compared to the R&D yield, consisting of the potential yield and operating efficiency. In addition, a rigid algorithm combining all indicators was used to calculate the R&D return value. Even though this framework might be transferable to other business units as well, it did not cover todays common understanding that innovation is not only driven by the R&D department but by the whole company (Kurz, 2013, p. 34). The frameworks described above obviously demand a sophisticated measuring or auditing system to collect and rate the required data. To simplify the data collection the bean counting approach attempted to just quantitatively measure the innovative output of an R&D department (Pappas & Remer, 1985, p. 18). This was done by collecting indicators such as patents, technical publications, rewards, etc. or simply by counting ideas (Fuchs, 2014, p. 40). All approaches linked to the innovation process provided a detailed description of the proposed indicators. Even though this facilitates applying them, no prove of implementation could be found. This might be due to the high effort required to collect all relevant information.

Within the second group – approaches clustered by dimensions – one can find approaches based on the balanced score card, e.g. the Innovation Balanced Score Card (Fischer et al., 2015, p. 646). The four aspects of the normal balanced score card (Financial, customer, internal business processes, learning & growth (Kaplan & Norton, 1996, p. 54) are evaluated based on the company's vision and – in this case – innovation strategy with a strong focus on increasing innovation success, e.g. time to market, market share gained by R&D (Žižlavský, 2016, p. 56). Another advance evaluating dimensions are innovation audits. Innovation audits do not only look at performance (an output measure) but also how this performance was achieved (a process measure). This is done by evaluating dimensions such as strategy, market, product, technology etc. (Goffin & Mitchell, 2017, p. 317). Exemplary indicators for the aspect market are change in market share, number of customer surveys, number of innovations based on customer ideas, etc. (Warschat, 2005, p. 17) or the 5 dimensions of A.T. Kearney's "House of Innovation" tool "Improve" (innovation strategy, innovation organisation & culture, innovation life cycle processes, enabling factors and innovation results) (Innovety, 2014, p. 5). Assessing the indicators mentioned in literature, it became obvious that dimensions related to internal company factors, such as enabling factors or innovation results, are evaluated by criteria also used in the process models, e.g., aspects of internal HRM, quality of project management, market share gained by the innovation or number of patents. By also taking the strategy and the market situation into account the advances using dimension have a broader focus than the ones of the first group. This aspect implies an even greater effort to collect the required data.

The group of other approaches is manifold. They usually have a specific focus such as:

- The "INNO" questionnaire assessing the innovation climate to allow a standardized an economic measurement of conditions facilitating innovations in organisations (Kauffeld et al., 2004, p. 156) by posing 21 questions clustered in the 4 dimensions (1) activating leadership,

(2) continuous questioning, (3) consequential implementation, and (4) professional documentation (Kauffeld et al., 2004, p. 153)

- The Team Climate Indicator, a multi-dimensional measure of facet-specific climate for innovation within groups at work (Anderson & West, 1998, p. 235). This advance used 38 five-factor items, for example clarity of the teams objective, chance to put forward new ideas, friendliness of team members (Anderson & West, 1998, 246, 254). This advance can be used for companies and single work groups, thus is covering two levels of the 4-level of analysis framework.
- Describing general determinants for a company's innovation ability (Jong et al., 2001). The model includes 50 determinants which were grouped into the 9 categories people characteristics, strategy, culture, structure, availability of means, network activities, company characteristics, innovation infrastructure and market characteristics. (Jong et al., 2001, pp. 9–10). Indicators such as willingness to take risks, innovation in the missions statement, result orientation, use of creativity techniques (Jong et al., 2001, p. 60) were supposed to support a company's innovation ability and can also be found in other approaches. Some of the questions were specifically adopted to the advance's specific objective, e.g. job rotation or exporting activities (Jong et al., 2001, 21, 26).

The approaches from this third group show a very specific focus not correlating with the objective of this research. For that reason they will not be included into the further discussion.

To sum up, one can find a great variety of frameworks to measure innovation within companies or business units. The holistic approaches integrating all aspects of innovation within a company are either structured by dimensions or following the innovation process. They differ greatly with regard to their complexity and need for data, thus the required effort to collect and analyse the data. This should be taken into account while selecting the right approach for a specific situation. Despite all differences the approaches are overlapping in their attempt to evaluate a company's innovation output and related enabling factors. Within these aspects similar indicators are used. In addition, a third group of advances with a specific focus is described in literature. With regard to the forthcoming discussion on measuring innovation performance of teams some indicators mentioned have already been applied on groups, thus might fit for this setting as well.

4.5. Descriptive review – organisational / project team level

The organisational and project team level is the most heterogeneous one. On the one hand this is due to the different kinds of teams which exists in companies: organisational and project teams. Organisational (work) teams are created for long-term interaction in the same work area, project teams are established to fulfil a defined goal which is frequently characterised by short-term interaction and cross-functional cooperation (Heathfield, 2018, p. 3). On the other hand Innovation projects are assessed on two levels: multi-project level and single-project level (Schentler, Lindner, & Gleich, 2010, p. 305).

Various numeric and nonnumeric approaches exist to assess the value of innovation projects on multi-project level, in particular with regard to R&D portfolio management (Dey, 2006, p. 91). Nonnumeric models used in the selection process are ranging from an assessment on the idea of interest to the management ("sacred cow concept") (Meredith & Mantel, 2012, p. 47) to more sophisticated approaches such as the concept of "sustainability" (Gale, 2009, p. 31). Amongst numeric models, well known and popular are those based on profitability, such as payback period or net present value (NPV). But also approaches taking more aspects into account, such as scoring models (Meredith & Mantel, 2012, 51ff) are used. Generally speaking, a wide variety of numeric and nonnumeric selection models 'have been developed to try to help managers to tackle the problem of project selection' (Trott, 2017, p. 329). Four project characteristics were efficient predictors of a projects commercial success: 1) expected profitability, 2) technological opportunity, 3) development risk and 4) appropriateness (degree to which a project is appropriate for the organization undertaking it) (Astebro, 2004, p. 320).

Assessing the success of single innovation projects is usually conducted by common project controlling tools. Fischer and Hauschildt for example describe an approach for operative innovation controlling focussing on projects (Fischer et al., 2015, p. 643). Similarly to the Integrated Evaluation Method indicators of a typical project controlling, such as milestone or budget controlling are proposed (Maier, Streicher, Jonas, & Frey, 2007; Vahs & Brem, 2015). Other advances measure innovative performance of projects by length of development time, costs of development etc. compared to market average or turnover (Fuchs, 2014, p. 31) or the return of investment of innovation projects (Hauschildt & Salomo, 2016, p. 262). These indicators are particularly working well with innovative product development projects. In addition, it is also stated that the only way to assess innovation activities is to evaluate the progress of these innovation projects (Littkemann & Derfuß, 2011, p. 588). This argumentation neglects the fact that innovation within companies is not always linked to projects, but the whole company including not only project but also all organizational work teams (Kurz, 2013, p. 34).

Although there have been many approaches published on innovation measuring the author is not aware of one specifically designed for measuring the innovation performance of organizational teams. There are only two other ways, whose primary focus is on providing concrete indicators to support decision makers and which might be also applied on organizational team level: the R&D return framework (Vahs & Brem, 2015, p. 648) and counting the number of implemented improvement ideas (Fuchs, 2014, p. 40) (see chapter 4.4). All the more so since the extensive literature on measuring team effectiveness is not relevant in this context: on the one hand measures of team effectiveness without innovation focus are excluded from the research (see chapter 2.2). On the other hand additional approaches focusing on innovation or describing a practical way to measure a teams' innovation performance specifically could not be found in the context of general measures for team effectiveness.

To sum up, approaches measuring innovation projects were closely linked to typical project controlling. Even though some of the indicators of approaches measuring innovation of companies might be transferred, no specific framework with regard to single work teams could be identified.

4.6. Descriptive review – individual level

Reviewing approaches with focus on the individual level one can cluster them based on the method obtaining the information and the indicators used. Common methods for rating innovation on individual level are either self-assessment, expert assessment, a combination of both or acquiring archival objective data. Jansen (2000, p. 1042) for example used a self-rating approach to evaluate the innovativeness of employees. However, in the recent years there has been a continuous increase in the use of third party ratings or the combination of self-assessment and third-party rating. Alge et al. for example used a peer review approach for a study on innovative performance (Alge, Ballinger, Tangirala, & Oakley, 2006, p. 225). The combination of both is also proposed, for example combining a self-assessment survey with an expert panel of professors and innovation specialist (Choi & Chang, 2009, p. 248) or a supervisory rating (Yuan & Woodman, 2010, p. 330). The EU funded FINCODA project created the Innovation Competence Barometer to support experts to rate the innovation capabilities of persons but also offers a parallel self-assessment (Butter & van Beest, 2017, p.1). Archival objective data were used to collect the value of objective measures such as patents or scientific publications.

Generally speaking, all frameworks use either output-related or behavioural focused indicators. Different schemes are used to describe the latter. They are either oriented on dimensions – such as creativity, critical thinking, initiative, teamwork or networking (Butter & van Beest, 2017, p.4; Zhou & George, 2001, p. 687) – or on the 3 stages of innovation (idea generation, promotion and implementation) (Janssen, 2001, p. 1043). Within both schemes the evaluated characteristics (for details please refer to table 2) are mostly based on the six items of individual innovative behaviour by Scott & Bruce (1994, p. 607):

- Searches out new technologies, processes, techniques, and/or product ideas
- Generates creative ideas
- Promotes and champions ideas to others
- Investigates and secures funds needed to implement new ideas
- Develops adequate plan and schedules for implementation of ideas
- Is innovative

These are extended and refined in the recent years. The 9 items scale developed by Janssen (2000, p. 1043) are linked to the three stages of the innovation process and assigned three items per stage. Items like "generating original solutions to problems" or "transforming innovative ideas into useful applications" were added to receive more specific information. In their 13 items scale to assess creativity George & Zhou (2001, p. 696) asked for even more additional indicators such as "suggests new ways to increase quality", "often has a fresh approach to problems" or "suggests new ways of performing work tasks". They stated that their approach was focused on measuring creativity and not innovation. By evaluating the items, it became obvious, that they included all items Scott and Bruce used, also including the ones related to realizing the new ideas. Thus, they covered all aspects of innovation and not only the creative steps.

The recently finished Fincoda project uses a differentiated approach of the five dimensions creativity, critical thinking, initiative, team work and networking to assess a persons innovation competence with the state of the art tool Innovation Competence Barometer (Fincoda Project, 2018). 29 questions were asked during the assessment which led to a rating of 34 items describing the dimensions in a five point Likert scale. Items were varying, some of them were similar to the scales mentioned above (e.g. "find new ways to implement ideas", "convince people to support an innovative idea" or "search out new working methods, techniques or instruments"). The vast majority were novel ones (Butter & van Beest, 2017, pp. 32–33) such as:

- Consult about essential changes
- Engage outsiders of the core work group from the beginning
- Ask "Why?" and "Why not?" and "What if?" with a purpose
- Use intuition and own knowledge to start actions or
- Show inventiveness in using resources

Offering an online tool for collecting and analysing the data supported the dissemination into companies successfully, despite the high number of indicators. Another group of authors was proposing output instead of behavioural orientated indicators. By transferring objective indicators also used on company level to rate individuals one tries to make the evaluation independent from personal bias. Typical indicators are number of patents, patent disclosures, ideas submitted to employee suggestion program (Zhou & Shalley, 2004, p. 174). It can be assumed that it will be difficult to rate every employee by these indicators due to the fact that even innovative persons might not participate on a suggestion programme and only very few have the chance to file a patent at all. As Zhou & Shalley formulate it, 'number of patents might not be relevant in nursing' (Zhou & Shalley, 2004, p. 174). To avoid this situation the combination of behavioural and output oriented measuring is also used (Goffin & Mitchell, 2017, p. 302). While this implies additional effort to obtain the data, it will provide the most holistic assessment of an individual.

To conclude, a wide variety of frameworks exists to evaluate innovation on individual level. Even though they can be differentiated based on the method the information is obtained, all behavioural focused approaches are using similar items to measure innovation ability. They distinct themselves by the number of items used. Using online tools for collecting and analysing data supports dissemination even with higher numbers of questions asked. Output-related measures are rarely used for the individual level of analysis. In addition, it can be stated that it is difficult to get a valid rating by using them exclusively. In addition, it can be stated that it is difficult to get a valid rating by using them exclusively because they will only provide meaningful results for people of specific proficiencies.

Overview over discussed approaches

Name / Author	Level of analysis	Phase(s) of value creation	Measu-ring techni-que(s)	Indicators used	Remarks / short description
Global Innovation Index (GII)	Country	Input & Output	Quantita-tive	The GII relies on two sub-indices—the Innovation Input Sub-Index and the Innovation Output Sub-Index— each built around pillars. Four measures are calculated: the two named above and the overall GII score and the innovation efficiency ratio. Thus, there is a strong focus on variables which are not applicable to measure the output of a single team. (Cornell University et al., 2014, p. 73)	'[The GII] helps policy makers and business leaders move beyond one-dimensional innovation metrics towards a more holistic analysis of innovation drivers and outcomes.' (Cornell University et al., 2014, p. 6)
Bloomberg Indicator	Country	All	Quantita-tive	Used metrics (equally weighted): Input: R&D, manufacturing, hi-tech companies, education, research personnel; Output: patents (Coy, 2015, p. 7)	
Economist Index	Country	Input	Quantita-tive	'The business rankings model examines ten separate criteria or categories, covering the political environment, the macroeconomic environment, market opportunities, policy towards free enterprise and competition, policy towards foreign investment, foreign trade and exchange controls, taxes, financing, the labour market and infrastructure. Each category contains a number of indicators that are assessed by the Economist Intelligence Unit for the last five years and the next five years.' (The Economist, 2014, p. 9)	'The business rankings model measures the quality or attractiveness of the business environment in the 82 countries covered by The Economist Intelligence Unit's Country Forecast reports. It is designed to reflect the main criteria used by companies to formulate their global business strategies.' (The Economist, 2014, p. 9)

Name / Author	Level of analysis	Phase(s) of value creation	Measu-ring techni-que(s)	Indicators used	Remarks / short description
Innovation Capacity Index	Country	Input	Quantita-tive	The ICI is built upon five pillars composed of 61 variables. The pillars are institutional environment, human capital, regulatory & legal framework, R&D and adoption and use of information and communication technologies. (López-Claros & Mata, 2010, p. 17)	'[The Innovation Capacity Index is] a tool for assessing the extent to which nations have succeeded in developing a climate that will nourish the potential for innovation.' (López-Claros & Mata, 2010, p. 1)
National Innovative Capacity	Country	Input	Quantitati ve	The innovation capacity index is measured by: • the proportion of scientists and engineers subindex • the innovation policy subindex (effectivness of intellectual property protection, ability of a country to retain its scientists, availability of R&D tax credits for private sector) • the cluster innovation environment subindex (pressure to innovate from domestic buyers, presence of suppliers of specialized research and training, prevalence and depth of clusters) • the linkages subindex (overall quality of scientific research institutions & availability of venture capital for innovative but risky projects) The data is taken from the Global Competition Review (GCR). (Stern et al., 2000, p. 4)	'This capacity is not simply the realized level of innovation but also reflects the fundamental conditions, investments, and policy choices that create the environment for innovation in a particular location or nation.' (Stern et al., 2000, p. 5)

Name / Author	Level of analysis	Phase(s) of value creation	Measu-ring techni-que(s)	Indicators used	Remarks / short description
European Innovation Score-board	Country	Input & output	Quantita-tive	The 2017 EIS measurement framework is built on 10 dimensions with a total of 27 indicators structured in 4 groups: • Framework conditions (8 indicators) • Investments (5 indicators) • Innovation activitites (9 indicators) • Impacts (5 indicators) (Hollanders & Es-Sadki, 2017, pp. 8–10)	'The annual European Innovation Scoreboard (EIS) provides a comparative assessment of the research and innovation performance of the EU Member States and the relative strengths and weaknesses of their research and innovation systems. It helps Member States assess areas in which they need to concentrate their efforts in order to boost their innovation performance.' (Hollanders & Es-Sadki, 2017, p. 8)
Innovation Balanced Scorecard	Company	All	Semi-quantita-tive & quantita-tive	Exemplary selection of indicators of the balanced score card with focus on increasing innovation success(Fischer et al., 2015): • Sales with new products in relation to required investments • Time to market • Market share gained by R&D	Approach describes implementation of a balanced score card with particular focus on innovation success

Name / Author	Level of analysis	Phase(s) of value creation	Measu-ring techni-que(s)	Indicators used	Remarks / short description
Input-Process-Output-Outcome Model (IPOO)	Company	all	Semi-quantita-tive & quantita-tive	Exemplary selection (Möller et al., 2011): • Input: quantitative: personal cost, number of ideas, training cost per employee; semi-quantitative: work experience of employees, quality of ideas • Processing System: quantitative: hours worked per project, number of results achieved in time; semi-quantitative: product / service quality, progress • Output: quantitative: number of patents, number of new products, average cost per patent; semi-quantitative: synergy effects, fundamental research results • Outcome: quantitative: increase in sales / profit, cost reductions semi-quantitative: product improvements, customer satisfaction	The quantified elements can be displayed for example in form of a spider web chart. Möller (et al., 2011) are characterising these indicators as „qualitative" Due to the fact that a quantitative visualization by using a spiderweb diagramme is used, they have to be categorized as semi-quantitative This measuring approach is also supported by other German authors, e.g. (Fischer et al., 2015; Vahs & Brem, 2015)
Innovation Audit Scorecard	Company	All	Semi-quantita-tive & quantita-tive	'Innovation audits look at not only performance (an output measure) but also how this performance was achieved (a process measure).' (Goffin & Mitchell, 2017, p. 317) Indicators are defined for various aspects, e.g. market, projectmanagement, product, innovation culture, know-how etc (Warschat, 2005)	The InnoAudit-Scorecard is an instrument to classify companies and to identify company-specific improvement potential. (Warschat, 2005, p. 13)

Name / Author	Level of analysis	Phase(s) of value creation	Measu-ring techni-que(s)	Indicators used	Remarks / short description
Goffin & Mitchell	Company	All	Semi-quantita-tive	The indicators of this approach equal the quantitative indicators of the IPOO model. The output and outcome indicators of the IPOO are put together in Goffin's aspect output (more detail appendix table 3) (Goffin & Mitchell, 2017, p. 316)	
Improve	Company	All	Semi-quantita-tive	Indicators are clustered in the 5 dimensions innovation strategy, innovation organisation & culture, innovation life cycle processes, enabling factors and innovation results. Exemplary indicators are e.g. time to market / profit, feedback loops, idea management, capacity for innovation, etc (Innovety, 2014, p. 5)	Measuring the company's overall innovation management performance with the Improve tool is based on A.T. Kearney's "House of Innovation" (Innovety, 2014, p. 5)
De Jong et al	Company	Input	Semi-quantita-tive	The model includes 50 determinants which are group into the 9 categories people characteristics, strategy, culture, structure, availability of means, network activities, company characteristics, innovation infrastructure and market characteristics (Jong et al., 2001, pp. 9–10)	The model describes the determinants for innovative ability (Jong et al., 2001, p. 9)

Name / Author	Level of analysis	Phase(s) of value creation	Measu-ring techni-que(s)	Indicators used	Remarks / short description
Commu-nity Innovation Surveys (e.g. CIS Finland, ZEW Germany)	Company	All	Semi-quantita-tive	Exemplary indicators are (Rammer, Crass, Doherr, & et al, 2016, pp. 8–11): Number of product / process innovations, innovation activities conducted in last 3 years, number of new products, surveys with regards to innovation barriers, etc.	The data is used to measure innovation output as a complement to more traditional measures such as patents. (Leiponen & Helfat, 2006, p. 9)
Michie & Sheehan, 2003	Company	Output	Quantita-tive	Firms were asked: 'during the past three years, how many product innovations has the company introduced?' and 'during the past three years, how many process innovations has the company introduced?'. (Michie & Sheehan, 2003, p. 129)	
Persaud	Company	Input	Semi-quantita-tive	Multi-scale instrument describing innovative capacity by 49 items (rated on a scale 1-5) clustered in four main aspects (Persaud, 2005, p. 414): • Autonomy • Formalization • Communication • Socialisation	

Name / Author	Level of analysis	Phase(s) of value creation	Measu-ring techni-que(s)	Indicators used	Remarks / short description
Laursen & Foss	Company	Output	Quantita-tive	Complex calculation based on categorization of companies in four different grades of innovativeness depending on the products or services introduced per period (Laursen, 2003, p. 249): • None innovations (non-innovator) • One product / service new to the firm • One product / service new to the Danish market • One product / service new to the world	
Martinez	Company	Input	Semi-Quantita-tive	Multi-scale instrument describing innovative capacity by 27 items clustered in three main aspects (Martínez-Román et al., 2011, p. 471): • Innovation capability (e.g. organization, human factor, knowledge) • Contextual factors (such as age, size, etc.) • Environment (e.g. competition, quality standards)	Model is based on an extensive review of former approaches and summerizing their input. Thus, the author will not cover these in detail
“Bean counting”	Business unit	Output	Quantita-tive	Exemplary indicators are: patents, technical publications, awards from peer groups, etc. (Pappas & Remer, 1985)	The objective is to quantitatively measure the productivity of the R&D personnel (Pappas & Remer, 1985, p. 18)

Name / Author	Level of analysis	Phase(s) of value creation	Measu-ring techni-que(s)	Indicators used	Remarks / short description
R&D Return Frame-work	Business unit	All	Quantita-tive	Indicators for the R&D Return Framework are (Vahs & Brem, 2015, p. 648): • R&D Productivity: Potential productivity and technology development efficiency • R&D Yield: potential yield and operating efficiency	A total value of the efficiency of the R&D department is calculated in this approach
INNO	Business unit	Input	Qualitative	The INNO instrument has 21 items clustered in four factors: (1) activating leadership, (2) continuous questioning, (3) consequential implementation, and (4) professional documentation. (Kauffeld et al., 2004, p. 155)	The INNO is focusing on the innovation climate in companies and allows a standardized and economic measurement of conditions facilitating innovation in organizations (Kauffeld et al., 2004, p. 156)
Team Climate Indicator	Business unit	Input	Semi-quantita-tive	The Team Climate Indicator measures 38 questions for example with regards to (Anderson & West, 1998, p. 246) team objectives, friendliness of team members, chance to put forward new ideas, appraisal of weaknesses, work evaluation, ways to tackle a problem, etc.	The TCI is 'a measure of group climate in organizations, and for team building and organization development interventions.' (Anderson & West, 1998, p. 255)
Outcome indicators for job center	Business unit	Process	Quantita-tive	'The outcomes we focus on are log job entry productivity as the quantity measure and the quality of service to job seekers (denoted JSQ), the quality of service to firms (denoted EMQ) and the business delivery target (denoted BDT) as the three quality measures'. (Burgess, Propper, Ratto, & Tominey, 2012)	

Name / Author	Level of analysis	Phase(s) of value creation	Measu-ring techni-que(s)	Indicators used	Remarks / short description
Fuchs	Business unit	Output	Quantita-tive	Number of implemented improvement ideas (Fuchs, 2014, p. 40)	
Tidd & Bessant	Business unit	All	Quantita-tive	Exemplary selection of indicators (Tidd & Bessant, 2014, p. 289): • Input: e.g. percentage of sales committed to R&D, investments in Training, recruitment of skilled staff • Process: e.g. number of new ideas, failure rates, number of overruns on development time / cost budgets, measures of continuous improvement (suggestions/employee, number of problem solving teams, cumulative savings, etc.) • Output: specific (e.g. patents, scientific papers, new products), process elements (e.g. customer satisfaction, improvements in quality), comparable (e.g. market share, quality performance, cost of product) and strategic success (e.g. revenue growth, improved quality, higher value added)	
(Meredith & Mantel, 2012)	Project	Output	Quantita-tive	Comprehensive overview over project selection models, e.g. • Nonnumeric models containing indicators such as competitive necessity or sustainability • Numeric models including indicators such as profitability/NPV, payback period or specific indicators needed for scoring	

Name / Author	Level of analysis	Phase(s) of value creation	Measu-ring techni-que(s)	Indicators used	Remarks / short description
(Astebro, 2004)	Project	Process & output	Semi-quantita-tive & quantita-tive	1) expected profitability, 2) technological opportunity, 3) development risk and 4) appropriability	
Integrated evaluation method	Project	Output	Qualita-tive	Indicators to assess the technology and market of an innovation, e.g. by value for customers, technological advance compared to existing products, competitor analysis, R&D lead, existence of required know-how, etc.	Technological aspects and market chances of an innovative product are evaluated by expert interviews and workshops (Vahs & Brem, 2015, 346ff)
Operative innovation Control-ling	Project	Process	Quantita-tive & qualitative	Approach is strongly focused on the project oriented control of innovation activities in the dimensions development costs, product quality and development time (Fischer et al., 2015, p. 643)	
Return on Invest-ment	Project	Output	Quantita-tive	Standard return on investment calculation for innovation projects (Hauschildt & Salomo, 2016, p. 421)	
Fuchs	Project	Output	Quantita-tive & semi-quantita-tive	Depending on availability innovative performance should be measured by not less than length of development time, costs of development, evaluation of innovativeness compared to the market average, turnover / profit achieved. (Fuchs, 2014, p. 31)	Based on the results from the single projects the innovation of the company is analyzed (Fuchs, 2014, p. 31)

Name / Author	Level of analysis	Phase(s) of value creation	Measu-ring techni-que(s)	Indicators used	Remarks / short description
Goffin & Mitchell, 2017	Emplo-yee	All	Quantita-tive & semi-quantita-tive	Use of a combination of output-related and behavioral focused indicators: • Output-orientated: scientific publications, patents, ideas generated, cost savings, project goals, process / service innovation • Behavioral focused: innovation performance rating, teamwork, competencies gained. (Goffin & Mitchell, 2017, p. 302)	Using employee level innovation metrics helps on deciding about rewards and recognition. (Goffin & Mitchell, 2017, p. 303)
Innovation Compe-tence Barometer	Emplo-yee	All	Semi-quantita-tive	Formative online self-assessment tool for students and professionals and tool for supporting structured behavioral interviewing to select innovators (supervisory rating). The self-assessment consists of 34 items clustered in categories creativity, critical thinking, initiative, teamwork and networking (Butter & van Beest, 2017, p. 33)	EU-founded project finished in 2017
Scott & Bruce	Emplo-yee	Process & output	Semi-quantita-tive	Exemplary items of the scale for individual innovative behavior in the workplace are (Scott & Bruce, 1994, p. 607), e.g.: • Searches out new technologies, processes, techniques and/or product ideas • Generates creative ideas, • Promotes and champions ideas to others	Base for extended scales, e.g. Jansen (see below) or 13 items scale by Zhou & George (Zhou & George, 2001, p.687)

Name / Author	Level of analysis	Phase(s) of value creation	Measu-ring techni-que(s)	Indicators used	Remarks / short description
Janssen	Emplo-yee	Process	Semi-quantita-tive	9 item scale for innovative work behavior (Janssen, 2001, p. 1043), e.g.: • Searching out new working methods, techniques, or instruments • Generating original solutions for problems • Acquitting approval for innovative ideas • Making important organizational members enthusiastic for innovative ideas • Introducing innovative ideas into the work environment in a systematic way	Scale to assess individual innovative behavior in the workplace, three items refer to each innovation stage. Scale is supposed to be completed by self-reports and supervisory rating.
Zhou & Shelley	Emplo-yee	Output	Quantita-tive	Exemplary objective measures are (Zhou & Shalley, 2004, p. 174): Number of patents, patent disclosures, research papers, ideas submitted to employee suggestion programs	Objective measures for employees creativity may be useful for R&D departments but not in all settings. For example, number of patents may not be relevant in nursing. (Zhou & Shalley, 2004, p. 174)

Name / Author	Level of analysis	Phase(s) of value creation	Measu-ring techni-que(s)	Indicators used	Remarks / short description
KEYS Scale	Emplo-yee	Input	Semi-quantita-tive	Multi-item scale (KEYS-scale) including indicators with regard to (Amabile et al., 1996, p. 1166): • Organizational encouragement • Sufficient resources • Workload pressure	

Table 9: Overview over discussed approaches

Definition of measuring technique: Semi-quantitative techniques are basically qualitative judgements that are converted to numbers. They differ from quantitative technique in that no attempt is made to use a sophisticated formula to complete the data. Qualitative techniques are intuitive judgements (Pappas & Remer, 1985, p. 15)

4.7. Summary of quantitative analysis and descriptive review

The above narrative review suggests, firstly, that there is a strong preference on quantifying the results of the measuring (approximately 94%, see Table 7). Approaches on company level do not only outnumber the other levels but are also the most heterogeneous ones with regard to their source and measuring technique. Secondly, even though some approaches on country level reduce the evaluation to core indicators a tendency to increase the number and variety of items to improve validity is visible on country, company and individual level. Holistic approaches covering a rising number of indicators describing dimensions or the innovation process led to high complexity and effort to collect data. Furthermore, the frameworks and indicators used were in some cases transferrable between the levels of analysis, e.g. number of patents was used on country level (in relation to population (Coy, 2015, p. 7)) on company level (as absolute number (Pappas & Remer, 1985, p. 18))(Pappas&Remer,Æ and – in case output indicators are used – on project team (Tidd & Bessant, 2014, p. 289) and individual level (Goffin & Mitchell, 2017, p. 302) as well. Transferring indicators might be a relevant aspect in the course of the study due to the fact that no framework specifically applicable to measure innovation on the organisational work team level could be found. Comparable advances focus on teams for innovation projects or on steering multi-projects, thus measuring mainly the process of implementing the innovation. Even though, some aspects of this project controlling might be transferrable on the situation of organisational work teams, it does not cover the output aspects sufficiently and therefore cannot replace an advance for operational controlling of a not project orientated work team.

4.8. Assessing approaches to measure work teams' innovation performance

Only few approaches specifically designed for measuring innovation exist on project team / organizational work team level. Applying common project management tools such as budget controlling, milestone planning or a ROI calculation on innovation projects seems absolutely feasible and sufficient. These tools have been created for all kinds of projects and will therefore support innovation projects as well (Hauschildt & Salomo, 2016, pp. 421–422). it can be stated that the gap of frameworks on organizational work team level needs further research. Innovation will be without doubt the major factor of competitive advantage for companies (Marin-Garcia et al., 2011, p. 25). Thus, the number of companies conducting activities to improve their innovation performance is high and will be rising in the next couple of years. Activities to foster innovation proposed for the team level are manifold, such as team workshops, coffee corners or whiteboard for ideas (Kahlfuss, 2013, 255–256)On the other hand it is important in the global competition to use resources effectively. Decision makers have to focus on investments with a positive outcome. ‘As Marco Iansiti of Harvard Business School has pointed out, ‚after all, what a company gets for the money it spends on R&D is what ultimately matters‘.’ (Goffin & Mitchell, 2017, p. 42). The two identified approaches (see chapter 4.5) are not sufficient: Fuchs’ (2014, p.40) concept of counting implemented ideas is an attempt to evaluate innovation performance with minimized effort. This finding neglects the merits of first stages of the innovation process by only including implemented ideas. However, it provides a basic framework and a direction for further research. On the one hand, the R&D return framework’s concept does not lead to meaningful results while transferred to other units of a company. It can be assumed that indicators such as number of patents or publications will work appropriately for R&D teams, but are inappropriate for other kinds of teams, in particular operative-orientated ones. On the other hand, it adds the significant aspect of considering the value of an innovation to the discussion. The shortcomings of both scales lead to the need of examining the transferability of the before mentioned items to enable measuring innovation performance of work teams.

The findings from literature, in particular the many different indicators proposed will be the base for the development of a measuring tool suitable for the setting of this research. The development of this instrument will be discussed in detail in chapter 5.

4.9. Limitations and conclusion of structured literature review

While other authors are focussing on identifying factors influencing the innovation level (e.g. Anderson et al (2014, p. 235)), the objective of conducting this review was to present a comprehensive overview over relevant advances and tools to answer the question whether the vast and still growing literature on innovation covered all aspects of measuring. However, this review has a number of limitations. First, the research was conducted by a single author which always inhibits a risk of bias in the selection of search strings, the coding process and the classification scheme. Second, following the systematic review approach (Pittaway et al., 2004, pp. 139–140), the approaches selected in this paper are based on well-cited articles or recent articles from mainstream management journals. They might by subject to biases such as paradigms in concept or methodology resulting in omitting approaches described in less well-known journals or those offered as business services. Third, the approaches on measuring innovation are so widespread and the literature so manifold and growing, that the same Scopus search today gives additional results and this review of 36 approaches is limited and does not represent the entire literature and methods available on innovation measuring today. By focussing on well-cited articles, university textbooks and well-known business methods mostly based on an academic background, and by using a systematic approach combined by a final check using Google to minimize the risk of author's bias, this study intends to be a good representation of existing approaches.

To conclude, the range and variety of approaches described in this review allow to measure innovation on nearly all levels of analysis. But it can be stated that a notable shortcoming could be identified: a specific framework to evaluate innovation on the work team level, especially with regard to non-R&D teams, is missing. To achieve the research's objective of measuring the innovation performance of organisational work teams the possibility of transferring existing or developing new and fitting indicators has to be evaluated.

5. Instrument development

As stated above (see chapter 2.2) the method used in this study to develop measures for an organisational team's innovation performance will follow Churchill's general design involving amongst others item generation, purification and development of a preliminary instrument (Churchill, 1979, p. 66). The whole development process was based on the findings from the literature review as described in chapter 4. Due to the fact that no measuring approach fitting to the setting of the research could be identified (see section 4.8), it will be evaluated whether elements of existing advances can be transferred and used in the planned quasi-experiment. This process consisted of four steps:

- Firstly, the set of indicators with a possible relevance in this setting were generated by selection of items identified in literature review (section 5.1)
- The fit of these indicators was assessed based on quality criteria for measuring instruments and the insights gained by the expert interviews (section 5.2)
- The two instruments for measuring innovation performance of teams in the research's setting were created (Counting tool: section 5.3; Valuing tool: section 5.4)
- The concept of utilising them was developed (section 5.5)

5.1. Generation of measuring items

Considering paradigms for scale development (Gerbing & Anderson, 1988, p. 187) this stage involved the generation of an inventory of items that could be used to measure the essence of a team's innovation performance. As described in chapter 4 and the overview over measuring approaches summarized in Table 9 the literature review resulted in an a large set of indicators currently used to measure innovation.

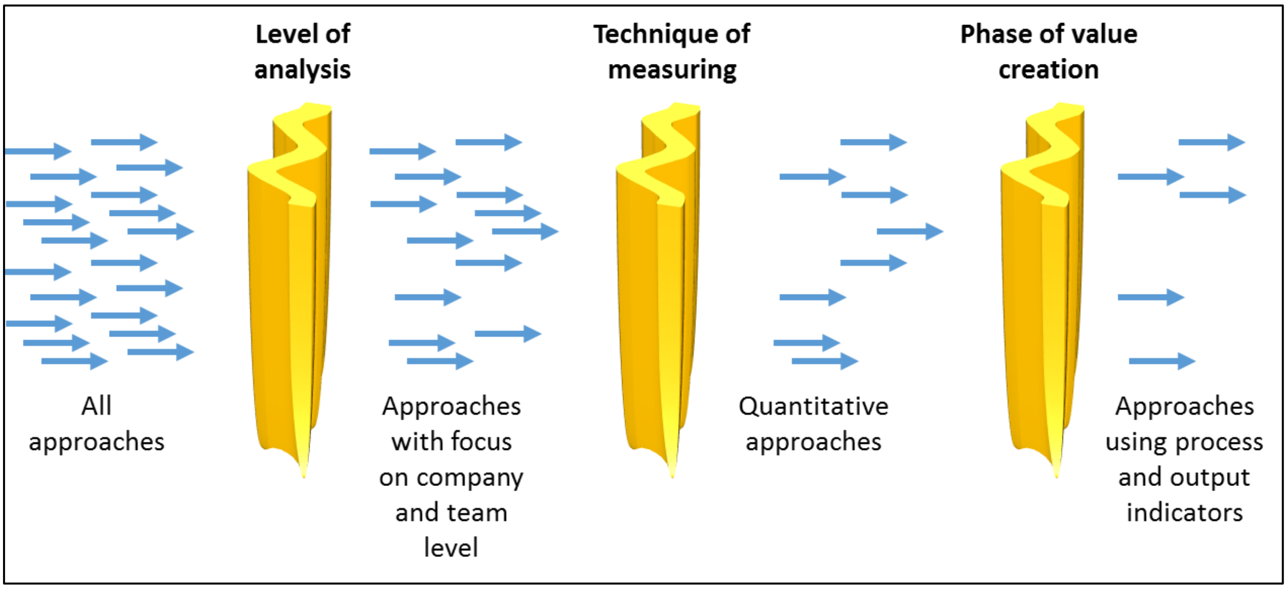

Figure 2: Process of generating pool of relevant items

The relevance of items to generate a pool of indicators to measure innovation performance of teams is assessed based on three filters: level of analysis, technique of measuring and the phase of value creation (see

Figure 2). With regard to level of analysis approaches with focus on company level should be assessed in more detail. Jointly with the existing items from team level the likelihood of a potential fit is higher compared to country or individual level. The individual level is concentrated on persons' behaviour. Behavioural focussed measuring of individual strengths differs significantly in method and process from evaluating organisational units such as teams or companies. On country level manifold heterogeneous organisational units are included in the assessment as well. Items describing infrastructure and enabling factors, such as legal regulations or availability of internet cannot be transferred to measure innovation performance of work teams. Indicators also fitting on organisational unit could be found particularly with relation to the outputs of innovation activities (see section 4.3). These indicators are also proposed on company level, so it will not be necessary to take the country level into account during the course of the discussion.

Measuring innovation performance of work teams supports management decision making. For example based on the actual level of a team's innovation performance activities to improve the team's innovativeness might be taken or incentives granted. This presupposes an objective evaluation of the situation. Generally speaking three different techniques to measure an indicators' value are used: quantitative, semi-quantitative and / or qualitative. As described above (see section 4.2) semi-quantitative techniques are basically qualitative judgements that are converted to numbers. They differ from quantitative technique in that no attempt is made to use a sophisticated formula to complete the data. Qualitative techniques are intuitive judgements ((Pappas & Remer, 1985, p. 15)). Due to the decision makers interest in unambiguous results quantitative measures should be used. Only if this is not possible, one should also analyse semi-quantitative approaches in more detail.

The approaches measuring innovation performance on company and business unit level use indicators to describe the input to facilitate innovation, the process the innovation is realised and the output generated (Goffin & Mitchell, 2017, p. 318). Input aspects are usually investments facilitating the innovation, e.g. by providing well-equipped laboratories or offering innovation workshops. Process aspects ensure a professional approach to achieve the innovation, e.g. staying in time and budget while creating a new product, thus a vice versa relationship exists. Assessing the output implies evaluating the value of the results of the innovation stages, such as ideas created or product implemented (Tidd & Bessant, 2014, p. 289). This should be linked to the three stages of innovation to assess which indicators can be used for measuring a team's innovation performance (Figure 3).

Due to the scope of the definition of innovation performance input indicators are not focussed on in the further progress of this study (see chapter 3).

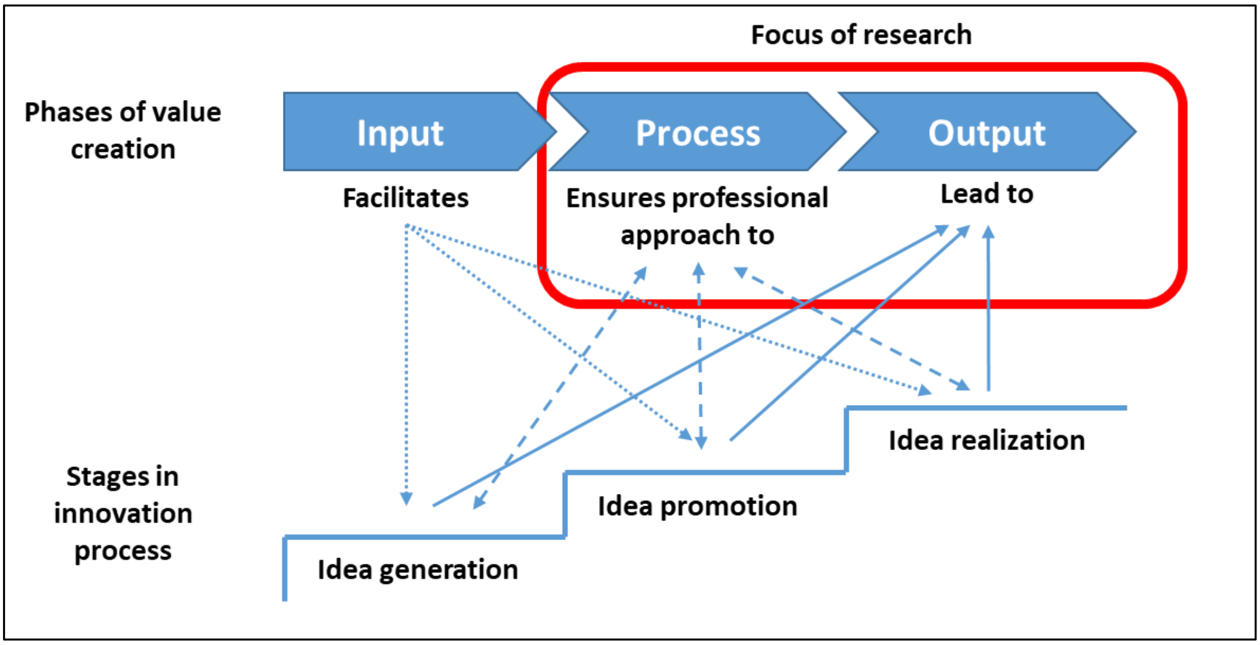

Figure 3: Relation between phases of value creation and innovation stages

To sum up, the pool of process and output items should be generated by the approaches focusing on company or team level and using a semi-quantitative or quantitative measuring technique. The following approaches fulfil these criteria: the bean counting approach, the Input-Process-Output-Outcome (IPOO) approach (Möller et al., 2011), the advances from a job centre (Burgess et al., 2012), Tidd & Bessant (2014), Fuchs (2014) and Goffin & Mitchell (2017) (please refer to appendix table 2). In a first step, the indicators proposed by these were analysed in detail (please refer to Appendix C)). Secondly, the indicators linked to process and output phase were clustered by their innovation stage (Table 10). Similar or identical indicators of different authors were only counted once. Indicators being suggested for more than one stage, e.g. development man hours per completed innovation were counted in each stage.

Five process indicators were proposed in literature to assess the idea generation stage of the innovation process, e.g. “the number of problem solving teams or “per cent of projects killed too late”. These items are characterized by the fact that they are covering all three steps of the innovation process. There is no indicator specifically created for evaluating the process in the idea generation stage. The two items mentioned as output indicators in this stage are quite alike: “number of new ideas” and the correlating ratio “number of new ideas per employee”. They were used to show the level of creativity mirroring the stage in the innovation process.

Four additional process items could be assigned to idea promotion stage, e.g. the number of overruns in development time and process average lead time for introduction. All four were also used for the realization stage, so for this cluster no items were specifically developed as well. Amongst all indicators on company and project team level, none could be identified measuring the output achieved during this stage. This is a significant gap requiring further discussion in the course of this research and will addressed in section 5.2.3.

Items proposed for the idea realization stage are manifold, 18 elements could be found for the process and the output phase each. Next to the nine process indicators mentioned above for the other two stages, additional ones such as quality performance or number of complaints exist. All output possibly derived from innovations was described by the literature as well ranging from implementing the idea (e.g. “number of prototypes” or “number of new products”) to the related economical impact (e.g. “ growth in revenue by innovations”, “earnings from patent licensing” or “growth in market share”).

To sum up, the analysis shown in Table 10 displays the limited number of output indicators for the idea generation and promotion stages and makes a broad range of indicators with regard to the idea realization stage and the underlying process visible. Particularly the lack of output indicators in the promotions stage requires further discussion in the course of this study.

	Process Indicators	Output Indicators
Idea generation	5 indicators: • Development men-hours per completed innovation • Number of problem solving teams • Per cent of projects where post-project reviews are conducted • Per cent of projects killed to late • Number of improvement to innovation projects	2 indicators: • Number of new ideas • Number of new ideas per employee
Idea promotion	9 indicators: • Failure rates in development process • Number of overruns in development time • Time to market • Process innovation average lead time for introduction • The five indicators also related to idea generation	0 indicators

	Process Indicators	Output Indicators
Idea realization	18 indicators: • Log job entry productivity • Business delivery target • Quality performance • Quality of service to job seekers • Quality of service to firms • Customer satisfaction measures • Savings accruing per worker • Cumulative savings • Number of complaints • The 9 indicators also related to idea promotion	18 indicators: • Number of patents received • Number of (technical) publications / quotes / speeches • Honors/awards from peer groups • Implemented improvement ideas • Number of new products introduced • Percentage of sales / profit from new products / services • Cost / selling price of product • Market share • Growth in revenue / market share derived from innovation • Higher value added • Improved profitability • Cost savings by process innovations • Quality improvements by process innovations • Return on innovation investment • Earnings from patent licencing • Number of new products compared with total number of product in the portfolio • Number of process innovations / new products / new services (compared to competitor) • Number of prototyps, testings, lines of code

Table 10: Overview over process and output indicators proposed by existing approaches

5.2. Purification of measuring items

In efforts to purify the instrument, multiple items were taken out and the list of items was elaborated on for clarity and appropriateness by two additional steps: application of criteria for measuring instruments and getting expert insights by structured interviews.

5.2.1. Purification by applying quality criteria for measuring instruments

This section discusses whether the indicators to measure innovation performance proposed by todays literature are sufficient. The number of indicators is not linked with the quality of an approach measuring innovation performance in work teams. The indicators used for measuring the innovation performance of a company or business unit are manifold and not all can be possibly transferred on work team level. Due to the particularities of the innovation process approaches to measure innovation should fulfil specific requirements (Vahs & Brem, 2015, p. 328). Generally speaking, evaluation criteria should be:

- coherent with the (strategic) goal of the approach,
- minimize the effort of data gathering and
- provide an objective picture of the situation (Gleich & Schimank, 2015, p. 57; Goffin & Mitchell, 2017, p. 315)

As described above, measuring innovation performance of work teams supports decisions on activities to improve the team's innovativeness. This requires an objective and ideally prompt picture of a team's actual innovation performance level. The analysis of the output indicators of the realization stage shows that some of them are covering a long time period. For example, an "increase in market share" by an innovative product or the "earnings from patent licensing" will only be known long after the innovation was realized. Even though different terms were used by different authors, within the literature it was differentiated between the innovation's concrete result (e.g. a product, prototype, change in process) and its effects (e.g. gain in market share, cost reduction) (Hauschildt & Salomo, 2016, p. 344-345; Tidd & Bessant, 2014, p.289). To facilitate decision making and fulfil this strategic goal, it can be assumed that only indicators describing the actual result are relevant in this context. Hence, outcome indicators cannot be regarded as aligned with the strategic goal and will therefore be excluded in the course of this work.

In the idea generation and promotion phase indicators such as development man hours per completed innovation (Goffin & Mitchell, 2017, p. 316), number of new ideas (Fuchs, 2014, p.40; Pappas & Remer, 1985, 18), or failure rates (Tidd & Bessant, 2014, p. 289) can be measured without major delay. Thus, all mentioned items are in line with the strategic goal. As discussed above, outcome indicators focusing on the effect of an innovation such as market share derived from innovation (Tidd & Bessant, 2014, p. 289) or cost savings by process innovation (Goffin & Mitchell, 2017, p. 349). Hence, with regard to output indicators of the idea realization stage, only indicators describing that the innovation is tested or realized will be considered. Some of the indicators were defined as ratios, e.g. "number of ideas per employee or new products compared to existing products in portfolio". To clarify the discussion, just the underlying criteria they are composed of will be discussed in the progress of the paper, thus in the stated example "number of ideas".

Secondly, the effort for data gathering could obviously be minimised by reducing the number of indicators. As shown in Table 10, the number of process indicators proposed is high. However, they will not all fit to every situation, e.g. quality of service to job seekers is only applicable in the very specific context of a job agency, time to market usually requires the introduction of a product or service. But they can be clustered in the categories linked to time (e.g. time to market, overruns in time), cost (e.g. overruns in costs) and quality (e.g. number of complaints) aspects. Even though most authors described their measuring approach using all of these indicators (Goffin & Mitchell, 2017, p.316; Tidd & Bessant, 2014, p.289), this cannot be followed. To account for the potentially difficult and long process from idea generation until final implementation one should better select a minimum number of indicators depending on the situation but ensure that all aspects of the magic triangle of project controlling (time, cost and quality) (Wegmann & Winklbauer, 2006, p. 32) are covered.

Thirdly, the items selected should proved an objective picture of the situation. The lack of output indicators in the idea promotion stage as well as the little number of indicators in the idea generation phase indicates a lack of objectivity while measuring innovation performance in work teams as proposed by the literature. While developing the measuring tool, it will therefore be necessary to define additional indicators. The planned expert interviews might be a potential source for the required information.

To sum up, purification by applying the quality criteria for measuring instruments implied that:

- Only output (and not outcome) indicators should be considered, thus, indicators describing that the innovation is tested or realized, to be aligned with the strategic goal of supporting short term decision making
- The number of indicators should be minimized to reduce effort of data gathering, in particular, the new approach should allow that the selection of process indicators is based on the magic triangle (time, cost, quality) but is flexible depending on the specific situation
- Additional indicators are required, particularly in the stage of idea promotion, to ensure an objective picture of the situation.

In effort to put forward an operational reliable and applicable approach, the construct, item generation and item purification was validated by five qualitative expert interviews.

5.2.2. Purification by expert interviews

In efforts to further purify the instrument the researcher sought to get additional insights in measuring innovation by expert interviews as described in section 2.3.

The interviews started with documenting the particularities of the interviewee's situation and his personal background. Taking this information into account allowed to transfer the appropriate and fitting ideas and insights on general research, in particular to measures and approaches to be developed later (Kubicek, 1975, p. 19). It was also required to validate the selection of the persons based on the criteria mentioned above.

The main focus of the interview was put on the second part, measuring innovation. The general logic followed the method to order the questions from "general to specific" to facilitate answering the first "simple" questions, engage the interviewees and motivate them to continue (Pew Research Center, 2018, p. 1). The first couple of question attempt to obtain additional ideas on innovation measuring which were not identified by the literature review (e.g. "Which approaches do you know to measure a team's innovation performance"). The following questions integrated academic know how into the discussion, such as the three step innovation process and the differentiation between output and process indicators. The interviewee's personal preferences on key performance indicators were obtained by a ranking question. To further validate the concepts of "idea counting" and "process indicators" relating hypotheses were presented to the interviewees. Based on these theory driven aspects the framework contained three concrete examples of incremental innovations for rating, to test and further refine the application of the ideas discussed, in particular by asking for information on objectivity of the rating ("…that every rater would come to the same result for the innovation performance value."), definition of indicators (e.g. "output", "budget" or "quality") and weighting of different indicators within a rating scale (e.g. "Which approaches do you know within controlling to weight indicators?" or "How would you bring these output indicators together to define a single innovation performance score?”). This part closed by addressing the question of an innovation’s value and the importance of taken it into account by the measuring logic (e.g. “Which ideas do you have to evaluate the benefit of an innovation?” or “…the likelihood of very valuable ones [innovations] is higher while creating more innovations. Do you agree with this statement?”). This questions were included due to the fact that depending on the interviewees position on this aspect, the character of a potential tool might change significantly.

Thirdly, the framework covered the topic applicability of the measuring tool. To facilitate dissemination the application of the newly developed tool should be as easy and comfortable as possible. An analogy to the application of common controlling tools, not only focussed on innovation controlling, was built to identify the main requirements and a scaling question (“On a scale from 1 to 10 are they fulfilled?”). To challenge the interviewees preceding answers and to be as specific as possible also the questions “Would you use such an approach in your business unit? Why? What could be improved?” were set up. For the complete questionnaire please refer to Appendix B).

The results of the interviews with the five experts in innovation management were structured using a thematic analysis matrix (see chapter 2.3.2) to enhance the knowledge on the topic.

The interviews’ focus was put on measuring, so important aspects and key indicators were discussed intensively. Based on the experience of three interviewees, the number of ideas or initiatives is of high relevance to evaluate innovation output, even though the ideas might not have been implemented. Furthermore, economic aspects, such as “EUR paid as reward” (interviewee 1 and interviewee 4) or “an increase in market share” (interviewee 4) were added. While these items have relevance to the stages of generating and realizing innovations, the promotion stage of the innovation process was not covered. This was aligned with the results of the literature review (see section 5.1). With regard to integrating the team’s competence in managing the innovation process

/ project, only interviewee 2 did not feel any necessity. The other interviewees proposed to include process items into the measuring as well.

These results were confirmed by the assessments of exemplary indicators' relevance which was conducted by three of the experts (see Table 11). The course of the interview did not allow to conduct the rating in the other two interviews. To sum up, "number of ideas", "patents" and "percentage of projects finished within time and budget" were rated top. Even interviewee 2 agreed on relevance of the process indicator "projects finished within time and budget". In addition, the indicator "cost savings from process innovation", which gives a clear value to the output of the innovation also achieved a top rating by two out of the three experts.

Indicator	Interviewee 1	Interviewee 2	Interviewee 3
Number of ideas	T(op)	T	T
Number of patents		T (for Service & R&D)	T
Number of positive decisions with regards to innovative proposals			
Honors / awards from peer group	L(ast)		T
Market share gained by innovation		T (for Service & R&D)	
Cost savings from process innovations	T	T (for production)	
Failure rates while attempting something new			
Time to market		T (for production)	
Percent of projects finished within time and budget	T	T (for production)	
Development man-hours per completed innovation			
Quality performance	T		
Number of publications	L		

Table 11: Rating of exemplary indicators

Two of the experts project highlighted the importance of measuring on individual level and the relevance of the ICB. Assumingly, this was caused by their current involvement in the Fincoda project and does not cover the scope of this study. Additionally, the sum of rewards paid was also regarded as practical indicator from two business experts, due to the fact that it gives a clear indication of the number and value of ideas created. Of course, this will only work if an innovation reward scheme is established in a company.

To get additional insights into rating innovativeness, the experts were asked during the interview to assess the teams' innovation performance of three examples (see question 8, Appendix B)). While all experts already assigned points for starting with an activity of the innovation process, the majority awarded ½ point for beginning but not finishing the steps of the innovation process, and another ½ points for finishing it. Only interviewee 2 were focussing strongly on finishing the generation and realization stage by granting less value to the beginning of a step (0.2 points). But he had a different view on the promotion stage: there, a full point should also be rewarded if the dismissive decision were not in the responsibility of the team. It can be assumed that determining the responsibility for failure will be difficult to assess in daily practice. They all agreed, that the three stages of the innovation process were of equal value, thus an additional weighting factor was not required.

Three out of five experts also see a high importance of evaluating the value of an innovation. Even though interviewee 1 deems counting ideas to be of higher relevance (see Table 12), she also rated "Cost savings from process innovations" as top indicator (see Table 11). One expert indicated that the value might be assessed by using a concept similar to the Failure Mode and Effects Analysis (FMEA) due to the fact that it has the situation of uncertainty in common. The other three interviewees felt that the value estimation should be conducted either by expert assessment. Interviewee 3 stressed that the value should be rated particularly high if the innovation improves the life of customers.

Statements with regard to the applicability of the tool put a focus on the ease of application. The instrument should be “easy, simple and reliable” (interviewee 1), should have “a high frequency of reports” (interviewee 2) and demand “little effort” (interviewee 4). In the development of the tool in the course of the study, this aspect should be particularly considered. Additional aspects like “very visual for workers” (Interviewee 2) or “able to use as a filter” (interviewee 4) were also mentioned.

Finally, the experts had encountered manifold activities to foster innovativeness in companies, ranging from incentive schemes to workshops and LEGO sets or from simple suggestion boxes to intranet tools and internal start-up labs. On overview of the results was summarized in Table 12.

Topic	**Interviewee 1: J. Marin-Garcia**	**Interviewee 2: R. Butter**	**Interviewee 3: M. Salenius-Ranki**	**Interviewee 4: M. Wiedenfels**	**Interviewee 5: M. Heutger**
KPIs/important aspects to measure: - **Output** - **Process**	- number of initiatives - EUR paid as reward - Innovation Competence Barometer (ICB, for persons) - process should be included as well, because budget and time are tight	- ideas implemented - ideas not implemented	- focus on ICB - process indicators particularly relevant, if there is a need for innovation	- EUR paid as reward - market share / turnover - process / project mgt crieria important	- counting ideas - projects finished
Experiences from rating examples	0,5 points for 1/2 step	- 0,2 points for 1/2 step - 1 by promotion if it is not my fault if the decision is not taken	always full points given	0,5 points for 1/2 step	0,5 points for 1/2 step
Weighting indicators	no weighting needed	no weighting needed	no weighting needed	no weighting needed	no weighting needed
Evaluation of innovation value	- ROI or expert assessment - counting more relevant	Valuing innovation important, e.g. by concept like FMEA	- commercial advantages assessed by experts - if possible, evaluate whether innovation improves the life of the customer	In favor of evaluation approach, because value of innovation is very important. counting is not enough	trend radar

Topic	Interviewee 1	Interviewee 2	Interviewee 3	Interviewee 4	Interviewee 5
Important aspects of applicability	easy, simple, reliable	time line, high frequency and very visual for the workers		- little effort - should be able to use as filter between valuable and useless ideas	
Measures relevant in praxis	- tool in intranet - incentive	- suggestion box (almost everywhere) - workshops (1-3 days)		Incentives (inventor participate on commercial success by law)	internal start-up lab, workshops, whiteboard, LEGO-Sets, open space, close customer contact

Table 12: Structured results of expert interviews

To sum up, the in-depth knowledge of the experts supported the purification of the literature review's results in particular with regard to the following aspects:

- Process indicators are of high relevance while measuring a team's innovation performance. "Counting ideas" was seen as one of the most important items to measure innovation output. Even though it was created to evaluate innovativeness on the individual level, the Innovation Competence Barometer was brought into discussion as supporting tool for measuring innovation performance of teams.
- When reaching the next step of the innovation process, each step should be granted the same value while rating (either ½ or full points) and all stages are of the same value, so a weighting factor is not required.
- Assessing the innovation's value is of high relevance, the idea of using a concept alike the FMEA was raised in the discussion.
- The application of the instrument should be easy to handle and without much effort for data gathering. The experts argue in alignment with the criteria for tool quality as discussed in section 5.2.1.

5.2.3. Implications for development of instrument

The main results of assessing the criteria for measuring instruments and from the expert interviews with regard the fit to the specific situation of work teams were:

- Various indicators are used by the proposed advances. Due to the fact that none of the existing framework can be used solely to measure innovation performance of work teams, an adequate combination has to be sought.
- Indicators can be found to evaluate the output and the underlying process of the three stages of innovation. The overview of the indicators shows the existence of a limited number of

output indicators for the idea generation and promotion stages and a broad range of indicators for the idea realization stage and with regard to the underlying process. Due to the fact that it was mentioned twice during the interviews, the items of the Innovation Competence Barometer (Fincoda Project, 2018, p. 1) will be taken into consideration while closing the gaps in required indicators.

- Process indicators should not be generally fixed but determined depending on the situation to reduce the effort to collect the required data. However, one must ensure that the aspects time, costs and quality and the expected difficulties in the realization process are covered.
- The manifold indicators for idea realization are
  - addressing either the final product (e.g. implemented ideas) or a step to prepare the final product (e.g. prototype built). The experts already awarded points for starting a phase and thus supported the literatures view.
  - focusing either on counting the realization (e.g. patents, implemented ideas) or assessing the innovation's value (e.g. yield, growth in revenue, profit)
- Particularly the business experts pointed out the importance of both alternatives with regard to the last aspect: while counting gives clear and objective results, the value of the innovation is also an important decision criterion which should not be left aside. From their point of view, counting gives an objective picture of the situation, but it does not consider the value created by the innovations. It were not comprehensible that creating and implementing a new ground-breaking product will be equally important as implementing a small process optimization saving 10k EUR per year.
- Indicators were generally regarded as equal in weight.
- A measuring instrument should be easily applicable

To sum up these aspects implies that the counting of innovations along the innovation process is as important as determining their value for the company. Secondly, additional items are required to ensure objectivity and close gaps. To ensure the ease of application in daily practice, the decision was taken to bring forward two approaches for measuring an organizational team's innovation performance: one based on counting the innovations (Counting tool) and one aiming to integrate the value of the innovations (Valuing tool). Splitting the focus of value and number into two instruments allows application according to the decision makers preferences. In cases that both aspects are required a joint application is also possible. The following sections elaborate on the detailed description of the tools.

5.3. Approach 1: Counting innovations

'The number of ideas and the number of good ideas therein are usually strongly correlated' (Thompson & Choi, 2006, p. 173).The result of Thompson & Choi's study implies that counting innovations is an efficient way to identify innovative teams. Based on the considerations addressed in the preceding section 4.8, section 5.2.1 and section 5.2.2 the existing instruments are not sufficient. In particular, the identified items (see Table 10) do not provide an objective picture of the output and are too abundant with regard to describing the process. Thus, on the one hand following the experts rating of examples appropriate items should describe every stage of the innovation process and should also give credit to preparing the finalization of each stage. On the other hand, the number of output and process indicators has to be minimized.

To reflect the first aspect and to ensure alignment with the strategic goal (see section 5.2.1) output indicators of the realization stage will only be taken into consideration if the innovation is tested or implemented. Transferring this idea to the other stages of the innovation process, the instrument fulfils the criteria if every stage is described by one item with regard to preparing/testing/beginning the stage and another one when finishing it. Using only one item per the stage's phase (thus two per stage, one for the preparation and one for the finalizing phase) ensures to minimize the effort of data gathering. The concrete items will be identified from literature along the steps of the innovation process and the requirements and input described by the interviewed experts.

The idea generation stage will be finished if an idea is created. This should be counted by the item "number of ideas generated" which was proposed in the literature (e.g. Fuchs, 2014, Goffin & Mitchell, 2017) as well as mentioned by the experts. To ensure that the new advance provides an objective picture of the situation, it can be recommend to also collect "hints/proposals for settings to be optimized" during the idea generation stage. Following the idea of the lean six sigma approach one begins with a deep analysis of the situation while it is required not to think in solutions. This step is essential to find unbiased and optimal solutions while progressing with the project (Meran, John, Staudter, Roenpage, & Lunau, 2013, p. 56). Transferring this idea to measuring innovation performance on work teams it might be the case that the team identifies a problem but needs experts to create an idea for a solution. By just measuring ideas generated this innovative input would be neglected. Still the team should be regarded as innovative, thus this criterion should be added to the list of indicators.

As discussed above (see section 4.7), a serious gap in research could be identified with regard to measuring innovative output of the promotion stage. This has to be closed to provide an objective picture of the situation. Thus, two additional indicators are required, one for the preparation and one for the finalization phase. With regard to the latter, situations exists in which teams create innovative ideas and manage to get a positive decision for implementation but the idea is not realized, e.g. due to a lack of IT capacity. Based on the definition proposed above, this team shows a higher innovation performance compared to teams just creating ideas without getting a decision. Giving credit to this work team would be in analogy to indicators for the individual level (see section 4.6). Scott & Bruce (1994, p.607), the Innovation Competence Barometer of the Fincoda project (Butter & van Beest, 2017, p. 33) and other authors describe items such as “promotes and champions ideas to others” or “convince people to support an innovative idea”. The result of successfully conducting such a behaviour will be either positive feedback or the positive decision which would finalize the promotion stage and would allow to start the implementation of the idea. Thus, transferring this idea to the setting of this research could close the identified gap in research by defining the item “number of ideas with positive decisions reached” for measuring the finalization phase of the promotion stage. Due to the fact that receiving positive feedback for an innovative idea will prepare the stage for achieving a positive decision, it can also be assumed that “number of ideas with positive feedback” will be the fitting indicator to describe the related preparation phase.

The process ends with the implementation of the innovation. Output indicators as proposed by the literature have strong focus on new products, services or processes (see Table 10). The implementation of new organisational methods, such as changes in business practices or workplace organisation, can also be seen as innovations (OECD & Europäische Kommission, 2005, p. 46). Thus, also introduction of methods such as knowledge management, which might not have a specific measurable effect on products, services or processes should be taken into account. All four of these can be summarized by the term “idea”, which will reduce the complexity of the instrument and is in line with the wording of the other items. Thus, the indicator for the finalisation phase of the realization stage will be defined by “numbers of ideas implemented”, a definition also proposed by the experts (see Table 12). Interviewee 2 also added, that ideas not implemented should be valued as well. This is in line with some authors who proposed indicators describing the not yet complete implementation of the innovation, like number of prototypes (Vahs & Brem, 2015, p. 644) or testing (Möller et al., 2011, p. 48). These suggestions should be seized while defining the fitting item for the preparation phase of the realization stage. Generally speaking, notably in large innovation projects, the innovation is tested before final implementation, so the new measuring approach will also include the indicator “number of ideas tested”.

An innovation’s output like publications or quotes will not necessarily only occur in the outcome phase, which is not covered by the planned approach (see section 5.2.1). These do also not really fit to the indicators as defined above due to the fact that they usually are not directly linked to the innovation. It is therefore proposed to summarize them by the indicator “additional inventive value shown”. Even though the literature categorizes the group of indicators proving an additional inventive value within the output phase (Tidd & Bessont, 2014, p. 289), these indicator might also be categorized as result of the idea generation stage. They do not necessarily represent the realization of the innovation but might also be on a theoretical level. Nevertheless, achieving one of these indicators proves a high validity and quality of the innovation so the literature’s categorization will be followed in the course of this study.

Finally, the use of context-specific process indicators was proposed by the literature and the experts. These should be based on the magic triangle to ensure that an objective picture of the situation and potential difficulties in realisation are represented. This also implies that the number of potential indicators is not limited to the indicators mentioned in the evaluated approaches, so depending on the situation, additional indicators can be used. Especially with regard to the process during the idea generation stage team members must embrace a creative process of taking risks, experimenting, and frequently experiencing failure. Thus, one should also think of extend the list to indicators derived from creativity measuring such as mistakes or failures (Thompson & Choi, 2006, p. 110).

To conclude, the following indicators will be used to measure an organizational teams innovation performance by counting ideas:

- Idea generation (output):

  Preparation stage: number of problems identified

  Finalization stage: number of ideas generated

- Idea promotion (output):

  Preparation stage: number of ideas with positive feedback

  Finalization stage: number of ideas with positive decisions reached

- Idea realization (output):

  Preparation stage: number of ideas tested

  Finalization stage: number of ideas implemented

  Additional inventive value shown (e.g. published articles, filed patents)

- Process quality indicators (process):

  Selected indicators depending on specific situation covering the aspect time, budget & quality of the innovation process.

These indicators cover the preparation and finalization step of all stages of the innovation process. Some indicators addressed in literature, such as filed patents (Fischer et al., 2015, p. 644) or scientific publications (Goffin & Mitchell, 2017, p. 302) are not linked with these steps. Even though they are unlikely to have practical relevance in work teams outside R&D, they are also included for the sake of completeness and transferability to all kinds of teams. To do so they were grouped into the cluster "additional inventive value achieved". An overview over the items of the Counting tool and the reasoning beyond it is summarized in Table 13.

Stage	Indicator (What will be measured?)	Explanatory statement (Why does the item fit?)
Idea generation (output): • Preparation phase:	Number of problems identified	Based on the lean six sigma philosophy that the base for generating solutions lies in the identification of the problem. This innovative output must not be neglected to ensure an objective picture of the situation.
• Finalization phase	Number of ideas generated	Item was mentioned (e.g. Fuchs, 2014, p- 40) and by the experts. It also reflects the term of the stage.
Idea promotion (output): • Preparation phase:	Number of ideas with positive feedback	Transfer of items from measuring innovativeness on individual level (e.g. “promotes and champions the ideas to others”, Scott & Bruce, 1994, or Fincoda project) to ensure objective description of the situation. “Positive feedback” can be assumed as preparing the stage for a “positive decision”
• Finalization phase	Number of ideas with positive decisions reached	See preparation phase
Idea realization (output): • Preparation phase:	Number of ideas tested	Indicators mentioned in literature review, such as “number of prototypes”, “testing”, “lines of code” describe a realized output but not the final product. This can be clustered and rephrased to “number of ideas tested”
• Finalization phase	Number of ideas implemented	Rephrasing and clustering of the output indicators mentioned by literature and experts to describe the realized innovation (e.g. new products / services / processes / organizational method introduced)
	Additional inventive value shown (e.g. published articles, filed patents)	Summarized term of all output indicators mentioned in literature which prove a high validity and quality, even though the innovation might not be realized yet
Process quality indicators (process)	Selected indicators depending on specific situation covering the aspect time, budget & quality of the innovation process.	Cluster of indicator to allow that the selection of process indicators is based on the magic triangle (time, cost, quality) but is flexible depending on the specific situation

Table 13: Indicators of the Counting tool

5.4. Approach 2: Valuing innovations

Due to the fact that the value of the innovations was in the experts' focus, the second approach should be designed to assess the teams' performance by the value of on their innovations. The list of items (see Appendix C)) contains indicators describing the value of an innovation, such as yield achieved or gain in market share. These can be evaluated either (long) after introduction or estimated according to the expected impact in the uncertain future. In addition, the mentioned indicators do not cover the whole innovation process but focus on the benefits after realization. To fulfil the above-mentioned requirements a pragmatic approach was sought which provides an assessment of an uncertain future. Based on the results of the expert interviews the commonly used concept of the Failure Mode and Effects Analysis (FMEA) was selected for further investigation for its fit as a base for innovation value evaluation.

It assesses potential risks (failure modes) and their effects resulting in the Risk Priority Number (RPN) (Teng & Ho, 1996, p. 9). 'It uses linguistic terms to rank the probability of the failure-mode occurrence, the severity of its failure effect, and the probability of the failure being detected on a numeric scale from 1 to 10. These rankings are then multiplied to give the RPN. Failure modes having a high RPN are assumed to be more important and given a higher priority than those having a lower RPN' (Bowles & Peláez, 1995, p. 203).

The concept of the FMEA can be transferred to the situation of innovation measuring, because:

- The objective of the FMEA is to identify all failure modes within a systems design (DoD, 1980, p. 3), the Valuation tool aims to cover all innovations created by the team
- FMEA supports decision makers by providing results in a timely manner (Teng & Ho, 1996, p. 9). Prompt results is one of the requirements of the innovation performance measurement tool, as mentioned above.

- The FMEA's outcome - the Risk Priority Number (RPN) – allows to ranking risks (Bowles, 2003, p. 380), the new tool aims to fulfil a decision makers interest by ranking teams based on their Innovation Performance Score (IPS).

Even though the FMEA approach is assumed to have shortcomings in both the ways in which calculations are made and the ways which the results are interpreted (Bowles, 2003, p. 385), the concept is widely accepted and frequently used in daily business (Sobral, Teixeira, Morais, & Neves, 2017, p. 2). Improved concepts using for example fuzzy logic (Bowles & Peláez, 1995, p. 204) might give superior results but are not in line with the requirement "minimal effort / practicability" as mentioned above. These advances need far more effort but evaluating innovations remain 'a combination of uncertain science and experience' (Trott, 2017, p. 329). So due to its trade off of practicability and validity of the results from a managerial perspective (Bowles, 2003, p. 380) the standard FMEA approach fits best to the situation in study.

This concept can only be used if it is possible to replace the three categories of this ordinal measurement approach by ones giving an objective picture of an innovation. The need for evaluating innovations also occurs while deciding on projects within a R&D portfolio. To increase practicability the four project characteristics mentioned above (1) expected profitability, 2) technological opportunity, 3) development risk and 4) appropriateness ; see chapter 4.5) can be further condensed to the two aspects of the priotization matrix typically used by management consulting firms for project selection: impact representing expected profitability and technological opportunity and feasibility summarizing development risk and appropriateness (Geissdoerfer, Bocken, Steingrímsson, & Evans, 2015, p. 311).

Thus, if there is a fit with the FMEA concept, the criteria impact and feasibility will be used within the Valuing approach. To assess risks the FMEA is evaluating the potential severity of the risks occurring 'to provide a qualitative measure of the worst potential consequences resulting from design error or failure' (DoD, 1980, p. 15). Transferring this idea to innovations, one should choose the criterion profit increase or impact respectively instead of "severity". In addition, the criterion probability of occurrence considers a similar aspect as the criterion "feasibility" does while priotizing: a future prediction whether it (risk or implementation respectively) will happen (Bowles, 2003, p. 382). So using these criteria will be in line with the idea of the FMEA concept. Only the third aspect "likelihood of detection" has no connection with the innovation process. While other criteria mentioned in the literature such as time to implementation (Goffin & Mitchell, 2017, p. 43) or failure rates (Tidd & Bessant, 2014, p. 289) might be considered, they can be integrated in the assessment of the criteria impact and feasibility. Not only the output but also the way the results are achieved have to be taken into account while evaluating organizational teams (Goffin & Mitchell, 2017, p. 314). While the criteria "feasibility" and "impact" describe an innovation's expected output, the aspect of process is not yet covered. Thus, this study proposes to use "process quality" as 3rd aspect within the "evaluating approach" in addition to the aspects "impact" and "feasibility". These aspects' scores are derived in analogy to the FMEA concept by assessing them on a scale from 1-10. Even though the assessment should take company specific aspects into account, this work proposes guidelines to shall give orientation and aims to support consistence in rating for assessing:

- the criterion impact
- the criterion feasibility
- the criterion process quality

If we implemented this innovation, what favourable impact would it have on fulfilling our objective? The objective of an innovation will in most cases be an economic one. Thus, the impact can be either assessed by experts or a standard tool such as Return on Investment or Net Present Value might be applied. While striving to increase knowledge sharing, team spirit or quality an economic impact will be very difficult to determine. Financial forecast are, generally speaking, necessary but of limited value (Trott, 2017, p. 329). Baring this in mind, an expert assessment – also taking the comparison with the impact of the other innovations into account – can be recommended. While applying the FMEA this will also facilitate the application in daily business and is well proven (Anderson et al., 2014, p. 1318).

Even though linking the rank with a concrete value is extremely depending on the situation (e.g. team size, team role, company size), the following general financial and non-financial (indirect effects or relevance to public / customers) criteria are proposed (see Table 14):

Rank	**Impact**	**Criteria: Impact effect**
9-10	High-flying	▪ ROI / NPV far above company investment regulations and market standards ▪ Ground-breaking improvements expected ▪ Innovation will be mentioned in (nearly) all media (e.g. Mars-mission, Nobel price relevant research)
7-8	High	▪ ROI / NPV (significantly) above company investment regulations ▪ Significant improvements expected ▪ Innovation will be mentioned in some media
5-6	Moderate	▪ ROI / NPV according to company investment regulations ▪ Good improvements expected ▪ Customers / (most of the) employees will take notice from the innovation
3-4	Low	▪ ROI / NPV below company investment expectations ▪ Little improvements expected ▪ Customers / most of the employees will not take notice from the innovation
<2	Minor / none	▪ ROI / NPV far below company expectations ▪ Minimal or none improvements expected

Table 14: Guidelines for impact assessment

While there is a variety of concepts to assess a venture's feasibility, the criteria always have to be adapted to the situation. Generally speaking, 5 areas should be considered while assessing feasibility: Technical, Economic, Legal, Operational and Scheduling (TELOS principle) (Taylor, 2007, p. 1789).

- Technical:

  Has the company / team the technical expertise to handle completion of the innovation?

- Economic:

  Are there sufficient (financial) resources to bring the innovation to an end?

- Legal:

  Does the proposed innovation conflicts with legal requirements or demands from worker's council?

- Operational:

  How well will a proposed innovation solve the problems or take advantage of the opportunities identified during scope definition and how will it satisfy the requirements?

- Scheduling:

  How reasonable is the project timetable / specific deadlines? How far is the project already progressed in its planned approach. Generally speaking, the further the innovation progresses in the stages idea generation, idea promotion and idea realization, the higher the score for feasibility.

The score for feasibility should be assessed by a value from 1-10 taking the TELOS aspects into consideration. While this can be based on a detailed analysis, the author recommends to rely the score on a summarized expert evaluation to make it pragmatic and manageable in daily business and also take into account the achieved stage of the innovation process. The following general criteria are proposed oriented on the FMEA probability of occurrence (Bowles, 2003, p. 383) and the Gaussian Function (Weisstein, 2018, p. 1) (see Table 15):

Rank	Feasibility	Criteria: likelihood of implementation
9-10	Realization (almost) guaranteed	About 99 in 100 innovations will be implemented at this stage
7-8	High	About 95 in 100 innovations will be implemented at this stage
5-6	Moderate	About 70 in 100 innovations will be implemented at this stage
3-4	Low, failure likely	About 50 in 100 innovations will be implemented at this stage
<2	Failure (almost) guaranteed	≤ 0,5 (less than 1 in 2 innovation will be implemented in this situation)

Table 15: Guidelines for feasibility assessment

As described above the process criteria should be individually selected based on the typical aspects of project controlling: time, budget and quality (Wegmann & Winklbauer, 2006, p. 32). „Normal“ process controlling is used to keep the project on track, thus showing differences to plan. However, to be able to evaluate a team's innovation process score, an additional step is needed, which transforms the controlling results with regard to time, budget and quality into a single performance score.

If a project controlling is already set up to ensure the proper process of realizing the innovation, the resulting status reports should be used as a base for assessing the process quality score and assigning an appropriate rank to quantify it. If this does not exist an expert assessment can be recommend to achieve valid results by investing a reasonable effort (Anderson et al., 2014, p. 1318). The following guidelines are proposed (see Table 16):

Rank	Process Quality	Criteria: Accordance to plan
9-10	Better than planned	1 or more aspects have a positive deviation to plan
7-8	In (approved) plan	All 3 aspects (time, budget, quality) are in plan
5-6	Minor negative deviation	1 aspect not in plan, likelihood to catch up with plan; Project is currently being set up.
3-4	Negative deviation	2 aspects not in plan or 1 extremely out of plan
<2	Out of control	3 aspects out of plan or equivalent bad situation

Table 16: Guidelines for process quality

5.5. Concept of utilization

As described above these two approaches resulted from literature and the elaboration during the discussions within the expert interviews. The "Counting approach" measures a team's innovation performance by adding up the steps progressed on the way to implement the innovation without differentiating between the quality of the innovations. The concept can be utilized by summarizing the described output indicators to a total score: every time an innovation reaches the next stage 1 point is counted. Achieving a stage implies, that the stages below were also reached, e.g. generating an idea implies that a problem was observed so one should assign 2 points. To assess the "process quality score" of the innovation a scale from 1-10 is proposed. The guidelines of the valuing tool should be applied for this assessment.

The "Valuing approach" assesses the quality of each innovation by using the criteria impact, feasibility and process quality to describe the teams innovation performance score (IPS). In accordance to the concept of the FMEA, the items should be rated based on the guidelines given and than multiplied with each other resulting in a score between 0 – 1,000. Each innovation's score is added to the teams total. The exemplary use of the concepts is described in Appendix section Appendix D).

6. Hypotheses and experimental design

The research's objective is to develop a functioning instrument for measuring work teams' innovation performance. But will the developed tools be applicable and will they work properly? To create a strong proof within research, true experiments are regarded as 'yardstick against which non-experimental design is assessed' (Bryman & Bell, 2019, p. 53). True experiments are characterized by a randomized sample. To appropriately test the tools and measure innovation performance work teams had to be observed. Thus, experimenting with a randomized sample would imply to create work teams for a certain period of time which is long enough to allow development and interaction. Alternatively, by using a quasi-experimental design, one can strengthen casual inferences when random assignment and controlled manipulation are not possible or ethical (Grant & Wall, 2008, p. 659) and still maintaining internal and external validity without interrupting ''real life'' through intrusive intervention (Grant & Wall, 2008, p. 655). The ecological validity of quasi-experiments is still very strong (Bryman & Bell, 2019, p. 53). Due to the fact that the complexity of organizing a quasi-experimental design is manageable while the robustness is nearly equal and based on the developed instruments, a quasi-experimental design was chosen to test the applicability of the approaches.

During the experiment the innovation performance of selected teams will be measured by both tools in two phases: in phase 1 all teams will be observed "as is". In the 2nd phase, one out of six measures potentially able to encourage innovation performance will be implemented in nearly all the teams. After the experiment, the experiences made by the involved team managers will be collected by means of a lessons learned interview. This setting does not only allow to test the develop approaches, it also provides the opportunity to evaluate the impact of measures to foster innovations on work teams. To begin with, hypotheses are stated to describe the expected experimental results (section

6.1). In a second step, the design of the quasi-experiment is developed (section 6.2) and thirdly the framework of the lessons learned interviews with the participants described (section 0).

6.1. Experimental goal and hypotheses

While the experiment's focus is on assessing the tools, the setting also provides the opportunity to report on the activities' effects on the team's innovation performance. Thus, on the one hand the main focus of the research is to evaluate whether:

- The Counting tool is of sufficient quality to measure a team's innovation performance and whether
- The Valuing tool is of sufficient quality to measure a team's innovation performance.

On the other hand the experiment's design should allow to resume the academic discussion on factors influencing innovativeness of teams. There was extensive research on ways how innovation could be fostered in organizational teams which was summarised in the meta-analysis discussing team-level predictors for innovative work behaviour conducted by Hülsheger et al (2009). The authors stated that team process variables (such as vision or task orientation) showed a stronger relationship with innovation than input variables, e.g. team size or team longevity (Hülsheger, Anderson, & Salgado, 2009, pp. 1134–1135). One process variable identified showing a positive effect is „support for innovation" ($p=0.47$). It was broadly defined as 'expectation, approval and practical support of attempts to introduce new and improved ways of doing things in the work environment'. This could be achieved by encouraging and valuing new ideas and their implementation (Hülsheger et al., 2009, p. 1131).

The choice of organisational practices to encourage innovation is manifold, they range from creating an innovation culture (Kahlfuss, 2013, p. 129) to trainings in creativity techniques (Bergmann & Daub, 2008, 114ff). Based on the model of Warschat (Warschat, 2005, p. 11) – in analogy to Kanter (Kanter, 2006, p. 76) – the working environment can be split into 4 possible aspects:

- Innovation culture

  Values, code of conduct and attitudes of management and employees having an effect during the innovation process.

- Innovation strategy

  Description of the strategic goals and the general way they should be reached.

- Innovation structure

  Organizational structure and management systems with regard to innovation.

- Innovation process

  Structure and implementation of innovation projects / activities.

Alternative structures are also discussed in the literature (Kahlfuss, 2013, p. 129), even though they mostly consist of equivalent aspects.

In analogy to the measuring approaches described above (see chapter 4) these organisational practices were targeted on improving innovativeness either on the whole company or a business unit, on a team or an individual. Due to the study's focus on team performance, only activities on process level were evaluated for their fit with the experimental setting (see **Fehler! Verweisquelle konnte nicht gefunden werden.**). Furthermore, only activities on this level were likely to have an impact short-term and direct enough to be measurable during the experiment.

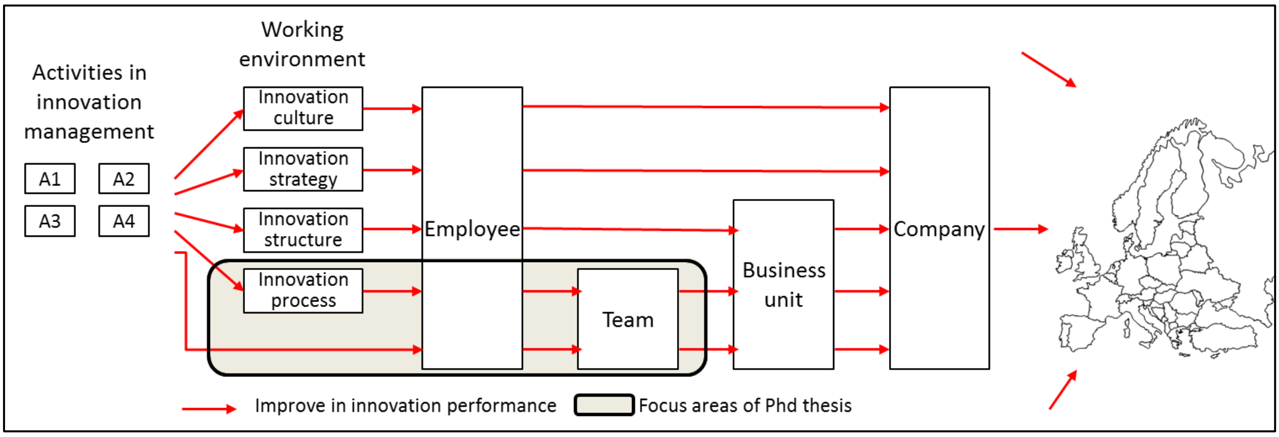

Figure 4: Process of improving innovation capacity

By selecting concrete interventions in this quasi-experimental setting, this study could bring forward the academic discussion by focussing on specific measures and organisational practices which are supposed to support innovativeness of teams. By specifying the hypothesis 'Support for innovation is positively related to innovation' stated by Hülsheger et al. (2009, p. 1131), the following unspecific but directed hypotheses were selected to assess the effect of the measures:

Hypothesis 1:

- H_0: The selected measures in total do not have impact on innovation performance
- H_1: The selected measures in total have an impact on innovation performance

Hypothesis 2 to clarifies the evaluation of hypothesis 1.

- H_0: Each of the selected measures has an impact on innovation performance
- H_1: Not all of the selected measures will have an impact on innovation performance

For both hypotheses µ1 represents innovation performance in phase 1, µ2 in phase 2. For this study, the independent variable is the measure selected while in some teams no activities were implemented. The dependent variable is the innovation performance as measured by the counting tool (dependent variable 1) or the evaluation tool (dependent variable 2).

To adequately test the hypotheses a triangulation research strategy was carried out to increase confidence into the research results (Scandura & Williams, 2000, p. 1249). The experimental data gained will be strengthened on the one hand through the application of both tools and on the other hand through the data from structured interviews with the participating team leaders. Thus, the study research consists of two steps:

- A quasi-experimental design during which the tools will be applied by team managers
- Follow-up interviews with the involved team managers to supplement the measured results

6.2. Research design of the quasi-experiment

The design as quasi-experiment was chosen. The teams in the study were existing operational work teams, which were formed over the last couple of years. Therefore, a randomization of the team members was not possible, the setting did not fulfil the criteria of a classical experimental design (Bryman & Bell, 2019, p. 51). By using a quasi-experimental design, one can strengthen casual inferences when random assignment and controlled manipulation are not possible or ethical (Grant & Wall, 2008, p. 659) and still maintaining internal and external validity without interrupting ‘‘real life’’ through intrusive intervention (Grant & Wall, 2008, p. 655). The ecological validity of quasi-experiments, in which random assignments are absent, is still very strong due to the fact that they are no artificial interventions in social life (Bryman & Bell, 2019, p. 53). Due to these facts this evaluation research study will be conducted by a quasi-experimental design while taking the potentially confounding factor of random team assignment particularly into account.

6.2.1. General design

A separate sample pretest – posttest control group design is used within this study to minimize potential sources of invalidity (Campbell & Stanley, 1967, 55):

O X O

O=observation or measurement X=exposure to an intervention

The teams in the study were existing operational work teams, which were formed over the last couple of years, thus, a randomization of the team members was not possible. The experiment was set at one site of a Top 5 German insurance company in November and December 2017. Competition is high in the saturated German insurance market. In the last couple of years the situation for established companies became increasingly difficult due to:

- The low interest period made it extremely difficult to compensate the guaranteed interest rates on life insurances with conservative financial investments, so the CFO of AXA Insurance, Mr. Harlin, stated in 2017: 'Our sector is suffering from the long-term impact of a persistently low interest-rate environment' (Mazars, 2017, p. 1).
- The raise of so called Insurtechs like Clark (clark.de) or Nexible (nexible.de) which promise customers a cheap and easy-to-use insurance experience. Based on a modern IT-infrastructure and their green-field approach without obligations to existing customers they can respond flexible to market needs. Due to their old and historically grown IT-systems companies are not able to react as quickly to the chances and threats of digitization and automating (Catlin & Lorenz, 2017, p. 6).

In reaction to this situation, managers' of established companies invest into a modern IT-infrastructure and are pushing the employees to a culture of innovation. Not only innovative products and services to enhance the customer experience and protect the customer relations are seen as essential (Mazars, 2017, p. 1) but also process innovations to reduce costs are in focus to increase competiveness. ‘Innovation enables insurers to streamline operations, transform functions, create efficiencies and develop superior capabilities’ (Bieck, Cornelius, Patel, & Uramatsu, 2016, p. 5). Innovations in this sector vary from radical ones like the introduction of the new product “cyber-insurance” or AI-based automatized claims handling in motor damage insurance to incremental ones like changes in existing products or process improvements. Even though traditionally ‘the insurance industry has been described as slow in adopting new ideas […] that is however changing with the entry of new technologies’ (Muindi, 2018, p. 3). That was the reason why the insurance sector was generally a suitable field for research in innovation. The company chosen for conducting the quasi-experiment is specifically undergoing this process. The new CEO started a turnaround programme in 2015, replacing nearly all members of the board, the top management and investing heavily in IT-infrastructure and new products and calls for efficiency increase in existing operative processes to cut costs. Consequently, all business units of the company are also in a continuous search for incremental innovations. This setting did not only fit to the planned research but also ensured the management’s support.

Within its site in Hamburg, Germany, common operative insurance tasks such as handling applications as well as claims and other customer requests, are processed by the employees. The employees are organised into teams of 10-17 persons and led by a team leader. Three until up to 12 teams are structured into a business department led by a senior manager. Educational-wise all employees conducted an education as insurance clerks. Even though the organisation underwent various structural adjustments in the last couple of years to react to the changing market environment, the composition and the tasks of the teams themselves changed very little. While workshops to improve customer experience exists on a quarterly base, there is no structured approach to valuing and implementing innovations created by the employees known in the company.

Conducting the study with operational work teams provided the opportunity to test applying the two measuring approaches in a “down-to-earth” environment with people focussed on “getting their work done” and generally not involved in innovative projects. On the one hand this should facilitate the transfer of the results. On the other hand, employees are a major source for innovation (Kurz, 2013, p. 34), so the results will be of general relevance for the academic discussion.

No recommendations for the time frame of separate sample pretest-posttest control group designs exists the author it aware of. Every experimental research is designed for the very specific situation, ranging from (Owens, Dearth-Wesley, Herman, Oakes, & Whitaker, 2017, p. 437), continuous observations for some months (Roethlisberger, 1939, p. 17) to lifelong observations (Iacono & McGue, 2002, p. 482) or even longer. The design must ensure that the evaluation research represents the situation as close as possible (Bryman & Bell, 2019, p. 57). Within the insurance company, a monthly rhythm for operative reporting was established to allow prompt responses. Its strategic goal was therefore coherent with the objective of the new instrument (see section 5.2.1). While applied in reality, it could be assumed that the innovation performance of teams would also be reported in this rhythm. To minimize interruption of the normal processes, to test the application in the most realistic environment possible and to fulfil the requirements for quasi-experimental designs (see above) the in accordance with the insurance company defined duration of one measuring phase was one month for the pre-test and another one for the post-test period. The experiment was timed for Q4 2017. Even though the month December is disrupted by bank holiday days around Christmas, the fourth quarter ensures a high employee attendance rate due to the fact that it is uncommon in Germany to go on holiday before the Christmas days. In addition, the season for a flu epidemic is supposed to be in the first quarter of a year, so the good availability of the workforce increases the likelihood of idea generation and allows team managers to cope with additional work and initiatives. Particularly the latter aspect ensured that the operative work was not disturbed by the quasi-experiment, a crucial aspect from the company's point of view.

Quasi-experimental settings like the one design for this work are at risk of confounding factors. Generally speaking, potential aspects inhibiting the risk of confounding effects are (Rack & Christophersen, 2007, p. 28):

- Issues of team selection
- Issues of selecting measures for intervention
- Issues of communication with participating team managers
- Descriptions of the template to measure innovation performance
- Issues of covert research
- Issues of further external influences

These will be described in more detail and measures to reduce the risk of confounding effects and improve the quality of the results will be discussed in the following sections.

6.2.2. Issues of team selection

In preparation to the study, operational work teams needed to be identified for participation on the study. To minimize and control confounding effects the possibility of using deep-level composition variables based on the underlying psychological characteristics (Bell, 2007, p. 596) was discussed. Due to strong company-internal agreements the involvement of the workers' council was sought. Unfortunately, the council denied the evaluation of employees' characteristics, so surface-level composition variables were used. These are defined as open demographic characteristics that can be reasonably assessed after a short exposure (Bell, 2007, p. 596). The following criteria were set for the selection of teams:

- Identical or very similar tasks
- Homogeneous composition with regard to age and job tenure

Within the German insurance industry specific wage groups ("Tarifgruppen") are agreed on between trade unions and employer's associations. Identical or similar tasks based on the job description are granted identical wage groups. To ensure a homogeneous sample the selected teams should be allocated to the same wage group(s). While the wage groups range from group 1 to group 8, it is implied that the higher the group, the more demanding the work and thus the higher the salary. In reality, the classification of jobs within the insurance company starts with wage group 3 and ranges until wage group six for normal operational work. The wage groups 7 and 8 are only available for employees with special functions or team managers. Employees earning a salary within group 5 or 6 are supposed to have customer contact per phone or writing and work on standard tasks as well as on more complex ones asking for problem solving skills and creative solutions (see sample task description in Appendix E). Due to the study's focus on innovation, this group of employees was focused on during the experiment. 33 teams in total being organized in five departments fulfilled this criterion.

In literature, several methods are proposed to determine the minimum sample size. Hair et al. stated a minimum of five times as many observations as numbers of variables to be analysed, but proposed a more acceptable sample size ratio of 10:1 (Hair, Babin, Anderson, & Black, 2014, p. 100). Other authors see sample sizes larger than 30 and less than 500 are appropriate with a sample 10 times or more as large as the number of variables (Sekaran & Bougie, 2016, pp. 296–297). Generally speaking, the 10:1 ratio can be seen as general consensus (Schreiber, Nora, Stage, Barlow, & King, 2006, p. 326).

The experiment's dependent variable is the innovation performance as measured by the tools, the independent one the activity implemented (see section 6.1), so a minimum of 10 participating teams is required.

Supporting the experimental test of the tools implies additional work for the involved team managers. It turned out that 9 teams were not available due to an extreme high workload due to a planned IT-release at the end of the year. Thus, finally, the team managers of the remaining 24 teams were asked for their interest in participating on the study and coping with the additional work. While the managers of 19 out of the potential 24 teams volunteered, one became ill for nearly the whole duration of the experiment. In the end, 18 existing operational work teams consisting of 264 employees were included in the study. Thus for the assessment of hypothesis 1 a minimum sample of 10 teams is required, which will be overachieved. Increased sample size always produce greater power for a statistical test (Hair et al., 2014, p. 10), so all these teams were included into the experiment.

But the sample size is unlikely to be enough for the validation of hypothesis 2 because 6 different measures for intervention plus a control group were planned. Nevertheless, the analysis will be conducted to check whether significant results can be obtained.

These teams were split into a group with intervention and a control group, having similar characteristics (see Table 17).

Aspect	**Teams with intervention**	**Control group**
Number of teams	16	2
Number of employees in total	235	29
Group size average (range)	14.7 (11-20)	14.5 (13, 16)
Average age	46.3 years	46.3 years
Standard deviation age	8.5	9.1
Gender	64.7% female	65.5% female
Average job tenure (years within the company)	24.7 years	24.3 years
Standard deviation tenure	9.6	10.2

Table 17: Characteristics of operative working teams

Ideally, the composition of the selected teams is homogeneous with regard to age and work experience. Gender split is assumed to not have an influence on innovativeness of teams and the size of the groups cannot be analysed due to only 1 degree of freedom (see Table 18).

Source	DG	Sum of Squares	Mean spare	F	Sig.
Org. unit	16	2409.9	150.6		
Gender	1	0.000	0.000		

Table 18: Sum of squares analysis - gender

To begin with, the homogeneity of the variance is tested by using the Levene-test. A significance level of 5% is chosen due to the fact that it is well-proven.

	Levene Statistic	df1	df2	Sig.
Age	1.184	17	246	0.278
Work experience	0.973	17	246	0.489

Table 19: Test of Homogeneity of Variances

The "p-value" of age and work experience is > 0.05, so there is no statistically significant difference in the variances between the 18 teams (see Table 19). This allows to do the ANOVA test which shows that there was no statistically significant difference in age for the teams (significance p > 0.05), however work experience differed statistically significant (see Table 20).

		Sum of Squares	Df	Mean Square	F	Sig.
Age	Between Groups	1680.7	17	98.9	1.4	0.146
	Within Groups	17602.3	246	71.6		
	Total	19283.0	263			
Work experience	Between Groups	2928.3	17	172.3	1.9	0.016
	Within Groups	21897.1	246	89.1		
	Total	24825.4	263			

Table 20: ANOVA

Looking at the single values of the teams, it can be assumed, that this was caused by team 8 (average tenure 14.1 years) and/or team 7 (average tenure 17.5 years, see Figure 5).

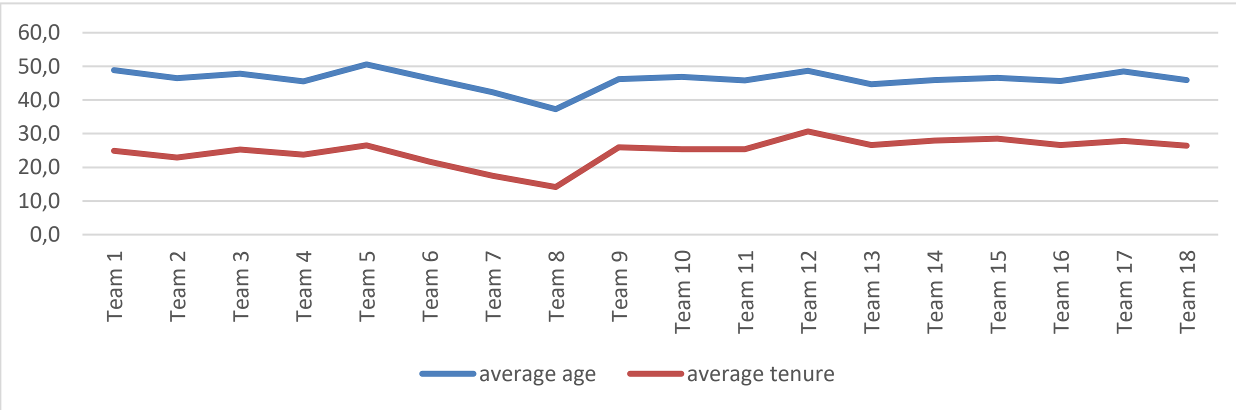

Figure 5: Overview of mean age and tenure per team

To sum up, even though the selection of teams of this quasi-experiment is not randomized, the single teams do not differ statistically significant in age, while they do with regard to job tenure. The results of teams 7 and 8 should be particularly taken care of because they might be influenced by this confounding factor.

6.2.3. Issues of selecting measures

Between the experiment's phase 1 and phase 2 the planned intervention is the implementation of a measure / organisational practice to increase a team's innovation performance. As stated above in section 6.1 the current academic discussion was resumed by focussing on specific measures on process level. Measures aiming on the innovation process mentioned in the literature were for example team workshops (Trott, 2017, p. 171), box for ideas (Laursen, 2003, p. 250), innovation management (Kahlfuss, 2013, p. 256), attending conferences regarding future trends / innovations (Goffin & Mitchell, 2017, p. 278), innovation labs / hubs (Schulz, 2014, p. 69), etc. Measures to foster a team's innovation performance should focus on creativity in the early phases of the innovation process and on supporting a systematic approach in the later steps (Kahlfuss, 2013, p. 257). In addition, the selected measures must have a fit to the environment of organizational work teams and the setting of the quasi-experiment. While the fostering of creativity had a very good fit to the situation, improving a team's systematic approach is a long-term task and difficult to achieve within the duration of an experiment if it is not there already. Due to frequent process optimization projects conducted in the last years within the insurance company, it could be assumed that the teams are used to handle projects. So implementing measures particularly supporting the idea generation were selected as well.

In conclusion, 13 potential measures were identified. Based on the abovementioned criteria "addressed step of innovation process" and "fit to setting/culture" the following six suitable ones were selected for the quasi-experimental setting (see Table 21) and the causal theoretical logic of the expected effect is described (Hambrick, 2007, p. 1351):

- Box for ideas

  Placing the box in the open-plan office is a clear visual for employees reminding them on the task to identify innovations. It might also foster the participation of "shy" persons on the problem identification or idea generation process due to the option to stay anonymous.

- 10% worktime to create and work on new ideas

  Used by innovation driven companies this measure allowed focus on being innovative without distraction or a work overload from daily tasks (Black, 2016, p. 2).

- Pin board

  Similarly to the box for ideas, the pin board can be seen as a strong visual reminder on innovation as well. In addition, it also allows discussion, further development and tracking of the implementation of the ideas by other members of the team due to the fact that the actual ideas and their implementation level is available to everybody.

- Excel-table on a shared drive

  The positive effect of the Excel-table lies in the fact, that existing ideas are visible for the team allowing further development and tracking of the ideas implementation.

- Team workshop

  Conducting a team workshop for generating innovations combines the effect of investing working time and the interactive part (e.g. discussions) of the pin board. Depending on the chosen topic, the workshop might foster idea generation as well as focussing on actual implementation of innovations

- Incentive

  reward of 1 egg surprise in case of creating or implementing innovation (max. 1 per person)

  aims to encourage employees to present and work on innovations in a playful manner

Activity	**Addressed step of innovation process**	**Fit to setting / culture**	**Remark**
Team workshop	All, depending on topic	Yes	
Idea management system	All	No	Investments and changes required which go far beyond the scope of the experiment
Box for ideas	Idea generation	Yes	“closed/individual” measure, ideas not known to public in first step
Pin board	All	Yes	“Open” measure, ideas and their status are visible and can be discussed
Mailbox for ideas	Idea generation	Yes	Measure not selected, due to the fact that it is very similar to box for ideas but not as visible
Shared Excel list for ideas	Idea generation	Yes	
Participation on congresses / trainings	All, depending on topic	No	No cultural fit with operational working teams in the insurance company (“team members never attended a congress”)
Customer survey	Idea generation	No	Duration of conducting and evaluating the survey too long
Innovation labs / hubs	All	No	Investments and changes required which go far beyond the scope of the experiment
Incentive	All	Yes	
Time for innovation	All	Yes	
Implement innovation spaces (e.g. coffee corner, table soccer)	All, mainly idea generation / sharing	No	Investments and changes required which go beyond the scope of the experiment
Stage gate process	Promotion	No	Already existent in company

Table 21: Assessment of potential interventions

The measures were allocated to the teams in the following step. (see Table 22, please refer to Appendix F) for detailed description of the selected measures):

Measures to foster innovation (implemented in phase 2)	**Number of participating teams:**
Box for ideas	4
10% worktime to create and work on new ideas	3
Pin board	2
Excel-table on a shared drive	2
Team workshop	3
Incentive: reward of 1 egg surprise in case of creating or implementing innovation (max. 1 per person)	2
None (control group)	2 (3)
Total	**18** (19)

Table 22: Measures for intervention in phase 2

Originally, three teams were selected as control group without intervention, to fulfil the requirements of the pretest – posttest control group design (Campbell & Stanley, 1967, p. 55). The measures were alloted to the teams by random. This procedure was chosen to avoid that an organisational practice particularly fitting to the workteam could be selected by the team leader. Unfortunately, the team manager of one of the teams supposed to be part of the control group became ill for nearly the whole duration of the experiment and he could not participate. So, at the end, the control group consisted of two teams. To sum up, six measures focusing on the innovation process were selected as intervention in phase 2 of the experiment. They are nearly equally assigned amongst the teams. In addition the teams 7 and 8 potentially differing in job tenure were assigned a measure and not part of the control group, so the risk of confounding effects should be minimized.

6.2.4. Issues of standardising the setting

Several steps were undertaken prior and during the experiment to ensure standardized and equal conditions in the experimental setting amongst the teams:

- Ensuring commitment and participation by continuous and standardized communication
- Equal "presence" of the measures / organisational practices in the teams
- Validating rating of results

After identifying the 24 teams suitable for participating on the experiment, the heads of the departments were addressed in a request to support this university study. While the detailed communication approach is described in the backup section (see Appendix G)), it can be summarized in the following steps:

1) Presenting information on the experiment's background to potential team managers to ensure commitment by voluntary participation
2) Presenting detailed information to participating team managers and explaining the measuring template (see section 0)
3) Sharing of lessons learned in 2nd week of experiment
4) Presenting detailed information on phase 2 to team managers
5) Providing standardised mails as weekly reminder to be sent by the team leaders to activate their teams in phase 2
6) Feedback interviews with each team manager to clarify documented information and conduct interview for evaluation (see section 0)
7) Debriefing in the teams – information about background of experiment

Secondly, the author ensured that the measures for intervention were comparably present in the awareness of the teams. The teams were all located in an open-plan office. To enable equal conditions, “visible” measures were put on the same places within the group space, e.g. the box for ideas and the pin board close to the entrance of the team’s space, the incetives on the team leaders’ desk. The author’s on site presence during the duration of the experiment allowed a continuous contact with the team leaders and availability in case of further questions.

Thridly, procautions to ensure inter-rater reliability were taken. On the one hand, clear descriptions, examples and guidelines how to conduct the ratings were provided (please refer to following section) and access to the author arranged by his continuous on site presence. On the other hand, a two step rating process was introduced. Primarily, the team leaders rated the innovations created in the team according to the provided guidelines to their best knowledge. In addition, they described the each innovation and its status of implementation. During the lessons learned interviews (see section 6.3) these descriptions were clarified and complemented when necessary. The final ratings were achieved through validation by an independent manager of the company responsible for business process optimisation projects.

To sum up, even though a field experiment is unlikely to provide fully standardised conditions, precautions were undertaken to standardise the implementation of the measures used as intervention and suppport inter-rater reliability.

6.2.5. Description of the template to measure innovation performance

The measuring templates also containing comprehensive descriptions and examples were handed over to the team managers in German language. Next to a general part including the time period and organisational unit, it had 3 parts (for the complete template see Appendix H)):

1) Fields to insert innovation performance as measured by the Counting tool
2) Fields to insert innovation performance as measured by the Valuing tool, also including place for a short description of the innovation
3) Explanations, how to use the tool including the guidelines to support a consistent rating

Within the first part the team leader was supposed to count 1 point every time an innovation reached one of the steps within the innovation process (Number of problems identified, number of ideas generated, number of ideas with positive feedback, number of ideas with positive decisions reached, number of ideas tested, number of ideas implemented and additional inventive value shown (see chapter 5.3)). To support standardisation of the ratings, definitions for innovation and each step of the process were provided. For example, ideas should only be included in the rating if they were likely to have a potential value for the team or the company (George, 2007, p. 441) and were unique relative to other ideas currently discussed in the workteam (Shalley & Gilson, 2004, p. 34) or "positive feedback received" was defined as ideas, which were regarded by the team leader or colleagues to be submitted for decision to the manager or the steering committee.

The second part contained the indicators of the Valuing tool impact, feasibility and process quality. Discussing and testing the template with team leaders showed that the scoring of an innovations impact was very difficult. Team leaders of operational working teams are lacking sound experience of developing realistic business cases. To ensure a valid scoring, the team leaders were asked to describe the innovations potential, the scoring was then done in the lessons learned interview in discussion with the researcher and putting it into context with the scoring of the other innovations.

The second part also asked for a short description of the innovation to clarify understanding and support documentation for the further process. This field as well as the indicator process quality are also required parts of the Counting tool. To facilitate the team leaders' work and increase the acceptance, both of them had only to be filled once, even though a distinction of the ease of handling of the both tools might be difficult in the users' perception. The explanations in the third part include the defined guidelines and described the use exemplarily by rating short case studies during 2 phases. The combination of continuous communication (see section **Fehler! Verweisquelle konnte nicht gefunden werden.**) and clear guidelines aims to secure an equal rating during the course of the experiment independent from the person of the rater.

6.2.6. Issues of covert research

'Research ethics are the guidelines that are utilised to ensure that, during the complete research process, effective communication to all participants and recipients of research processes and results are ensured, that all research participants are free from harm in all its forms and formats, and to ensure that researchers "actively do good" in society or societal groups.' (van Deventer, 2009, p. 46). The ethics code of the Horizon 2020 programme is used as guideline to discuss ethical questions (Houghton, 2014, p. 2) arising from the planned research design. Two aspects have to be evaluated:

- Conducting an experiment within an insurance company
- Participation of team members

To begin with, to comply with general ethical issues the research must be legal (Houghton, 2014, p. 40). The department heads within the company are part of the executive team and have the task – similar to other companies – to continuously improve their processes and the freedom to act accordingly. This also includes initiatives to identify innovations. By discussing the course of the experiment upfront and asking for support, the aspect of "legality" was sufficiently taken care of. Even though the information is not required based on internal regulations, the member of the workers' council responsible for the site's operative working teams was informed about the experiments research design, so also from a moral perspective ethical aspects of conducting the experiment were covered.

In addition, this research was focusing on human behaviour, thus the aspect humans is touched (Houghton, 2014, p. 7). Ethical issues may arise due to the fact that a covert research design is planned – the team members were not informed about participating the experiment (Spicker, 2011, p. 119). While they are aware that for example phoning behaviour is being monitored they were not aware that the team, not the individuals, were measured with regard to the team's innovativeness. On the one hand, not informing the team members was essential to avoid the Hawthorne effect and ensure a „normal" behaviour (Olson, Verley, Santos, & Salas, 2004, p. 31). On the other hand, covert research is criticized for not giving the people the chance to consent and / or withdraw effectively (Social Policy Association, 2009, p. 4). Observing the people during this experiment did not cause any risk for them. It also did not limit there freedom or hamper their work. In this case while balancing the arguments, the covert research can be conducted due to the lack of risks to participants (Spicker, 2011, p. 130). To comply with ethical standards as good as possible the team members were informed about the background of the experiment after phase 2 (van Deventer, 2009, p. 48).

6.2.7. Issues of further influences on team innovation

The results of the quasi-experiment might also be influenced by characteristics of the teams. Hülsheger et. al (2009) identified 15 team-level variables discussed in academic literature which can be clustered in either input variables or team process variables. To sum up, it could be assumed that input variables did not have a confounding effect in this setting. Based on the information gathered in the initial discussions with the five department leaders, the results of the latest employee attitude survey and the personal observations made by the author during the course of the experiment, not only that the significance of the effect was generally not proven for the variables team longevity, background diversity and task interdependence, but also did the teams not differ with regard to these variables and the variable goal interdependence. The aspect tenure of the item job-relevant diversity was already discussed in section 6.2.2. Specific focus required the potentially confounding factor team size. These are generally large and also varying ranging from 10-17 persons. Nevertheless, the existence of an effect on team performance was only proven for project and management teams, not for production teams (Stewart, 2006, p. 42). Due to the fact that the organisational work teams participating on the experiment could be categorised in the third type, mitigating actions did not have to be taken for team size as well (for details please refer to Table 23).

Team level-variable - Input	Significant effect (Hülsheger et. al, 2009)	Approach to mitigate influence on results
Team size	Yes	Team sizes were generally large ranging from 10 – 17 persons. While Hülsheger (2009, p. 1138) identified a generally positive effect this is not the case for production teams (Stewart 2006, p.42). Thus, on the one hand, the team sizes were all large and on the other hand, no effect was proven for the type of teams participating on the experiment, it could be assumed that the variable did not have an effect in this setting.
Team longevity	No	No changes in internal team structure took place for more than 7 years, so this aspect had no relevance.
Job-relevant diversity	Yes	The employees' working background was very similar due to the fact that they conducted an education as insurance clerk within the company and were working in their teams for many years. Thus, the teams equal each other in their task-related level of diversity in function, knowledge or skills. The aspect tenure was already covered in section 6.2.2.
Background diversity	No	Differences amongst the teams with regard to age, gender, or ethnicity could not be recognised.
Task interdepend-dence	No	The team members worked individually on their own tasks, so, this aspect was not relevant
Goal interdepend-dence	Yes	The team members all have a fixed salary and their individual goals, so, this aspect was not relevant

Table 23: Team innovation input variables (source: Hülsheger et al., 2009, p.1138)

A similar picture could be drawn for most of the process variables identified of having a potential influence on teams' innovation: the teams did not show different characteristics implying the risk of a confounding effect. The latest employee attitude survey (survey conducted in 2016, results published on department level) resulted in very positive feedback with regard to the items working in teams (85% up to 100% confirmation amongst the five participating departments), mutual support and trust ((82% up to 91% confirmation) and cross-departemental cooperation ((72% up to 86%

confirmation). It could therefore be assumed that influences by participative safety, cohesion and relationship conflict were unlikely. While task conflict showed no significant effect, it could also be reckoned that the standardised centrally organised communication of the teams' goals and of objectives as well as the standardised obligatory yearly feedback and assessment system (additional feedback was regarded as possible but uncommon) reduced a potential effect of the variables vision and task orientation to a minimum. Definite judgements on the extent of the team members' internal and external communication could not be rendered due to the workers' council missing allowance to conduct a detailed survey amongst the participating teams (see section 6.2.2). However, it could be speculated that no major differences amongst the teams existed due to the fact that they were all located in the same building with an open-plan office layout and all employees were likely to have social contacts to an similar extent based on their comparable backgrounds. Finally, as stated above (see section 6.2) support for teams during the innovation process was identified as influencing factor. Relevant actions were the encouragement and valuation of new ideas, recognizing and rewarding them and support during the whole process by managers, supervisors and coworkers (Hülsheger et al., 2009, p. 1131). Next to influence of the planned interventions supporting innovativeness, the team leaders' attitude towards innovation inhibited the risk of being a confounding factor during the quasi-experiment. The following mitigating actions were taken to reduce the effect:

- Voluntary participation

  The concept of voluntary participation aimed to attract only these persons with an inherent interest in innovation and the experiment to achieve a homogeneity in attitude towards the experiment.

- Standardised communication approach

  By providing standardised emails as part of the communication concept (see Appendix G), a level of encouragement in communication was ensured.

- Design of the feedback interview

  The feedback interview was designed to indicate the team leaders' attitude towards the team's innovation performance and to take this into account during the analysis (for details see below, section 6.3).

To conclude, most of the variables identified by Hülsheger et al. (2009) did not impose a risk of a confounding effect (see overview in Table 24). Mitigating actions were planned to reduce the potential effect of different support for the experiment amongst the team leaders. In addition, the aspects discussed in the section were predominantly relevant for the inter-team comparison. The chosen pre-test post-test design also allowed intra-team comparison of the interventions' effect, which could be seen as independent from these factors with the exception of the comparison with the control group. A particular focus was needed to be given to them.

Team level-variable - Process	**Significant effect (Hülsheger et. al, 2009)**	**Approach to mitigate influence on results**
Support for innovation	Yes	Measures / organisational practices used for intervention were supposed to support team innovation and were integrated in the experimental design. This variable is also covering all aspects of 'expectation, approval and practical support' (Hülsheger et. al, p. 1131). While the general environment in the company was influencing the teams in a similar manner, the team leaders' interest in innovation and support for the experiment might be a major differentiating factor. Its risk of influencing the results was reduced by voluntary participation of interested team leaders, standardised communication and taking the team leaders view on the team's innovation performance before and after the intervention into account.
Partici-pative safety	Yes	The company's latest employee attitude survey showed similar (and very positive) results with regard to mutual trust and involvement in team internal decision making amongst all teams. This aspect is unlikely to have an influence on the results
Vision	Yes	Objectives and goals were continuously and centrally communicated to all teams, no differences in level of commitment could be identified.
Task orientation	Yes	Task orientation was evident by the yearly, standardised feedback / appraisal system existent in the company. It is obligatory for all employees. This aspect was not different amongst the teams and will not be an issue during the experiment.
Cohesion	Yes	The employee attitude survey showed a positive attitude towards the team internal work amongst all participating departments. This aspect is unlikely to have an influence on this study's results
Internal & external communi-cation	Yes	The extent of variety of communication between the teams remained unclear due to the fact, that a detailed survey was not agreed on by the workers council (see section 6.2.2). However, it can be speculated that no major differences amongst the teams exist due to the fact that they were all located in the same building with an open-plan office layout and all employees are likely to have social contacts to an similar extent based on their comparable backgrounds.
Task conflict	No	This aspect shows no variety amongst the teams due to the lack of task interdependence (see above).
Relationship conflict	No	Based on the results of the latest employee attitude survey the employees appreciated the team internal cooperation on a very positive level in all operative units. Influence on the experimental results is therefore unlikely.

Table 24: Team innovation process variables (source: Hülsheger et al., 2009, p.1138)

6.2.8. Summary of quasi-experimental design and confounding factor analysis

To sum up, the quasi-experiment follows a separate sample pretest-posttest control group design. Six potentially confounding factors were identified and considered in this research: the selected operational work teams are quite homogeneous and close to a setting of randomly assigning them. Even though, a statistically significant difference in work-experience exist, the groups can be treated as 'essentially equivalent' (Koepsell, 2005, p. 3). A variety of measures for intervention was selected which were appropriate to test the tools and the measures effect and also fulfilled the conditions of the specific setting within the insurance company. Comprehensive communication and clear guidelines and examples were given to the team leaders to ensure consistent rating and not informing the team members while performing the quasi-experiment was in accordance with ethical standards. Finally, measures were taken to reduce the influence of different support by the team leaders on the experimental results. To conclude, it could be stated that the setting up of the quasi-experiment reduced the risk of confounding factors to a minimum.

6.3. Research design of the feedback interviews – process and template

After the end of phase 2 a 30min interview was scheduled with every participating team leader. It aimed to evaluate the following aspects:

- Helping the researcher to understand the content and status of the team's innovations and the tools' scores
- Finalizing the filling of the templates in case a team leader had open questions
- Obtaining the team leaders' subjective assessment on the effect of the measure
- Receiving feedback to the tools' ease of handling and the process of the experiment

To ensure the first two aspects the team leader explained all documented innovations to the researcher. Open questions were clarified and the complete filling of the template ensured. In case the measuring template was not used (e.g. due to a lack of innovations) the process of the experiment were discussed and reasons for the result sought.

A structured interview template was created to discuss the remaining aspects (for the complete template please refer to Appendix B)). The evaluation of the measures' impact three scaling and one qualitative questions were asked. The team leaders were asked to rate the potential impact on a 4-point-Likert-scale (completely agree to completely disagree) and the assessment of the team's innovation before and during the measure on a scale from 0 (low) to 10 (high). Guidelines were given to ensure consistent rating on this aspect. Qualitatively, the reasons for the measure's impact or the requirements for improving the impact were collected.

The measuring template's ease of handling was evaluated by three 4-point-Likert-scaled questions referring to the completeness of the measuring template and the effort required to fill the Counting or the Valuing tool's part of it. A field to document additional comments was included as well.

Thus, the quasi-experimental design strengthened by the structured follow-up interviews is suitable to evaluate the applicability of the Counting and the Valuing tool. In addition, using this set up will also provide the opportunity to assess the impact of the selected measures on innovation performance.

7. Results of the quasi-experiment

The description of the quasi-experimental results covers the following aspects:

- General observations of the experiment and the lessons learned interview with the team leaders (section 7.1)
- The evaluation of the tools' quality and applicability (section 7.2)
- The analysis of the implemented measures' impact on the teams' innovation performance (section 7.3)
- The evaluation of the hypotheses (section 7.4) and limitations and directions for further research (section 7.5).

Thus, the results of the triangulation research strategy (see section 6.2) were based on the results of the Counting tool, the Valuing tool and the lessons learned interview with the participating team leaders. The findings are clustered and jointly discussed in the respective topic sections.

7.1. General observations

As stated above structured lessons learned interviews were conducted with the team leaders after finishing the experiment. During the meeting with the 14 team leaders and the coordinator for the remaining 4 teams the documented innovations were clarified and experiences using the tools and implementing the measures discussed. In detail, the innovations and their actual status were described by the team leaders, open questions with regard to how to fill the template were clarified and a structured interview guideline with regard to the lessons learned was filled. A total of 112 problems were identified by the teams in both phases and for 103 of them an idea for solving them was created. Then, there is a first significant drop with regard to achieving a positive feedback (63 in total) and a second drop with regards to the following steps of the innovation process (39 and less). In addition, it could be observed that a far higher percentage of innovations from phase 1 (16 out of

28 ≙ 57%) passed through the final stages compared to innovations from phase 2 (12 out of 81 ≙ 14%, see Figure 6). This might be due to the longer time available for implementation (two instead of 1 month) or the quality of the ideas generated and will be assessed in more detail in section 0.

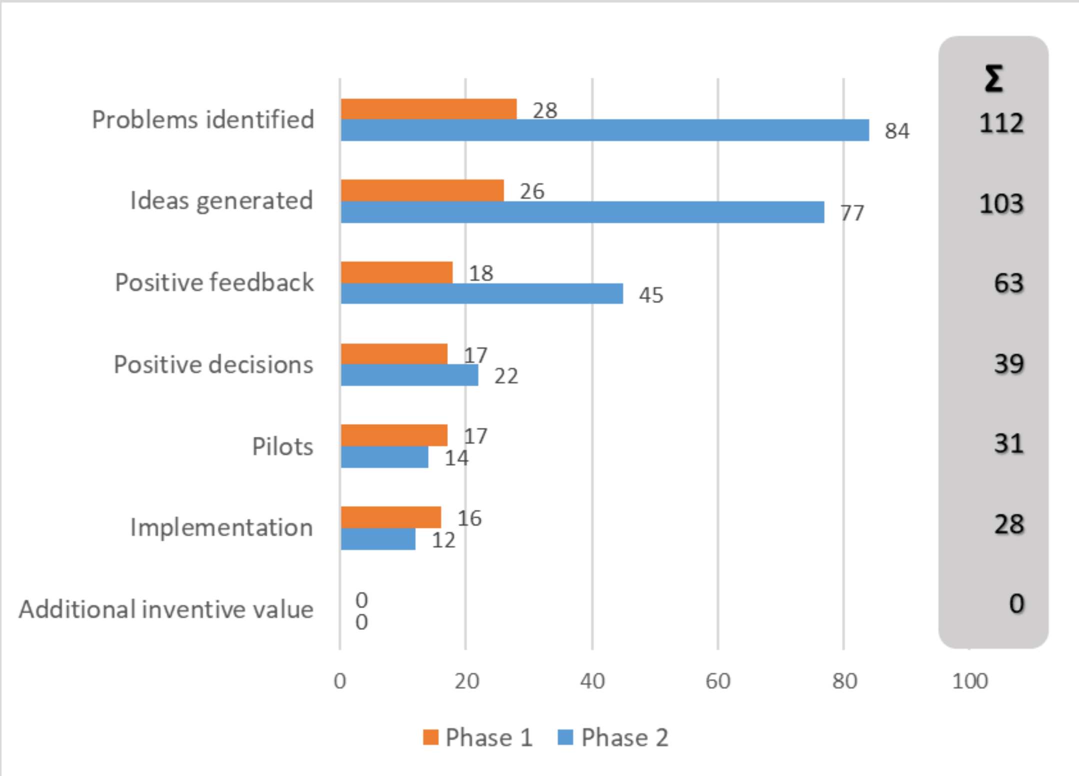

Figure 6: Innovations created and their level of maturity achieved

Furthermore, it became obvious that:

- The generated innovations were mostly incremental ones focussing on quick wins, e.g. a different structure for team meetings or adjustments to the planning of the telephone service times
- The additional space provided to describe the innovation's "impact" was used and the actual score frequently developed during the lessons learned interview

- The scaling of the Valuing tool's indicator "quality of project management" was hardly assessed by the team leaders because of the quick implementation not requiring setting up a project or the fact that the projects to implement more complex innovations hadn't hardly started until the end of the quasi experiment. To handle this situation a standard value of 6 was defined for this indicator. This also influenced the assessment of the Counting tool: the result are lacking its process indicator due to the identical definition (see chapter 5.3)
- The counting tool's item "additional inventive value shown" did not receive any score, thus had no relevance in this setting of work teams.
- In some teams no innovations at all were brought up during the experiment while others were scoring high in innovation performance. Reasons for a lack of innovativeness were discussed with the team leaders but explanations were vague or not provided.
- Despite of this fact, a strong correlation of the tools' results could be found (r=0.82). Never the less, the use of the two different approaches resulted in a different ranking order of the Top 3 teams based on the mean of both phases:
    - Counting: team 1, team 4, team 14
    - Valuing: team 14, team 4, team 17

The team leaders' assessment on the tools' applicability and the measures' impact are integrated into the analysis in the following sections.

7.2. Quality and applicability of the tools

The quality of the tools is determined by their objectivity, reliability and validity. The application of the tool can only be recommended if the results measured by the tool are likely to reflect the team's actual innovation performance. This implies objective, reliable and valid results (Himme, 2007, p. 485). The Counting tool's results are based on quantitatively measuring the number of innovations created by a team within a phase, e.g. the ideas created are counted by the team leader, thus the tool is based on numerical data. This is not possible with regard to the Valuing tool. Its elements feasibility, potential and project management can only be measured indirectly. The team leader quantifies the three indicators for each idea based on the defined scale, an ordinal scale is used. Thus, the tools differ in their character and will be evaluated separately.

7.2.1. Quality of the Counting tool

Objectivity of a tool is described by intra-rater, inter-rater objectivity and objectivity of interpretation (Himme, 2007, p. 485). Due to the fact that the members of the team did not have any knowledge of the experiment and that the researcher did not exert influence on the team leaders during the measuring phases, intra-rater objectivity is given (Rack & Christophersen, 2007, p. 27). Inter-rater objectivity implies no degrees of freedom during the experiment's evaluation and objectivity in interpretation requires that the same results lead to the same interpretation of the results (Himme, 2007, p. 485). This is always given with numerical data, thus the Counting tool is objective.

As stated above, a research instrument is considered to be reliable if the results are consistent over time and an accurate representation of the total population under study and if the results of a study can be reproduced under a similar methodology (Joppe, 2006, p. 598). This is always the case with the numerical data of the Counting tool.

Validity requires that the research truly measures that which was intended to measure (content validity) or how truthful the research results are (criteria validity) (Himme, 2007, p. 491). As described above, the Counting tool intends to measure the innovation performance of working teams. Its criteria are derived from an extensive literature review and the interviews with experts in the field of measuring innovation (please refer to chapter 5.3). To prove the validity of the tool's indicators one has to find correlating outside factors. This would be the case, if teams with a higher innovation performance also have a high working productivity. Even though the teams participating in the experiment are homogeneous (please refer to chapter 6.2.2), the productivity of the teams is measured in different ways amongst these teams so an inter-team productivity comparison cannot be conducted. Therefore, it can be assumed that content validity is given for the Counting tool while criteria validity cannot be proven.

To sum up, the Counting tool is objective, reliable and also valid with regard to content.

7.2.2. Quality of the Valuing tool

The quality of the Valuing tool has to be assessed as well. While comparing the results of measuring the impact of the activities, the results of the valid Counting tool and the Valuing tool are correlating strongly (r=0.81, see Table 25). This is not the case by analysing the Valuing tool and the subjective impression of the team leader (r=0.68, see Table 26)

Variables	Total Counting tool	Total factor impact	l Total factor feasibility	Innovationscore
Total Counting tool	**1**	0.71	0.85	0.81
Total factor impact	0.71	**1**	0.88	0.91
Total factor feasibility	0.85	0.88	**1**	0.98
Innovationscore	0.81	0.91	0.98	**1**

Table 25: Correlation matrix (Pearson): Counting tool vs. Valuing tool

Variables	Total Counting tool	Total factor impact	Total factor feasibility	Innovation-score	Innovation Perfor-mance	Impact all activities
Total Counting tool	**1**	0.64	0.80	0.77	0.68	0.68
Total factor impact	0.64	**1**	0.86	0.89	0.68	0.56
Total factor feasibility	0.80	0.86	**1**	0.97	0.77	0.69
Innovationscore	0.77	0.89	0.97	**1**	0.74	0.63
Innovation Performance	0.68	0.68	0.77	0.74	**1**	0.61
Impact all activities	0.68	0.56	0.69	0.63	0.61	**1**

Table 26: Correlation matrix (Pearson) Valuing tool vs. impression team leader

In addition, the tool also fulfils the intra-rater objectivity, inter-rater objectivity and the objectivity of interpretation. Both tools were applied in the same experimental setting, so the Valuing tool fulfils the aspect of intra-rater objectivity (see above). The aspect inter-rater objectivity requires no degrees of freedom while collecting the data. Even though, there is the general risk that the potential, the feasibility or the process quality of an innovation is assessed differently, the guidelines for assessing these criteria are clearly described, inter-rater objectivity is therefore given. In addition, the innovations and the relating scores were described by the team leaders during the final interview (see above). This ensured that questions could be clarified and that the scoring followed the guidelines.

Applying the tool results in an unambiguous ranking order and a clear score – the team`s innovation performance score (IPS), so the interpretation of the results is objective.

However, any tool used to support decision-making is expected to have reasonable reliability and validity, particularly if its outputs are used quantitatively. The production of reliable findings refers to the consistency, stability, and repeatability of results (Madill, Jordan, & Shirley, 2000, pp. 1–2). While this is the case by simply counting innovations (see above), it is difficult to prove for the Valuing tool. The chosen quasi-experimental setting did not allow to conduct a test-retest method and the application of the split-half method or the calculation of Cronbach's α to confirm reliability (Himme, 2007, pp. 487–489). It might be possible to confirm alternative-forms-reliability. This is the case if different versions of an instrument measure the same phenomenon at the same time and have scores with approximately equal means, variances and α-coefficients (DeVon et al., 2007, pp. 160–161). It can be stated that the Counting and the Valuing tool are based on completely different methodologies. They are two different tools and not two versions of one tool, so they cannot be seen as substantially equivalent (Peter, 1979, p. 10). Reliability of the Valuing tool itself cannot be confirmed by this experiment, so the tool is also not valid (Rack & Christophersen, 2007, p. 27).

In addition, validity is not given because of the characteristics of the innovation performance score (IPS). It is calculated by multiplying the three ordinal scales feasibility, potential and process quality. To multiply it in a meaningful and valid way the differences between the ranks of the single criteria have to be identical (Bowles, 2003, p. 381). This is generally not the case with intervals of an ordinal scale (Shebl, Franklin, & Barber, 2012, p. 154).

To sum up, even though the results of both tools are strongly correlating, this experiment did not confirm the reliability and validity of the Valuing tool.

7.2.3. Additional quality criteria

Next to the main quality criteria objectivity, reliability and validity a tool should also fulfil the following additional criteria (Himme, 2007, p. 486):

- Scaling of values
- Economic viability
- Ease of handling

While the numerical data of the Counting tool obviously inhibits a clear scaling, this is also the case with the Valuing tool due to its ordinal scale. The economic viability of an instrument is strongly depending on the specific situation in the company. This criterion is fulfilled, if the costs for measuring and the time required are in balance with the output (Himme, 2007, p. 486). The expected output of using the instrument is a comparison of the innovation performance of teams. Based on this result, it should be possible to identify factors for fostering innovativeness. As stated above, finding means to improve the innovativeness of employees will be a major competitive advantage for the company (see section 1.1). Comparing this advantage with the expected costs and time required seemed to be in a fair balance according to the authors view. Documenting the team's innovations in the provided template was a quick and easy task for team leaders. Even though setting up the analysis and reporting systems implied effort, the continuous application would be possible with little additional resource needs or could even be part of the normal operational controlling system. In this situation the economic viability is given.

Finally, the ease of handling was evaluated by the degree of filling the template and was also asked for during the lessons learned interviews with the team leaders and will be discussed in the following sections.

7.2.4. Ease of handling – team leaders assessment

As described in section 0, the questionnaire of the lessons learned interview with the participating team leaders contained the following questions regarding the tool's ease of handling (Scale 1-4):

- How would you assess the practicability of the Valuing or Counting tool respectively?
- How would you assess the completeness of the template?

While no additional aspects were required in the template, the practicability of the counting tool was regarded to be better. This is particularly obvious if only the evaluations from team leaders were taken into account, who actually had used the tool. 81% of the users completely agreed with the statement that the counting tool was practical compared to 45% with regard to the Valuing tool (see Figure 7):

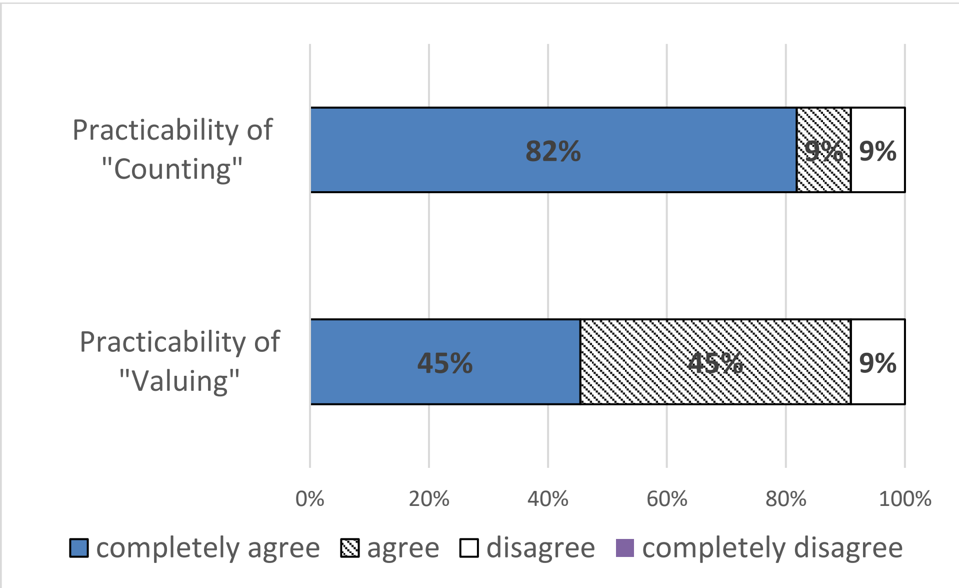

Figure 7: Tool-review: team manager using the tool

Nearly 40% of team leaders did not use the tool. For example, this was the case if the team did not have any innovation during the 2 phases. While some of the team leaders used it even with few innovations, others did not use it and just returned the cards / notes on which the innovations were documented. In all other teams the tool was used. This also includes all teams with a high innovation performance score. It can be assumed that the leaders of these teams dealt intensively with the template due to the fact that the more innovations were generated the higher the effort of filling the tool. So from a researchers perspective their feedback is highly relevant. Thus, while 80% of the team leaders deemed the Counting tool to be very practicable (arithmetic mean: 1.27), only approximately 40% of the team leaders agreed while assessing the Valuing tool (arithmetic mean: 1.63).

7.2.5. Ease of handling – degree of template filled

While the Counting tool's practicability received better feedback by the team leaders, this impression is also confirmed by analysing the degree of completeness the template was filled. Even though the percentage of team leaders applying the Valuing tool compared to the Counting tool is slightly higher (58% vs. 53%), one must state that if the team leaders not using the tool are taken out, the template of the Counting tool was filled in a complete way by all remaining team leaders, while only 11% were willing or able to enter all required information into the Valuing tool while in 53% more than two aspects were missing (see Figure 8, arithmetic mean: 2.3). In particular, the aspect quality of project management was quite often left out, the assessment of feasibility and potential was mostly done.

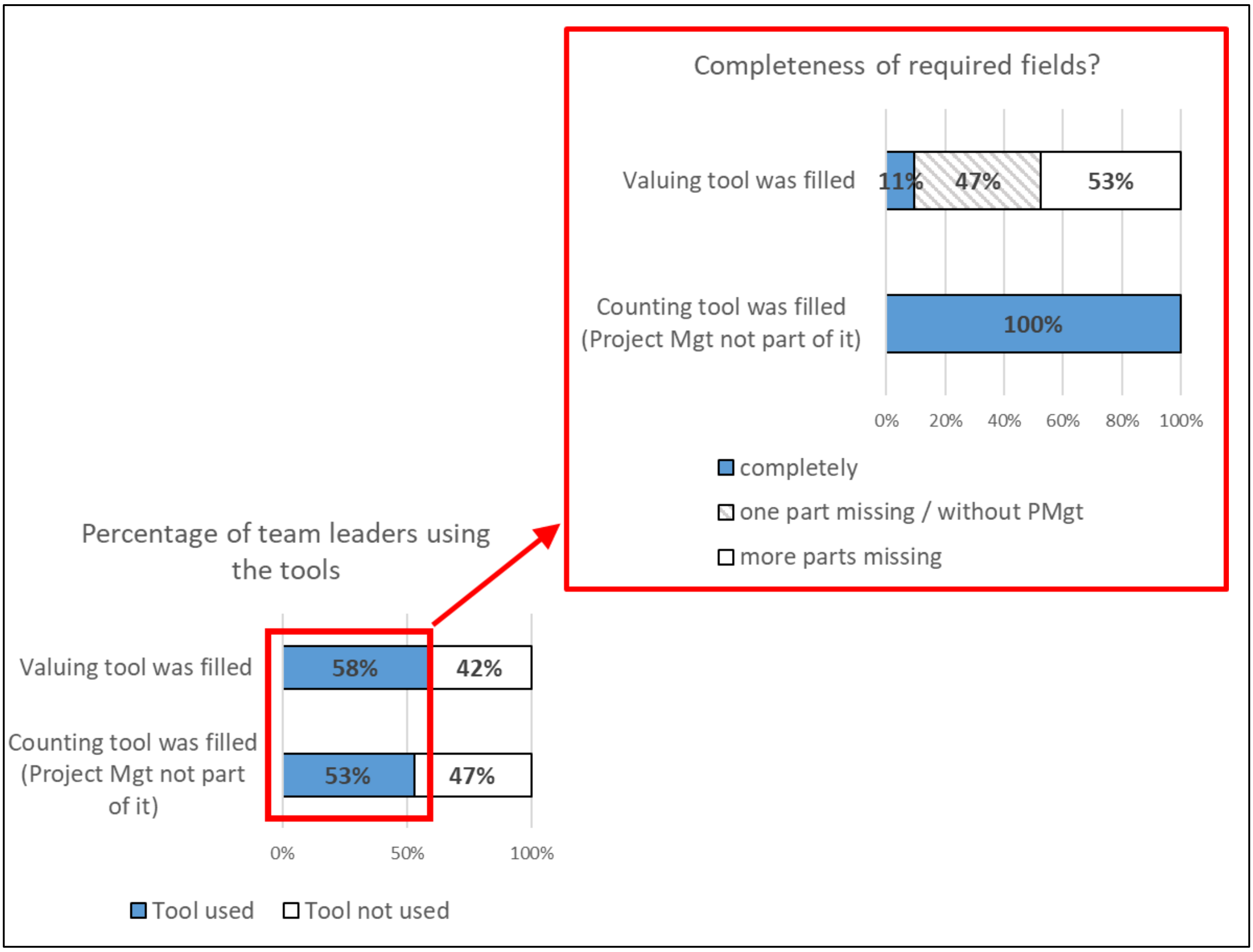

Figure 8: Degree of completing the template

To sum up, the valid Counting tool is easy to handle and fulfils the additional criteria for the quality of a tool. Even though the team leaders' subjective perception did not differ greatly, the evaluation of the degree of completing the tool confirms that the Valuing tool is more complex to handle.

7.3. Impact of measures to increase innovation performance

The experiment also provides the chance to assess the impact of different measures on the teams' innovation performance. The following measures were implemented in the teams (for detailed description please refer to Appendix F)):

No	Measure	Teams
1	Innovation box	team 3, team 7, team 10, team 15
2	Time for ideas	team 16-18
3	Pin board	team 6, team 11
4	Excel table	team 8, team 13
5	Incentive (chocolate egg)	team 12, team 14
6	Team workshop	team 1, team 4, team 5
7	None	team 2, team 9

Table 27. Implemented measures

The effects of implementing measures to improve innovation performance was assessed based on three different perspectives:

- The correlating results (r=0.82, see Table 25) of the Counting (visualized in Figure 9)
- and the Valuing tool (visualized in Figure 10)
- The team leaders assessment of the team's innovation performance before and after the implementation of the measure during the lessons learned interview (visualized in Figure 11) which did not correlate strongly (r=0.68, see Table 26).

These three perspectives resulted in a heterogeneous effectiveness of the implemented measures (see Table 28) which will be discussed in the following sections.

		Counting - points achieved		Evaluating - IPS achieved		Estimation team manager
	Measure	Phase 1	Phase 2	Phase 1	Phase 2	(stable: =; improved: +)
Team 1	Workshop	0	47	0	558	+
Team 2	None	0	0	0	0	=
Team 3	Box for ideas	0	3	0	48	=
Team 4	Workshop	26	16	402	420	=
Team 5	Workshop	0	9	0	162	+
Team 6	Pin board	11	17	120	480	+
Team 7	Box for ideas	14	8	258	192	=
Team 8	Excel table	0	0	0	0	=
Team 9	None	0	0	0	0	=
Team 10	Box for ideas	0	0	0	0	=
Team 11	Pin board	6	20	120	348	+
Team 12	Incentive	8	7	126	108	=
Team 13	Excel table	8	11	165	306	=
Team 14	Incentive	8	38	165	906	+
Team 15	Box for ideas	11	16	174	378	+
Team 16	Time for ideas	8	26	108	576	=
Team 17	Time for ideas	14	15	288	462	+
Team 18	Time for ideas	8	18	126	360	+

Table 28: Results according to perspective of evaluation – overview

7.3.1. Analysis of the effect of the measures in general

Generally speaking three types of development of the teams' innovation performance could be identified by analysing the visualized results of the different perspectives:

1. No innovations in both phases (4 teams – team 2, 8, 9 & 10 - , including 2 teams without measure) or no improvement in innovation performance visible to the team leader
2. Stable or reduced innovation performance measured by the two approaches in phase 2, combined with a consistent innovation level as assessed by the team leader (3 teams – team 4, 7 & 12)
3. Improved innovation performance measured by the tool in phase 2 (11 teams). This result was frequently aligned with the team leaders' assessment of an increase in innovativeness, but with regard to two teams (team 13 & 16) exceptions existed in which the team leader rated a stable innovation performance
4. The measuring and the assessment of the team leaders provided unremarkable results for team 7 and team 8 despite their job tenure below the average (see section 6.2.2).

The existence of the first group might be caused by:

- A lack of innovativeness in the team
- Limited support by the team leader (see section 6.2.7)
- Measures without effect

One reason might be that the teams just have a strong focus on daily work and little on improving the working environment. It might also be the case that team leaders of the control group without a measure had not been as eager as the others in supporting the experiment – despite the regular contact with the researcher. Finally, some of the measures just might not have an impact on innovation performance. For example, teams 8-11 are all part of the business unit 3 but only the measure pin board showed some impact within this group. The findings led the author to speculate, that the team leaders' support might have been missing in these teams. Even though operative teams generally speaking have a stronger focus on getting their work done than searching for innovations, it is very implausible that no problem in the working process was addressed by all teams over the whole duration of the experiment. Even though the two teams of the control group were part of this type, other teams created innovations in phase 1 as well, so having an intervention could not be seen as condition for creating innovations. By ruling the other two potential reasons out, the lack of support by the team leaders seemed to be the most likely cause, the actions to mitigate these as described in section 6.2.7 might not have been sufficient.

Three teams can be assigned to type 2: team 4 (workshop), 7 (box) and 12 (incentive). Even though there are slight differences between the scores of the tools it can be stated that the measures did not have an impact or had a negative impact. Even though this correlates with the subjective assessment of the team leader during the lessons learned interview, this is particularly noticeable with regard to the measure "incentive". In both teams with the measure incentive all incentives (chocolate eggs) had been distributed at the end of phase 2. While a significant rise in innovation performance could be observed in team 14, in team 12 a stable situation on a low level was measured. No explanation for this situation could be found during the lessons learned interview, so the author recommends to re-try this measure while setting up a similar experiment.

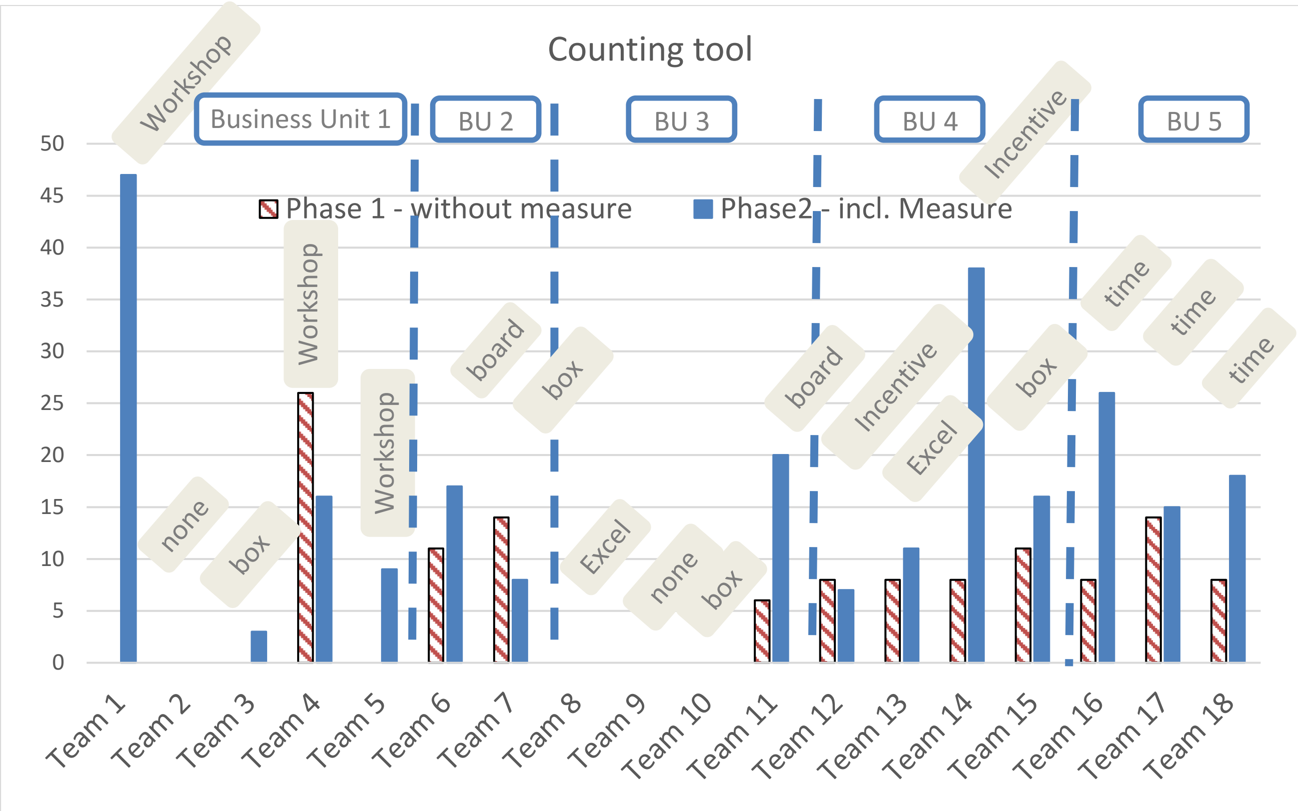

Figure 9: Results of the Counting tool

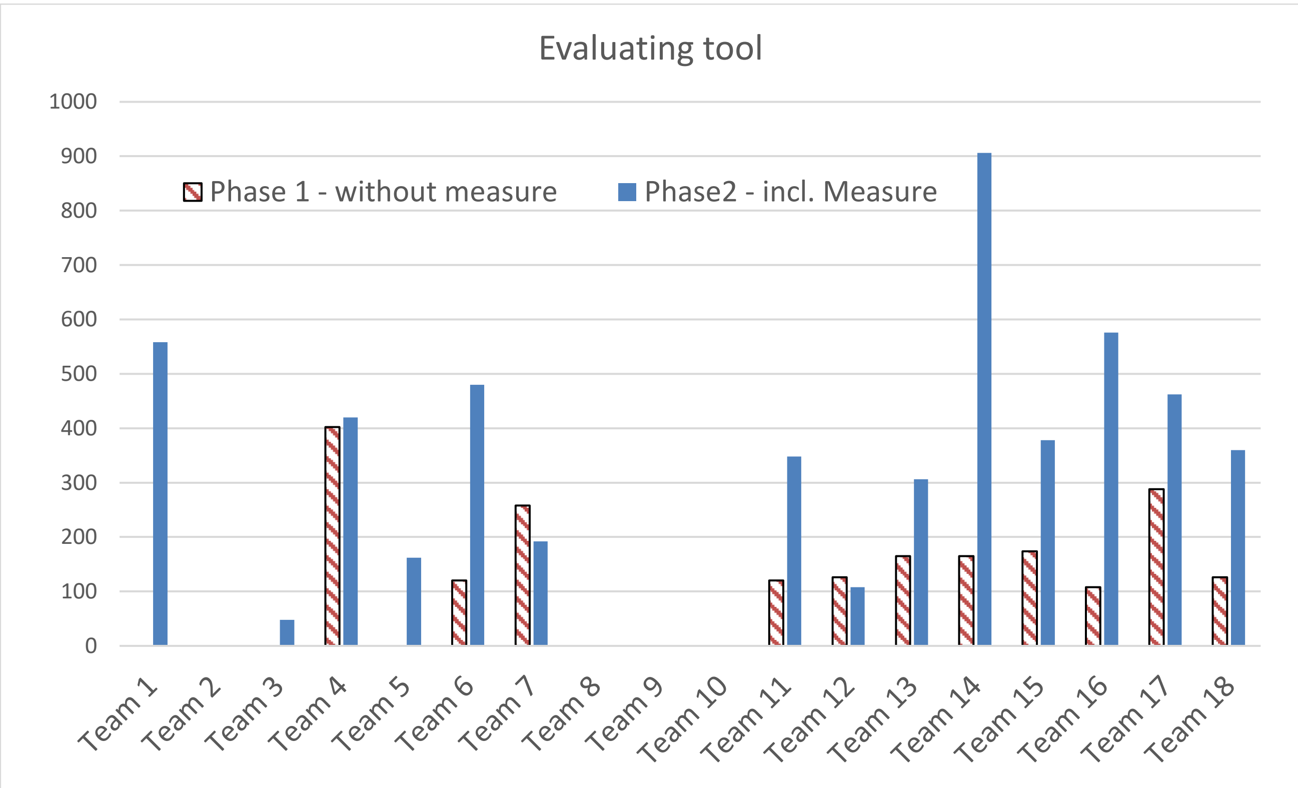

Figure 10: Results of the Valuing tool

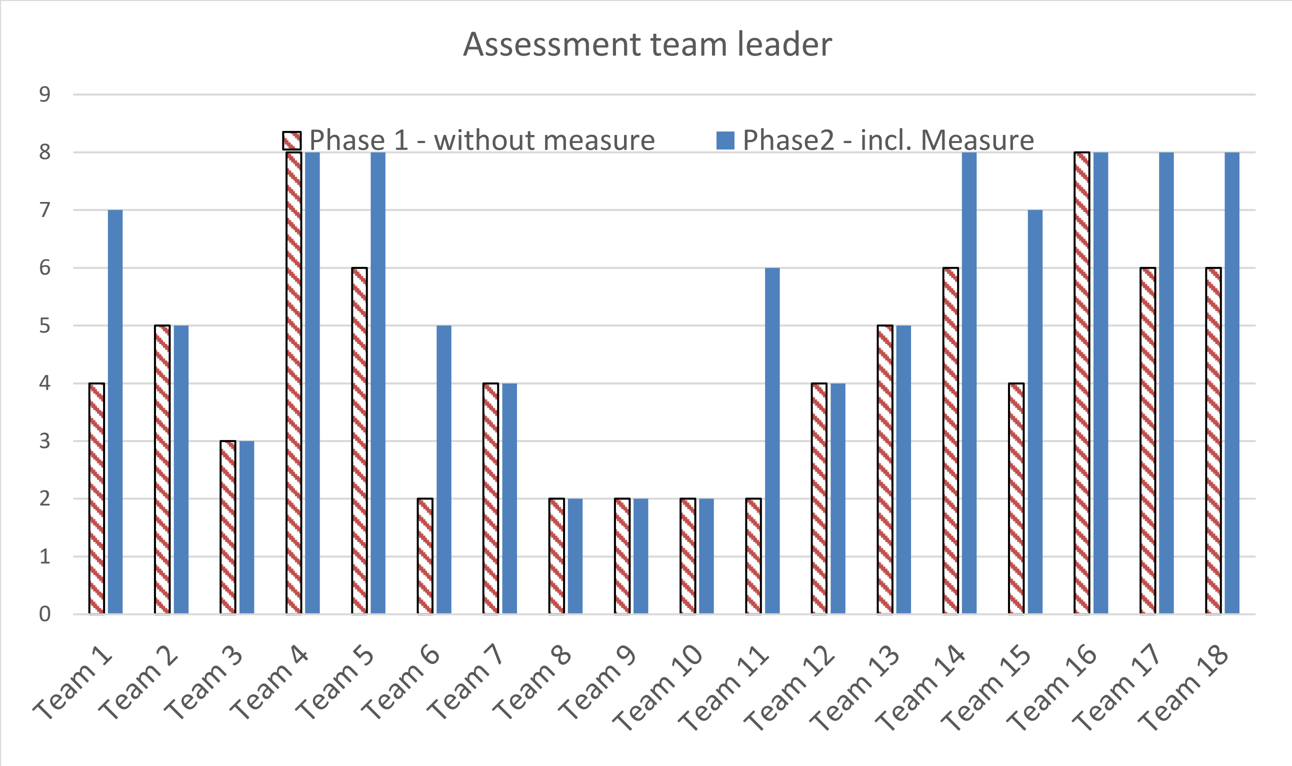

Figure 11: Team leaders' assessment on team's innovation performance per phase

In more of 70% of teams using the measuring instruments an improved innovation performance in the second phase (type 3) could be measured. Particularly in five teams a major increase was measured in phase 2:

- Team 1 (workshop): increase measured with the Counting tool from 0 points in phase 1 to 47 points in phase 2 (Valuing tool: 0 to 558)
- Team 11 (pin board): 6 points compared to 20 points in phase 2 (Counting tool; Valuing tool: 120 to 348)
- Team 14 (incentive): 8 points compared to 38 points in phase 2 (Counting tool; Valuing tool: 165 to 906)

- Team 16 (time for ideas): 8 points compared to 26 points in phase 2 (Counting tool; Valuing tool: 108 to 576)
- Team 18 (time for ideas): 8 points compared to 18 points in phase 2 (Counting tool; Valuing tool: 126 to 360)

These teams were distributed over four out of the five business units (BU) involved with no team in BU 2 and two in BU 5. Within BU 2, one team achieved an increase while the other team's innovation performance declined. Due to this fact and the homogeneity of the teams as discussed in section 6.2.2 it can be assumed that the general characteristics of the BU does not have an influence on the measurement. Even though the effect of the different measures will be analysed in detail in section 0), it could be seen that the box for ideas and the Excel table for ideas were obviously not the suitable measures to foster a high increase of innovation performance.

Not all team leaders had a perception of the change of innovation performance correlating with the measured results which led to relatively weak correlation of r=0.68. For example, the leaders of the team 13 & team 16 did not see a correlation with the implementation of the measure by rating a stable level of innovation performance (see Figure 11). One reason might be the time lag of one upto four weeks between the end of the experiment and the lessons learned interview. While the interview with the leader of team 13 took place on 29th Jan, the interview regarding the experiences in team 16 was conducted on Jan 8th. While the result of the former could be explained by the time lag, the latter seemed to be due to the team leader's very positive view on the team's innovation performance: she rated the high score of 10 before the implementation of the measure. Taking these details into account the correlation might not be as weak as it seemed on first glance.

The ANOVA analysis over all teams showed that the measures led to a statistically significant change in innovation performance. With regard to the Valuing tool, innovation performance score improved by the measure on a statistically significant level (p=0.037, see Table 29) based on the applied model parameters which show a significant difference in average means of the innovation score between phase 2 and phase 1 (see Figure 12).

Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	1	5208.0	5208.0	4.73	**0.037**
Error	34	37406.7	1100,2		
Corrected Total	35	42614.8			

Computed against model Y=Mean(Y)

Table 29: Analysis of variances (innovationscore)

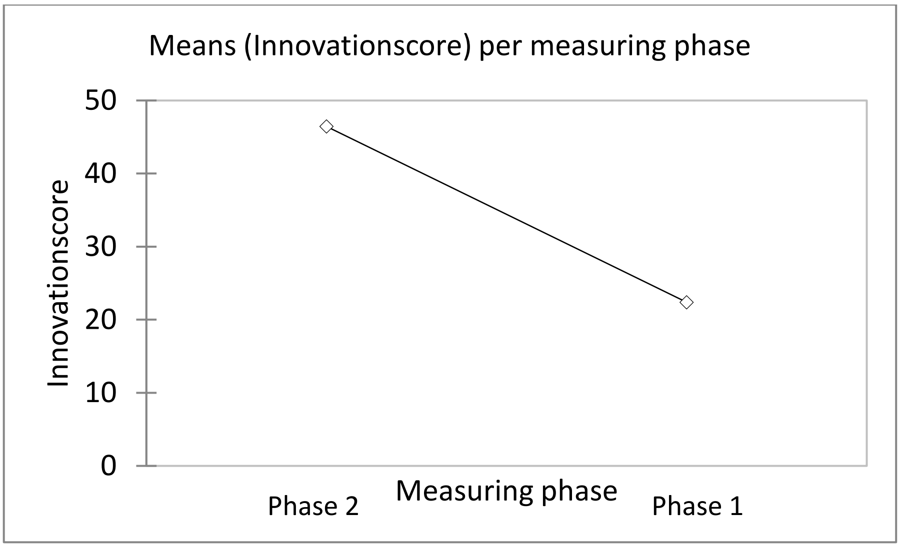

Figure 12: Comparison of the innovation's score means per phase

The Counting tool allowed a more detailed analysis based on the phases of the innovation process. It could be seen, that the improvements of the stages problem identification (p=0.007, see Table 30), idea generation (p=0.005, see Table 31) and positive feedback (p=0.037, see Table 32) were statistically relevant. The following stages were lacking this significance. This might be due to limited number of innovations reaching these phases. The higher number of innovations of phase 1 in later stages (see Figure 6) indicated that the later stages might have been reached after the measuring phases. Alternatively, the selected measures might also have a stronger impact on the early phases. In particular the box, the Excel table for ideas and the team workshop could be assumed to focus on the first phases by fostering creativity. Analysing the teams' innovation performance in which these measures were implemented, showed little or no impact (see Figure 9 and Figure 10), so it is unlikely that this was the reason.

Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	1	87.1	87.1	8.35	**0.007**
Error	34	354.4	10.4		
Corrected Total	35	441.6			

Computed against model Y=Mean(Y)

Table 30: Analysis of variance (Problems identified)

Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	1	72.3	72.3	8.9	**0.005**
Error	34	277.6	8.2		
Corrected Total	35	349.8			

Computed against model Y=Mean(Y)

Table 31: Analysis of variance (ideas generated)

Source	DF	Sum of squares	Mean squares	F	Pr > F
Model	1	20.3	20.3	4.71	**0.037**
Error	34	146.0	4.3		
Corrected Total	35	166.3			

Computed against model Y=Mean(Y)

Table 32: Analysis of variance (positive feedback)

With regard to influence of the organisation unit the ANOVA analysis results in a relevant influence of the team by using the Valuing tools innovations score (p=0.006). This effect does not occur while using the Counting tool (see Table 33), as example the results of the stages problems identified (p=0,187) and ideas generated (p=0.201) are displayed).

Source	DF	Sum of squares	Mean squares	F	Pr > F
Organisational unit – Counting	17	167.1	9.8	1.51	0.201
Organisational unit – Valuing	17	29211.3	1718.3	3.56	**0.006**

Table 33: Influence of organisational unit on innovation performance

A further evaluation of this effect requires to include the interaction of the parameters into the analysis. Because the limited number of degrees of freedom (17) this is not possible, so the cause remains unclear (see Table 34).

Source	DF	Sum of squares	Mean squares	F	Pr > F
Organisational unit	17	29211.3	1718.3	3.56	**0.006**
Measuring phase	1	5208.0	5208.0	10.80	**0.004**

Table 34: Sum of Squares analysis (innovationscore)

To sum up, three types of development of the teams' innovation performance could be identified during the experiment which were distributed over all five involved business unit: no innovations in both phases (4 teams), stable or reduced innovation performance in phase 2 (3 teams) and improved innovation performance in phase 2 (11 teams). By applying a significance level of $p<0.05$ a statistically significant effect by teams in which a measure was implemented, was found for the Counting tool with regard to the first three phases of the innovation process. The ANOVA analysis resulted in values for $p=0.007$ (problems identified), $p=0.003$ (idea generation) and $p=0.037$ (positive feedback received). The Valuing tool showed a generally positive impact of the measures on a team's innovation performance ($p=0,037$), however, a clear distinction whether the team or the measure caused this effect was not possible. Not all team leaders' agreed on this positive relationship during the lessons learned interview, so their assessment showed only a weak correlation with the measured results (Pearson $r=0.068$), while the results of the two tools were strongly correlating ($r=0.82$).

7.3.2. Analysis of the effect of each single measure

The setting of the experiment also allowed to conduct further analysis to evaluate the effect of each single measure. The development was assessed in two ways:

- Change of measured innovation performance in three categories: increase, no change, decrease (see Figure 13). Due to the fact that the results of the Counting and the Valuing tool are identical, they were analysed jointly.
- Subjective assessment of team leaders during lessons learned interview based on the 4-point Lickert-scale question "The measure improved the innovation performance of the team" (see Figure 14).

The analysis of the absolute figures showed that all five teams in which either the pin board or time for innovation were implemented an increase in innovation performance was measured by the tools (see Figure 13). Even though the Pearson correlation did not show a strong correlation between the result of the tool and the team leaders' assessment during the feedback interview (r=0.68, see above Table 26), the team leaders confirmed a strong positive effect of the pin board and the majority agreed on this effect with regard to the intervention "time for ideas" (see Figure 14). The other measures drew a heterogeneous picture: two teams using the box for ideas, one team using the other measures achieved a higher innovation score by applying the tools after their implementation but at least half of them had a stable or even decreased innovation performance in phase 2. The team leaders' ratings were aligned with these measured results or – with regard to the team workshop and the box for ideas – even more sceptical whether the interventions were able to improve the teams' innovation performance. The doubts on the effectiveness of the box of ideas were confirmed by evaluating the measured positive development in team 3, team 13 and team 15 – teams in which the Excel table and the box for ideas were implemented – in more detail. Even though the effects were >0, they were nearly not noticeable: with regard to the Counting tools results team 3 (box for ideas) achieved an increase of 3 points, team 13 (Excel table) of 3 points as well and team 15 (box for ideas) of 5 points (see Table 28). While the results might be inconsistent with regard to the measure "incentive" (see section 7.3.1), one could state that the other measures imply the risk of none or even a negative impact.

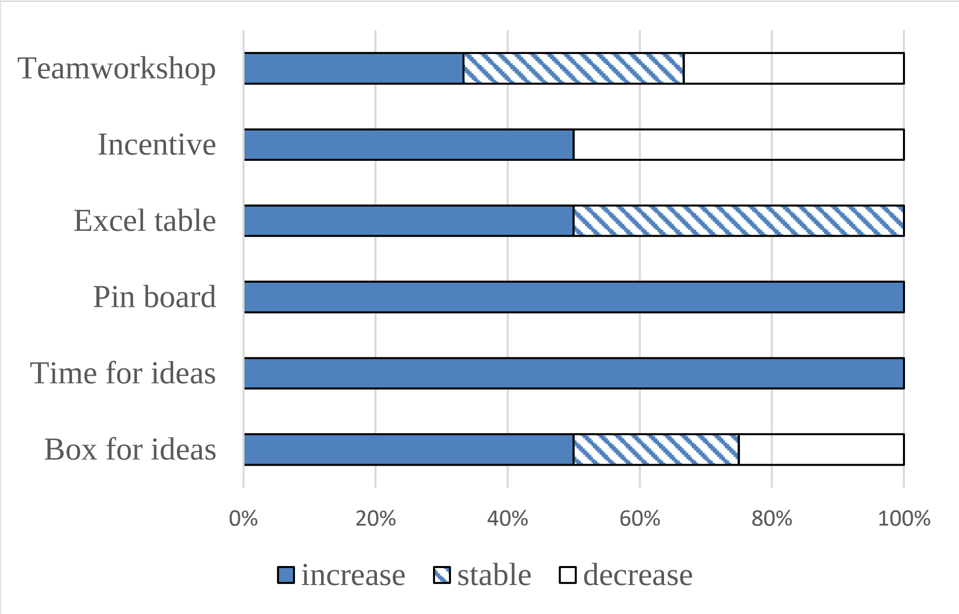

Figure 13: Measured changes of innovation performance depending on measure

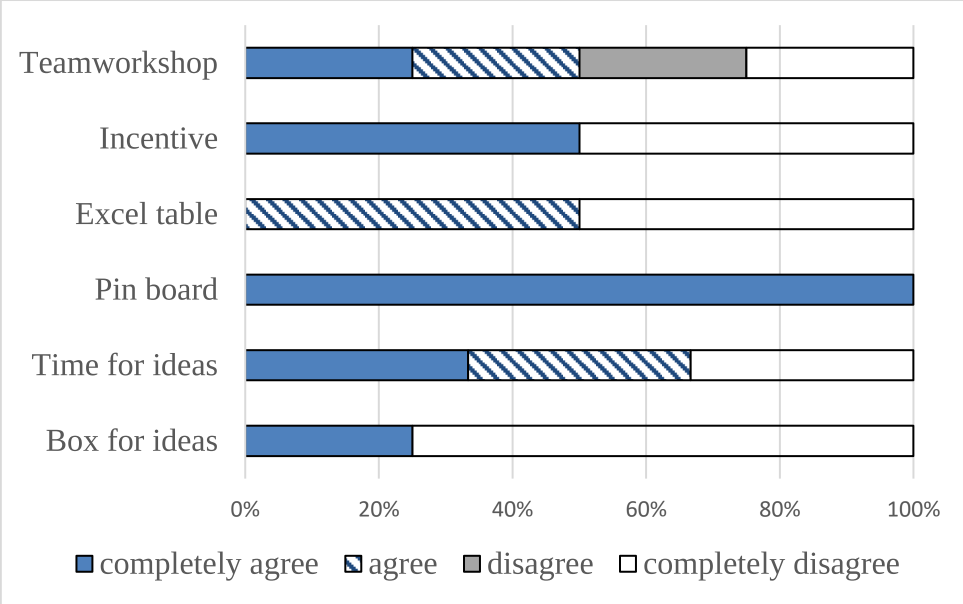

Figure 14: Team leaders' assessment of change of measured innovation performance

This picture is also confirmed by testing the impact of the measures for their significance. While focussing on the changes of innovation performance >0 as measured by the tool a significant effect between the innovation performance and the measures time for ideas and pin board was measured. Providing the pin board to the teams fostered innovation performance on a significant level during the first three stages. The p-values calculated were below the significance level of 0.05 in the phase problems identified (p=0.015), ideas generated (p=0.015) and positive feedback achieved (p=0.010). Similarly, this effect could be measured for the first two phases of the intervention "time for ideas" with very strong p-values of p=0.003 for problems identified and p=0.002 for ideas generated. The in-depth analysis showed that the points awarded dropped in the measuring phase 2 after "positive feedback" from 10 to 4 points (pin board) and after "ideas generated" from 27 to 6 points by implementing time for ideas while they remained stable in measuring phase 1 (see Figure 15: Pin board – ideas per phase and Figure 16). This might be either due to the quality of the ideas created or to the duration of the experiment. It can be assumed that the limited number of innovations reaching these stages explained the lack of significance of the measure.

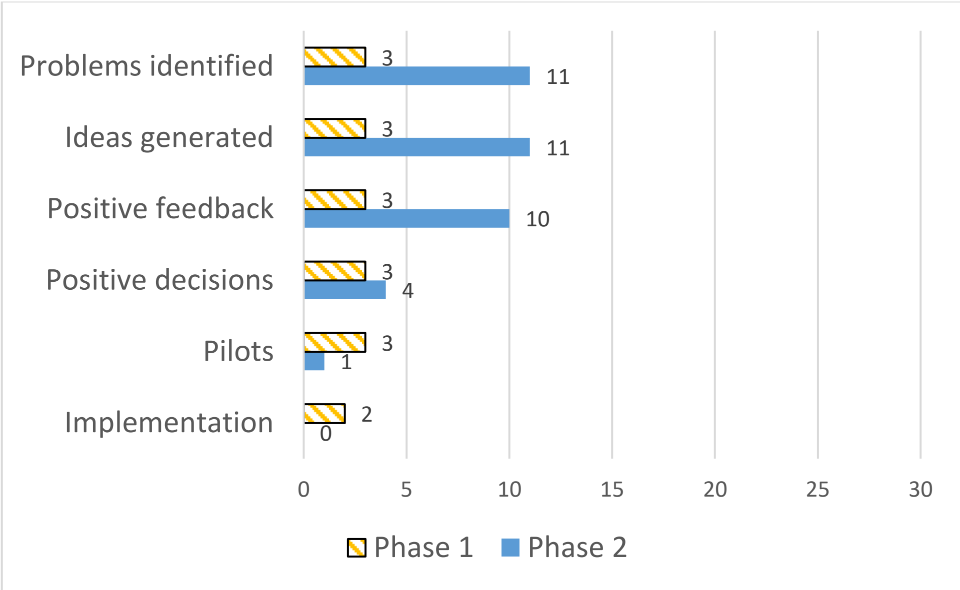

Figure 15: Pin board – ideas per phase

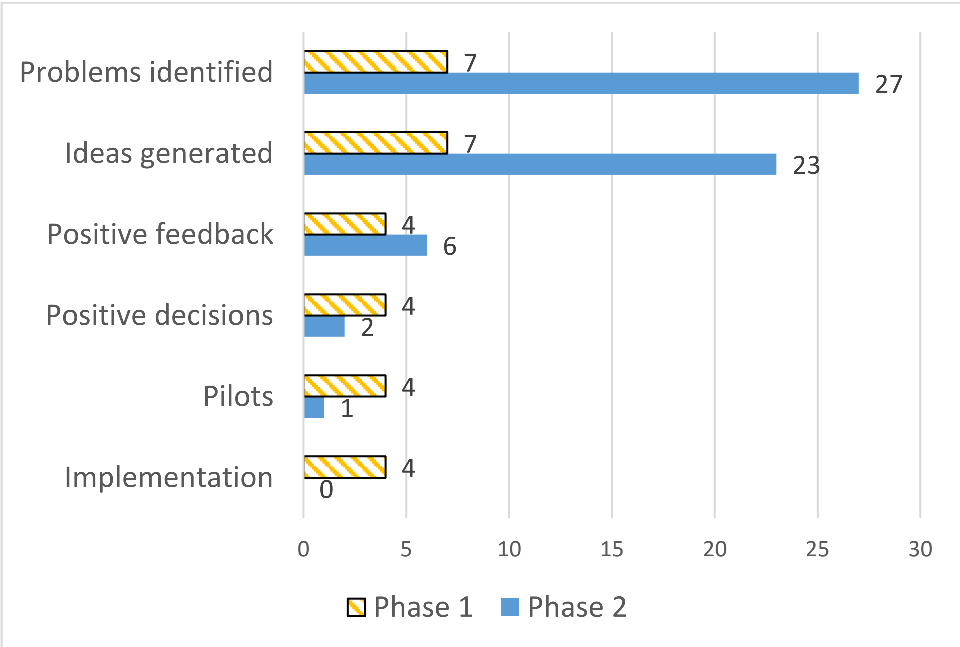

Figure 16: Time for ideas – ideas per phase

While this drop could be generally observed for all measures (see section 7.1) as well, it was assessed in more detailed for the ones with a significant effect. Two categories were defined:

- Implementation time

  Innovation will / is very likely to be realized during or after the experiment

- Feasibility:

  Innovation is (very) unlikely to be realized at all due to a lack of potential, (IT-) resources, etc.

The analysis showed that 73% of the ideas created by teams using the pin board and 59% of the ideas in teams with time for ideas were implemented or were likely to be implemented (see Figure 17). It can be assumed that the lower value of the measure time for ideas and the measured decline after phase 2 already were related, thus the ideas generated in these teams were not as feasible for implementation as the ones from the teams using the pin board had been. By comparing the defined categories in the teams in phase 1 (without the measure) and phase 2 (with measure), the percentage of feasible ideas did not differ greatly. Due to that fact it is more likely that the difference in feasibility is related to specific characteristics of the teams and not to measures themselves. A conclusion that ideas created by teams using a pin board are more feasible compared to the ones with time for ideas or that these measures were only fostering the idea generation process could not be drawn.

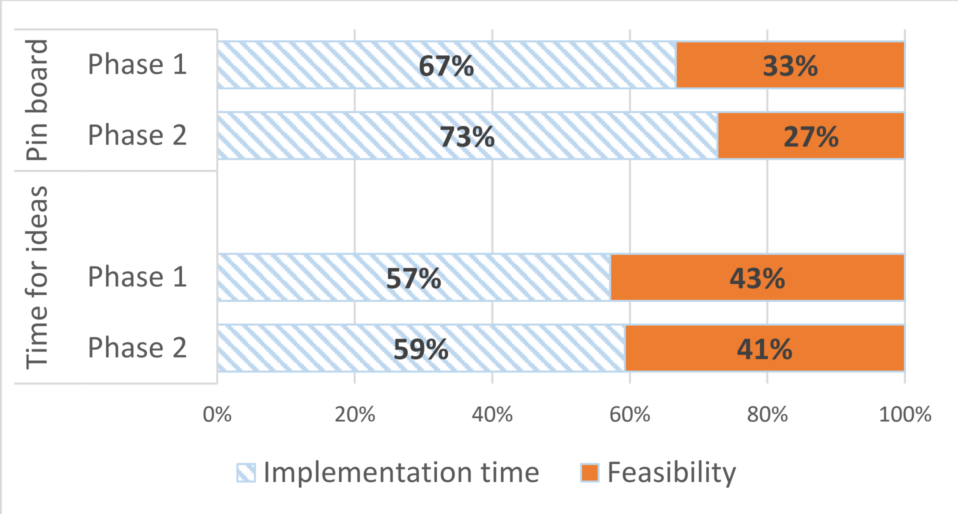

Figure 17: Assessment of reasons for failed realization per phase during experiment

The measured effect of these measures remaining phases as well as the other measures did not show a significance in this experimental setting (for an overview over the results please refer to Table 35).

	Time for Ideas	**Excel table**	**Incentive**	**Workshop**	**Pin board**	**Idea box**
Indicator	p-values	p-values	p-values	p-values	p-values	p-values
Problems identified	**0.003**	0.317	0.182	0.168	**0.015**	0.387
Ideas generated	**0.002**	0.309	0.169	0.168	**0.015**	0.325
Positive feedback	0.304	0.430	0.226	0.131	**0.010**	0.352
Positive decisions	0.789	0.789	0.385	0.184	0.415	0.500
Pilots	0.949	0.789	0.789	0.215	0.854	0.805
Implementation	0.992	0.789	0.789	0.215	1.000	0.805

Table 35: p value for measures per innovation stage

To conclude, all three perspectives showed an increase of innovations in teams a pin board was used. This was also the case for the tools' results of the intervention time for ideas, only the team leaders' assessment was heterogeneous and not as positive. These results were confirmed by the ANOVA analysis of the Counting tool's results for the first stages of the innovation process. The effect of the pin board was significant during the stages problems identified (p=0.015), ideas generated (p=0.015) and positive feedback (p=0.010), the effect of granting time for ideas only during problems identified (p=0.003) and ideas generated (p=0.002). The lack of significance in the later stages was most likely due to the innovations' implementation time which required time beyond the end of the quasi-experiment and not linked to the type of measure implemented or the quality of the innovations created. No significant effect could be proven with regard to the other measures.

7.4. Hypotheses evaluation and summary

The raised questions and stated hypotheses were assessed based on the described findings

7.4.1. Main findings, hypothesis evaluation and contributions

H0 of hypotheses 1 ("The selected measures in total do not have an impact on innovation performance") cannot be confirmed for the first three items of the Counting tool (p=0.037 and less, see above). It seems also verified by the Valuing tool (p=0.037), however the influence of the organisational unit is also significant (p=0.006), so it cannot be assessed whether the effect is actually resulting from the measures.

H0 of hypotheses 2 ("Each of the selected measures has an impact on innovation performance") cannot be proved to be correct by this experimental setting. Only the effect of the measures time for ideas and pin board is significant for the first two or three phases respectively.

7.4.2. Managerial implications for the application of the measuring tools

The study shows that it is possible to measure innovation performance of organisational work teams by the two proposed instruments. Based on the experimental results, some guidance for managers could be given.

The Counting tool was measuring innovation performance along the steps of the innovation process. It can be stated that this concept was plausible for the team leaders having to work with it. In addition, transferring the idea of the bean counting (Pappas & Remer, 1985, p. 18) into the measuring concept by letting the team leaders count every innovation stage reached per innovation was easy to understand. From a managerial perspective on the one hand this approach can therefore be recommended on the other hand one should minimize the risk that innovations are not counted. This risk was particularly present during phase 2 of the experiment. Due to the focus on the implemented measure the team leaders were at risk to only count the innovations directly linked with the measure (e.g. innovations put into the box for ideas) and disregarded other ideas created during the period.

Not only with this regard but generally speaking, keeping the level of communication high by preparing everyday examples for innovations with fit to the chosen setting, explaining them to the team leaders and keep constantly in touch with them to answer questions and remind them on the importance of measuring is crucial for the successful application of the tool.

With regard to the proposed indicators it can be recommended to adapt them to the specific setting. For example, in this work's experimental setting the indicator "additional inventive value shown (e.g. published articles, filed patents)" was of no relevance. It seems very relevant for organisational teams who have a focus on R&D. Similarly, in the setting of work teams, especially in situations the innovation performance is measured for a short time only or in which incremental innovations and quick wins are more likely, a representative value of the indicator "process quality" is difficult to obtain. In addition, even though this item was regarded as necessary by experts and literature (see section 5.2.3, using it would increase complexity of reporting the results. It were necessary to display two separate scores, one for the sum of innovations reaching the different steps of the process and another one as a rating for process quality. Due to their different generation they cannot be combined. Thus, it can be recommended to managers to refrain from these two indicators in similar settings in which measuring them is not constructive. The additional field provided in the measuring template to document a short description of the innovation created, should be also added while using the Counting tool for regular innovation performance measuring. It proved to support the common understanding of the innovation during the lessons learned interviews of this research.

In conclusion, the application of the Counting tool can be recommended from a managerial perspective, while the abovementioned guiding rules (see also Table 36 for summary) are observed.

Guiding rules for managers applying the Counting tool
• Keep the level of communication high with the persons rating the innovation performance high, in particular provide measuring examples linked to the specific work situation and remind them to include all innovations created, not only the ones related to the measures • Always use the six indicators related to the innovation process as measuring items and provide an additional field in the measuring template to allow documenting a short description of each innovation • Only use the indicators “additional inventive value shown” and “process quality” if suitable to the situation

Table 36: Counting tool - guiding rules

Some of the aspects mentioned above are also important managerial implications while applying the Valuing tool: keeping the level of communication high with the raters is also crucial as well is the contemplative use of the item “process quality”. Due to the fact that it is a key indicator within the formula to calculate the innovation performance score with the Valuing tool it cannot be taken out like it was suitable for the Counting tool in certain settings. While applying it within these settings, e.g. in which incremental quick to implement innovations are likely, it can be recommended to managers to set a fixed value, such as it was done during the described experiment (see section 7.1). Generally speaking, the application of the Valuing tool is considerably more complex (see section 7.2.4 and section 7.2.5). Next to the item “process quality”, also the indicator “impact” caused some difficulties for the raters. It was particularly difficult to assess the worth of the impact beyond one’s own nose of the own team in the greater context of the company. Nest to the rating guidelines the additional space provided during the experiment to describe the expected impact to ensure consistence in rating worked well by allowing the comparison of innovations in the greater context. With regard to the item “feasibility” the prepared guidelines were supporting the team leaders during the experiment and enable the assessment of the item.

While applying the Valuing tool managers should also be aware that the resulting innovation performance score might show high variances in absolute numbers because of the multiplication of three elements, e.g. team 3 had an increase of 3 points by using the Counting tool versus 48 points with the Valuing tool or team 14 of 30 points (Counting tool) vs. 741 by measuring with the Valuing tool (see Table 28). Finally, managers should also keep in mind that its validity was not proven by the conducted experiment (see section 7.2.2).

To sum up, several managerial implications can be extracted from this experiment to benefit from the Valuing tool’s capability to take the value of an innovation into account. In particular the contemplative use of the items “impact” and “process quality”, the high variance in absolute numbers evaluated and the unproven validity stick out (see Table 37).

Guiding rules for managers applying the Valuing tool
• Keep the level of communication high as recommended by using the Counting tool • The use of the three developed indicators will provide an integrated result also taking the value of the innovations into account. The rating of the item “impact” in the company context should be ensured by providing the additional field “description of impact” in the measuring template. • For settings with incremental innovation allowing quick win implementation a fixed score for the indicator “process quality” should be defined. • One should bear in mind that big changes in absolute numbers are possible due to the characteristic of the Valuing tool’s formula and • that the validity was not proven during this experiment.

Table 37: Valuing tool - guiding rules

The result of the experiment proved that measuring innovation performance of teams gives managers important information for decision making. The measuring instrument could be used both descriptively and diagnostically. Initially a baseline level of innovation performance amongst teams with similar working characteristics could be established. Then, the development of innovation performance could be documented over time to chart the organisation’s efforts. Diagnostically, based on the differences in development of the scores, successful measures in increasing innovation performance could be identified by analysing successful teams. These measures could be mapped and replicated to the other organisational units.

Based on this specific experimental setting it can be recommended to managers to use the Counting approach due to its higher validity, even though the Counting and the Valuing approach are correlating. In addition, the application of the Valuing approach was more difficult and partly not possible for the team leaders, in particular a standard value had to be defined for the indicator quality of project management. In case there is a need for integrating the innovations’ value into the innovation performance controlling, the Valuing tool should be applied while considering the guidelines stated above.

7.4.3. Managerial implications for implementing measures to foster innovation performance

The experimental results support practitioners in their investment decisions regarding specific activities. By confirming H1 of hypotheses 1, a statistically significant effect of the selected measures to improve a team's innovation performance in general could be found within this quasi-experimental setting. It is therefore absolutely recommendable to invest into implementing measures to foster a higher innovativeness. Depending on the measure selected the required investments are not expensive either.

By analysing the effect of the single measures, the measures time for innovation and pin board stuck out. The measure "time for innovation" is not only applied in innovation driven companies such as Google and 3M (Black, 2016, p. 2) but also has an impact in operational work team settings. While comparing both the pin board's cost benefit analysis yields in a better output by requiring a lower investment. A ready to use template for the pin board is attached in Appendix F).

The impact of other measures was very heterogeneous. While in team 1 great improvement was measured by conducting a team workshop, it showed little effect in team 5 and a negative effect in team 4. Similarly, team 12 and team 14 created the same level of innovation without intervention. After introducing incentives far more innovations were addressed in team 14 while team 12 remained on a nearly stable level. Further research, in particular a longer period of observation is required to clarify the effect of these measures. Practitioners should not invest effort in the measures box for ideas and the Excel table. Both did not improve a team's innovativeness. Thus, based on the experimental results practitioners should primarily prove the implementation of pin boards or time for ideas. In addition, the measure incentive should be considered as well due to their heterogeneous results. In addition, its advantage is the option to combine it with other activities as it is done by many companies having an innovation incentive scheme implemented.

Managers should also continuously communicate about the importance of driving innovation and refer to the measures implemented. While this was done during this study's experiment visibility of the measure seemed to be an important aspect for success. The experiment was conducted in an open work space environment. The pin board placed in the organizational work group was clearly visible to all team members. Innovations placed there could be read and discussed while passing by. The same situation could be created by placing the pin board in the coffee corner or the corridor. Similarly, all members of the team could notice when a chocolate egg ("Kinder Überraschung") was "awarded" as incentive. On contrary, the excel table was not visible at all and the number and content of innovations put into the box for ideas neither. The interventions time for idea and the team workshop represented interruptions in the team members daily routine, thus were also noticed by them. Furthermore, it can be recommended to managers to support the communication with regard to the importance of innovation by action or clear feedback. On the one hand, quick wins should be selected for implementation to point out that creating innovations is important and worth doing. On the other hand – if the innovation is not yet realized – the actual status should be given to the team on a frequent base and documented on the pin board if this measure was selected. The same applies accordingly if the innovation will not be realized: the decision and the reasons should be given to the team to avoid disappointment or the impression that generating innovations is of no earthly use.

In summary, the experimental results imply for managers that the measures in general and the pin board and time for ideas in particular had a positive impact and should be implemented in a visible way (see Table 38).

Guiding rules for managers for selecting measures to foster innovation performance
• Implementing a measure has a positive impact so should be done in any case if possible • The measures pin board and time for innovation proved to be the best choice. The investment for the pin board is likely to be lower, the template used in the experiment could be found in the backup (see Appendix F) • Keep the level of communication high by making the measures and the implementation status of the innovations visible to the team • Select innovations for quick implementation

Table 38: Selection of measures - guiding rules

7.4.4. Overarching theoretical contributions to the innovation literature

The theory of measuring innovation is moved forward in four aspects by this work:

1) Comprehensive overview over approaches for measuring innovation provided
2) Academic discussion on measuring innovation performance further developed
3) Design for a scalable quasi-experiment created
4) Positive effect of measures to improve innovation performance verified

Firstly, a fundament for further academic discussion was laid by identifying 36 existing advances for measuring innovation and clustering them in the four-levels-of-analysis framework. The result made it obvious that a significant gap in research is the lack of an specific framework to measure innovation on the work team level (see Figure 18).

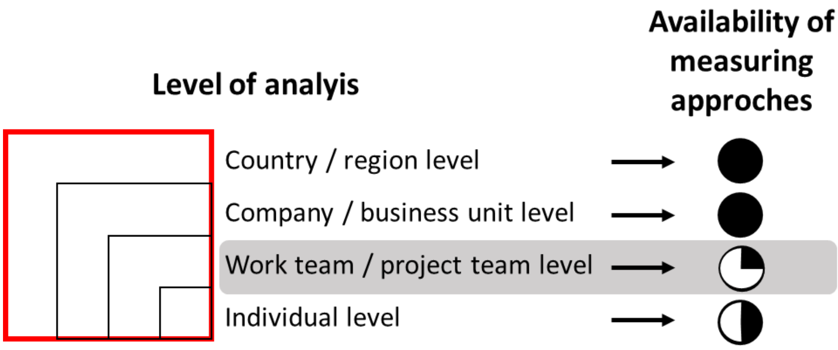

Figure 18: Availability of approaches per level of analysis – overview

Researchers can use this comprehensive overview, concentrate on their own research objective and also draw new conclusions.

Secondly, the academic discussion was moved forward in two ways: existing research was confirmed and suggestions to close the existing gap developed. Because of the fact that the results of both approaches were strongly correlating, the result of Thompson & Choi's study that 'The number of ideas and the number of good ideas therein are usually strongly correlated' (Thompson & Choi, 2006, p. 173) was confirmed. The Counting tool which is based on the three steps idea generation, idea promotion and idea realization was applied successfully. Thus, it was shown, that this interpretation of the innovation process, which is favoured by most authors in literature (see section 3) is reflecting business reality by the successful application of the Counting tool. It can be recommended to base further research on this three step process.

In addition, by developing the Counting tool and the Valuing tool the research supported the closing of the identified gap. In particular, the Counting tool complements existing research with a valid and practical instrument for innovation measuring on the work team level. On the other hand, the Valuing approach showed significant limitations in this specific operational work team setting. By applying it in a different setting and as soon as the validity of the Valuing tool will be proven, it will also be an additional advance which is of high relevance to researcher s and decision makers due to its integration of the innovations' value into the assessment of a team's innovation performance.

Thirdly, a design for a quasi-experiment to test measures to improve innovation performance of teams was created. This design is scalable with regard to the selected interventions, its duration and the teams included. It can also be transferred to other settings if research was continued on the topic. This work also includes the discussion of typical confounding factors such as issues of team selection or covert research, provides exemplary case studies (see Appendix D) and documents the developed measuring templates (see Appendix H). This exhaustive overview contributes to innovation literature by its value as reference to future research on the topic of innovation measuring and allows easy replication of the experiment.

Finally, the author is not aware of any experimental research based on academically standards which evaluated the effect of implementing measures to influence a team's innovation performance. By showing the different effects of measures to improve innovativeness and coming up with the recommendation for two the distinct measures time for ideas and pin board, the study provided new state-of-the science findings in the topic of innovation management and suggests additional opportunities for further research.

To sum up, this work adds valuable contributions to the academic discussion on innovation literature, in particular, the developed tools addressing the identified shortcoming of measuring innovation on team level, the transferrable design of the experiment and the assessment of the effectiveness of the 6 selected measures on the innovation performance of teams moved academic literature on innovation measuring forward.

7.5. Limitations and directions for future research

Although the study represents a step forward, several methodological and application issues require further consideration. The setting of the experiment had a number of limitations with regard to evaluating the quality of the tool on the one hand and the impact of the measures on the other hand which might be considered as potential weaknesses. The circumstances of the quasi-experiment caused some additional limitations to the study's result. The compilation of the teams could not be changed, so they were not homogeneous with regard to working experience, and the duration might have caused the documentation of many quick win innovations. In addition, the Valuing tool's validity could not be proven yet and its indicator process quality could not be properly applied and was given a standard value of 6 to avoid confounding effects. Even though this implies that the Counting tool's result was not integrating the process view as planned, this had a more severe impact on the Valuing tool. Despite the participation of all team leaders during the information meetings, the lessons learned interview and the regular contact during the duration of the experiment, 40% of team leaders did not use the tool at all. Most of these team leaders did also not count any innovations within their teams within the two months of the experiment. Whether this is because of a lack of the teams' innovation performance or the individual characteristic of the team leader could not found out.

There are a number of considerations to mitigate this. Firstly, even though homogeneity of teams was not achieved within this setting of a quasi-experimental design, the teams could be treated as 'essentially equivalent' due to their similar baseline characteristics (Koepsell, 2005, p. 3). Extending the duration of the experiment might provide the opportunity for additional findings but was not required. A statistically significant impact of the implemented measures could be documented in this chosen setting of 2 phases. Based on the assessment of the Counting tool's single stages' results it can be assumed, that the findings would be confirmed and not contradicted by a longer duration.

In addition, even though the validity of the Valuing instrument could not be proven, its development process focussed on ensuring a good content validity. It was conceived on a theoretical perspective, elaborated on by expert interviews and tested by the team leaders. Its soundness was also confirmed by the strong correlation with the Counting tools results.

Finally, the statistically significant impact existed despite the fact that some of the teams did not show any innovation performance during the experiment. One could also assume, that this is due to the lack of support for the experiment by the single team leaders. If these results were taken out, the level of significance would become even stronger with regard to the measures' impact on innovation performance or if the team leaders had used the tool properly.

Moving beyond this, these methodological issues raise interesting areas for further research on human resources in innovation performance, in particular by repeating the quasi-experiment on a larger scale with regard to number of teams and duration.

8. Conclusion

The objective of this research was to identify ways to measure innovation performance of work teams and to clarify the effect of measures fostering innovation performance. Based on defining the term "innovation performance" and a structured literature review the study shows manifold concepts and instruments to measure innovation. Even though fostering innovativeness on all levels of the organisation is essential to be able to compete in today's globalised business environment, approaches to measure innovativeness on the organisational work team level were neglected. Complementing and challenging the results from literature by interviews with academic and business experts in innovation measuring this work addressed the shortcoming. By conducting a triangulation research strategy two approaches for measuring innovation performance were developed: one based on the idea of counting innovations along the steps of the innovation process (Counting tool) and one also integrating the value of the innovations into the assessment (Valuing tool).

The measuring concept of the Counting tool was based on awarding one point for each of the following steps achieved along the innovation process: number of problems identified, number of ideas generated, number of ideas with positive feedback, number of ideas with positive decisions reached, number of ideas tested, number of ideas implemented and additional inventive value shown. To also allow the integration of the innovations' value into the assessment of a team's innovation performance, the Valuing tool was based on the concept of the Failure Mode and Effects Analysis (FMEA). The FMEA's indicators probability of the failure-mode occurrence, the severity of its failure effect, and the probability of the failure being detected (Bowles, 2003, p.380) were replaced by the indicators impact, feasibility and process quality. These were rated on a numerical scale from 1-10 to calculate the innovation's value. The points achieved for all innovations within the measuring period are summed up to the team's total innovation score in both tools. This concrete tool development not only added insights to the academic discussion by closing the gap of measuring innovation on work team level. It also allowed in-depth testing and application of the tools in real life. Due to their different constructs both tools can be used independently or jointly depending on the decision makers' preferences.

The testing of the approaches in this explanatory sequential design approach using qualitative and quantitative methods firstly addressed the quality of the tools. Secondly, the quasi-experiment with organisational work teams of a German insurance company also allowed to identify the impact of different measures to foster innovation performance in teams.

8.1. Overview over quality of tools

The research showed that the tools had different quality levels (see Table 39). The Counting tool's concept proved to be reliable and valid. Within this specific setting neither reliability nor construct validity of the Valuing tool could be confirmed. However, because it is including the value of each innovation into the measuring of the team's innovation performance, the Valuing approach covers an important requirement as addressed by the business experts. To seize on this aspect, the study showed a strong correlation between the results of the tools. In addition, one can assume a high content validity because of the chosen triangulation research strategy.

The Counting advance is also easier to apply by team leaders compared to the Valuing approach. The tools practicability not only received better feedback from the users, the level of completeness was also higher. Even though guidelines facilitating the assessment exist, assigning the scores is more difficult than counting the steps progressed of the innovation process.

Quality aspect	Counting tool	Valuing tool
Reliability	+	(–)
Validity	+	(–)
Ease of application	+	(0)

Table 39: Overview over quality of tools

Within an organisational work team setting, adjustments to both tools can be recommended. The Counting approach's indicator "Additional inventive value shown (e.g. published articles, filed patents)" is only relevant for very few organisational teams who are likely to have a focus on R&D, it should be left out in other settings. The experiment showed that the work teams created innovations which were characterised by incremental process improvements and the possibility of short-term implementation without significant investment. Thus, the item "process quality" was difficult to evaluate. It can therefore be assumed that literature and the interviewed innovation experts are right to assign a high relevance to it in other settings, especially if a long implementation phase is expected and high investments are involved.

8.2. Overview over measures' impact

This work also elaborated on measures to improve innovation performance of teams. The separate sample pretest – posttest control group design showed that the six selected measures to foster innovation performance implemented in the posttest phase had a significant positive impact, the hypotheses "The selected measures in total have an impact on innovation performance" was confirmed. Looking into more detail the measures pin board and time for ideas were the most advantageous within this quasi-experimental setting. They had a significant impact on the team's innovation performance, while this could not be proven for the other measures (box for ideas, Excel table, team workshop and incentives). While decision makers would be prudent to take the limitations of the study into account, this indication still helps them to use resources effectively and avoid investing in measures not causing any effect. Based on a cost-benefit analysis one could state that the measure pin board is cheap to implement yielding in a good output.

8.3. Further research topics

Not all questions of measuring innovation performance of work teams could be covered within this work or analysed in detail. In addition, the limit in time and scope does not allow a transfer to general applicability. There are topics for further in-depth research arising from the results. In particular, model generalization is an issue: increasing the number of participating teams and conducting the experiment for a longer period of time or in different branches or companies might give additional valuable insights. On the one hand the validity of the Valuing approach should be in focus, on the other hand the effect of the measures in general and of incentive and team workshop specifically could be clarified. Additionally, it would be worthwhile to analyse whether the indicator “process quality” can be obtained by long-term application of the tools or in different settings. Alternatively, another more applicable way to assess teams’ innovation performance based on the innovations’ value is another open field of research.

To conclude, although additional work remains in methodological and substantive areas, the findings reported here are encouraging. The two developed and applied tools to measure a work team’s innovation performance and the proof that measures to increase innovation performance, in particular providing teams with a pin board or time for innovation, have an impact, are valuable and easy to implement in daily business from a practitioners point of view. For researchers, this study provides the transferrable and scalable description of a quasi-experimental setting to measure innovation performance and closes a gap of measuring innovation on team level.

Appendix

Appendix A) Interviewees

Name	Position	Company / Institution
Juan A. Marin-Garcia (PhD)	Professor	Universitat Politecnica de Valencia, SP
Rene Butter (PhD)	Researcher	University of Applied Science, Utrecht, NL
Marja Salenius-Ranki	SVP Human Resources (fostering innovation amongst employees, recruitment & resource management)	Elomatic
Dr. Matthias Wiedenfels	CEO (in time of interview)	STADA
Matthias Heutger	Vice President	DHL Innovation Center

Appendix B) Interview guideline

Questions not disclosed to interviewees before the interview are written in italic letters.

Personal Background

1. Can you please give me a short summary about your concrete task / the task of your business unit?

2. Why are innovation particularly relevant for your company / for the segment it is operating in?

3. The whole thesis project circles around the topic of activities to improve innovation of teams, e.g. table soccer, free time for innovation, etc. In which activities to improve employee's innovation performance does your company invest?

   What is your personal interest in innovation / innovation activities

   Are you involved in controlling activities with regards to innovations / innovation activities?

4. How do you / does your company know whether the investments pay off?

Measuring Innovation

5. While measuring the innovation performance of a single team or business unit (4-15 employees) what are important aspects to take into account? Which criteria exists to make a measuring approach great / meaningful?

6. Which approaches do you know to measure a team's innovation performance?

7. The innovation process is defined by the 3 steps idea generation, idea promotion and idea implementation. Which indicators would you use to measure a team's innovation performance with regards to

    - measuring the teams output (output indicators)
    - the team’s performance while managing the 3 step process (process indicators)

    While measuring a team's innovation performance how would you rank the following possible KPI based on their meaningfulness? (The following options were provided on single notes in random order: counting ideas, number of patents, number of positive decisions with regards to innovative proposals, honors / awards from peer group, market share gained by innovation, cost savings from process innovations, failure rates while attempting something new, time to market, percent of projects finished within time and budget, development man-hours per completed innovation, quality performance, number of publications)

    From a literature’s point of view the most important aspect to measure the output is "counting the number of ideas" - what do you think?

Criteria with regards to the innovation process also exist. "Teams with high innovation performance will stay within time and budget while implementing the planned idea." Do you agree / disagree with this hypothesis? Why?

8. During the first day of the Hamburg Fincoda meeting it is all about rating. I also would like to rate 3 examples jointly with you during the interview. How would you therefore rate the following innovation performance examples (max 1 point per aspect) with regards to the Output and Process criteria:

    - Performance to generate, promote and implement innovative output?
    - Performance to manage the process with regards to aspects such as time, quality/goal reaching?

- Example 1: idea to reduce customer phone calls

  The change in wording in a text for customer letters is proposed; the new wording is agreed on by law unit, however the marketing business unit did not respond / agreed yet. The new text will lead to less customer phone calls

- Example 2: proposal for a new product feature

  The business unit claims proposes a new product feature which gets a very positive feedback from the responsible product development unit. However, the decision is not finally taken by product development yet

- Example 3: internal knowledge management

  The business units documents and knowledge are organised on an intranet website. a new structure for the web site is proposed which makes the knowledge management significantly easier. 20% of the new structure are implemented until now.

9. I also would like to discuss in the interview how the aspects could be described in a way that every rater would come to the same result for the innovation performance value.

   What would be your definition for an "output KPI" for innovation performance

   How would you describe the indicators for time with regards to "Idea generation / promotion / implementation"? How the indicators for budget and quality?

10. While rating examples like the ones above one comes up with one value per assessed aspect. It would be great to create one total out of these. Generally speaking to create a total one might want to give different values/weights to the assessed aspects. While weighting indicators within a controlling system what are important aspects to take into account? Which criteria exists which make the weighting approach great / meaningful?

11. Which approaches do you know within controlling to weight indicators?

12. How would you handle this with regard to the examples mentioned above?

    "*The following indicators are possible to measure innovation performance:*

    - *Idea generation: number of ideas*
    - *idea promotion: positive decision with regards to new ideas achieved & positive feedback / praise from other business unit*
    - *idea realisation: number of patents received, number of publications / speeches, honors / awards from peer groups, number of prototyps failures while pursuing something new*
    - *additional ideas from interviewee...*

    Should these output indicators be generally equally weighted or is one more important than another?"

How would you bring these output indicators together to define a single Innovation Performance Indicator score?

How would you handle the process indicators - are they equally weighted?

13. Especially if the innovation is not yet implemented it might be difficult to evaluate an innovation's benefit. What are important aspects to take into account? Which criteria exists, which make the evaluation approach great / meaningful?

14. Which ideas do you have / approaches do you know to evaluate the benefit of an innovation?

    On a scale from 1 (very rough) to 10 (very exact) how much effort should be put in evaluating the benefit of an innovation?

    How would it be possible to assess the value of an innovation from your point of view?

    Within the measuring period of a team's innovation performance there might be teams creating many ideas of average quality, while other have few but very valuable ones. In literature most authors believe that the innovation performance should focus on the number of ideas due to the fact that the likelihood of very valuable ones is higher while creating more innovations. Do you agree with this statement?

    Is a differentiation necessary within teams working in a similar setting? Why?

Application of a measuring tool

15. What requirements do you have with the regards to the applicability of a controlling tool?

    On a scale from 1-10 are they fulfilled?

    What should be done to fulfil the requirements better?

16. In which situations do you think is it generally necessary to measure the success of innovation activities?

Is the discussed approach applicable in these situations?

What else should be taken care of?

Would you use this approach in your business unit? Why? What could be improved?

Further questions

17. Are there other aspects which had you expected to be asked?

Are there further topics I should also cover in my dissertation?

Template for rating examples:

Output			Process		
Generate	Promote	Implement	Time	Quality	Cost

Appendix C) Overview over indicators

Indicator / item	Phase of value creation	Bean Counting	Tidd	Fuchs	Job Center	IPOO	Goffin	Step in inno process (G=generation, P=promotion, R=realisation)	Magic triangle (T=Time, Q=quality, C=Cost)
percentage of sales committed to R&D (process / product)	input		x			x	x	G	
investments in training	input		x					G	
recruiting of skilled staff	input		x			x		G/P/R	
per cent of projects delayed or cancelled due to lack of funding / human resources	input						x	G/P/R	
per cent of total employees involved in innovation projects	input						x	G/P/R	
per cent of personnel training in creativity and problem-solving	input						x	G	
number of journals bought	input					x		G	
number of participation on conferences and trade fairs	input					x		G	
number of problem solving teams	input		x					G/P/R	
log job entry productivity	Process				x			R	Q
business delivery target (achieved)	Process				x			R	Q
quality performance	Process		x					R	Q
quality of service to job seekers	process				x			R	Q
quality of service to firms / customers	process		x		x			R	Q
failure rates - in development process / in the market place	process		x					P/R	Q
number / percentage in overruns on development time / costs	process		x					P/R	C/T

Indicator / item	Phase of value creation	Bean Counting	Tidd	Fuchs	Job Center	IPOO	Goffin	Step in inno process (G=generation, P=promotion, R=realisation)	Magic triangle (T=Time, Q=quality, C=Cost)
customer satisfaction measures	process		x				x	R	Q
time to market	process		x				x	P/R	T
development man-hours per completed innovation	process		x				x	G/P/R	T
process innovation average lead time for introduction	process		x					P/R	T
savings accruing per worker	process		x					R	C
cumulative savings	process		x					R	C
per cent of projects where post-project reviews are conducted	process						x	G/P/R	T/C/Q
per cent of projects killed to late (after significant spending)	process						x	G/P/R	T/C/Q
number of improvements to innovation projects	process						x	G/P/R	Q
complaints: number and type	process					x	x	R	Q
number of patents received (Goffin: & commercialised)	output	x	x				x	R	
number of (technical) publications / quotes / speeches	output	x	x			x		R	
honors/awards from peer groups	output	x						R	
implemented improvement ideas	output		x	x				R	
number of new products introduced	output		x					R	
percentage of sales / profit from new products / services	output/ outcome		x				x	R	
cost / selling price of product	output/ outcome		x					R	

Indicator / item	Phase of value creation	Bean Counting	Tidd	Fuchs	Job Center	IPOO	Goffin	Step in inno process (G=generation, P=promotion, R=realisation)	Magic triangle (T=Time, Q=quality, C=Cost)
market share	output/ outcome		x					R	
growth in revenue / market share derived from innovation	output/ outcome		x				x	R	
higher value added	output/ outcome		x					R	
improved profitability	output/ outcome		x					R	
cost savings by process innovations	output						x	R	
quality improvements by process innovations	output						x	R	
return on innovation investment	output/ outcome						x	R	
earnings from patent licencing	output/ outcome						x	R	
number of new products compared with total number of product in the portfolio	output						x	R	
number of process innovations / new products / new services (compared to competitor)	output		x			x	x	R	
number of prototyps, testings, lines of code	output					x		R	

Indicator / item	Phase of value creation	Bean Counting	Tidd	Fuchs	Job Center	IPOO	Goffin	Step in inno process (G=generation, P=promotion, R=realisation)	Magic triangle (T=Time, Q=quality, C=Cost)
number of new ideas (product/service/process)	output		x					G	
suggestions or ideas per employee	output		x				x	G	

Table 40: Overview over indicators on company level

Appendix D) Case study: Assessing innovation performance

The innovation performance of an organisational work team is exemplarily measured during 2 phases. At the end of each phase the innovation performance score is determined by assessing the number and actual status of the team's innovations.

Innovation performance in phase 1

The following innovations were observed in the team during phase 1:

- Innovation A:
  The team noticed a limited number of phone calls on Friday afternoon and proposed to reduce the compulsory attendance from 6 to 4 employees. The idea received positive feedback, a decision is supposed to follow sooner than expected.
- Innovation B:
  The idea for a new product feature in car insurance contracts was generated and the decision taken to implement it. A profit of 140k EUR is expected and the project progresses nearly as planned.

By using the "counting approach" the team would achieve an output score of 7 in total, 3 for innovation A and 4 for innovation B due to the fact that both innovations reached the stage of positive feedback (so 3 points each) and a positive decision was taken on the 2nd innovation (1 additional point). The process quality score is 8 due to the fact that innovation A progresses faster than expected achieving a score of 9 and innovation B progresses nearly as planned (score of 6).

The team achieves an innovation performance core of 345 by applying the evaluating approach (see table 5). While innovation A has no measurable impact, the team spirit might rise by the reduced compulsory attendance on Friday afternoon. The decision to implement the change is not yet taken, but based on the TELOS aspects the realization should not be too difficult. Innovation

B’s impact is moderate but not ground-braking and the decision is taken. It can be assumed that introducing new features follow a standardised process, so the feasibility is high.

Innovation	Impact	Feasibility	Process	IPS (inno perf. score)
A	3	5	9	135
B	5	7	6	210
			Team IPS phase 1	**345**

Table 5: Phase 1 assessment results Valuation approach

Innovation performance in phase 2

The following innovations were observed in the team during phase 2:

- Innovation A:

  Even though the project progressed according to the plan, the idea was abandoned by the steering committee.

- Innovation B:

  The definition of the insurance terms for the new feature was nearly finished and the project progresses as planned. After refining the calculations the impact was reduced to 110k EUR.

- Innovation C:

  The team noticed that an unusual high number of customers were calling after they received a certain letter concerning the monthly insurance costs. Solving this problem would result in cost reductions of 250k EUR. The setting up of a Lean project is postponed due to a lack of available resources.

The counting approach results in the team's output score of 1 for this phase (see table 6). Innovation A was abandoned and innovation B has not finalized the preparation stage of implementation, so no additional points are assigned. With regard to innovation C a problem was identified resulting in 1 point. The team's process quality score is 7. The progress was according to plan in innovation A and B (score 8) and is lacking behind the timeline in innovation C (score 5) during the setup of the project. To evaluate the team's performance over both phases, the team achieves an averaged innovation performance score of 4 for output and 8 for process quality.

Indicator	**Phase 1**	**Phase 2**	**…**	**Average**
Number of problems identified	2	1		
Number of ideas generated	2	0		
Number of ideas with positive feedback received	2	0		
Number of ideas with positive decision reached	1	0		
Number of ideas tested	0	0		
Number of ideas implemented	0	0		
Additional inventive value shown	0	0		
Total Innovation Output Score	**7**	**1**		**4**
Total averaged Innovation Process Score	**8**	**7**		**8**

Table 6: Phase 2 assessment results Counting approach

The evaluating approach would score the team's innovation performance with 262 points (see table 7). The steering committee's decision to abandon innovation A reduces the feasibility score to 0, thus the product of the elements as well. While the impact of innovation B is slightly reduced (down to 4 points) the feasibility is due to the innovations progress slightly increasing (score of 8). The impact of innovation C is supposed to be high and is scored with 7 points, while the feasibility is currently low due to the fact that a potential solutions is not yet found (score 2). The process score is 5 points (see above).

Innovation	Impact	Feasibility	Process	IPS (inno perf. score)
A	3	0	9	0
B	4	8	6	192
C	7	2	5	70
			Team IPS phase 2	**262**

Table 7: Phase 2 assessment results Valuation approach

To be able to assess the team's performance over both periods it is proposed to calculate the average, thus the average IPS is a score of 304. Even though this is the most pragmatic way, it has to be tested whether this motivates teams to act counterproductive by prolongating the implementation of innovations.

Appendix E) Employee sample task description

I. Organizational Unit			
HQ ☐	**Service** ☐	**Operations** ☒	**Date:**
Job characteristics:			
Managing Dept.:	E8		
Management Level 1:	SUD		
Management Level 2:	USAHH/USBD		
Unit:	USA_HH/USB_D		
Managerial function:	yes ☐ no ☒		
Name of position: (max. 40 digits)	Clerk invalidity claims		

II. Description of main tasks
(If possible max. 5 tasks)
Examination of the obligation to pay benefits and determine the claims amount within the framework of new claims and follow-up processing (e.g. disability) in the case of more complex claims Decision on the liability within the designated payment competence Payment of benefits according to payment competence/preparation of payment Correspondence in standardized and free form Creation of appropriate loss reserves Telephone calls with clients, lawyers, etc.
Further remarks

III. Required skills and competences
(If possible max. 5 skills / competences)
Insurance clerk or comparable education Sound knowledge of the legal framework in private accident insurance Comprehensive basic medical knowledge Confident command of the German language, both written and spoken Knowledge of host and PC applications Negotiation skills Many years of experience in claims settlement
Further remarks

Appendix F) Description of measures

The following measures were selected (see Table 41):

Measures to foster innovation (implemented in phase 2)	Short description
Box for ideas	A box (see picture) and templates for innovation description in card-format were provided for the team. Described innovations were put into the box and counted at the end of phase 2. 
10% worktime to create and work on new ideas	Team leaders announced that every member of the team is allowed to spend 10% of their working time to create and elaborate on innovations either on their own or in groups
Pin board	A pin board (see Figure 19) and templates for innovation description in card-format were provided for the team. Described innovations were put onto the pin board and counted at the end of phase 2.
Excel-table on a shared drive	An Excel-table to describe innovations was made accessible for all team members. The description consisted of the following aspects: • Explanation including real life examples for innovations in operational working teams • Name of innovation • Short description of identified problem / innovation • Voluntary aspects: advantages, requirements for implementation, contact person
Team workshop	Team workshops were conducted during phase 2. They were planned to work on innovations amongst other topics. The team leader were in charge for selecting the method while discussing the innovation part.
Incentive: reward of 1 chocolate egg in case of creating or implementing innovation (max. 1 per person)	The team leaders announced, that everyone generating, promoting or implementing an idea would receive a chocolate egg ("kinder surprise"). The incentive was limited to a maximum of 1 per person, independently from the number of innovations brought up

Table 41: Description of measures

A detailed description of the selected measures was prepared containing the aspects (see **Figure 19**):

- How to communicate initially to the team
- What to do / how it works
- Proposal for periodical information (including sample mails)

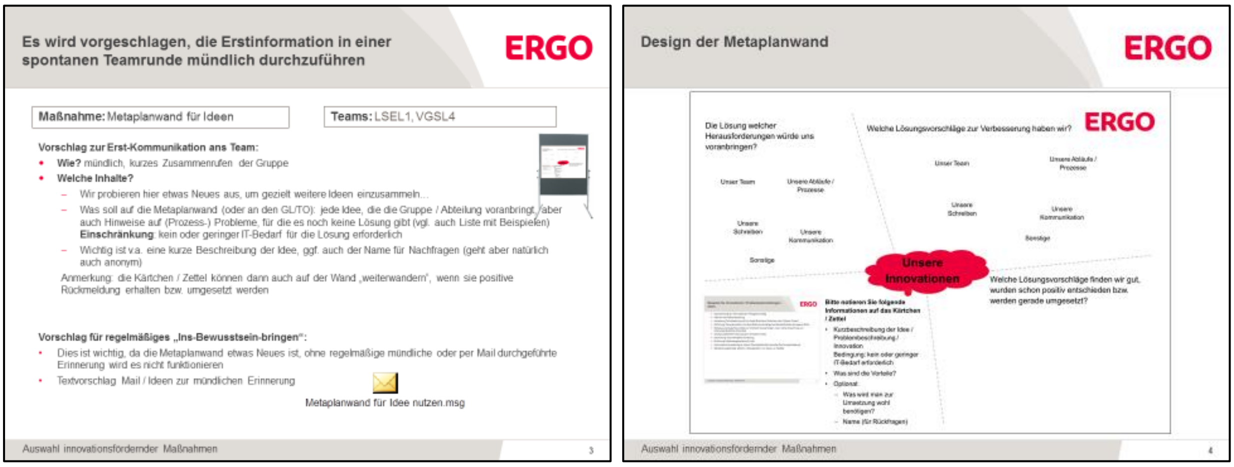

Figure 19: Handling information for team leaders – example pin board

Appendix G) Communication approach

The communication between team leaders and researchers lasted from October 2017 until February 2018. It consisted of 10 steps including the lessons learned interview after the experiment (see Table 42).

Time	Communication activity
Week 40/41 2017	Presentation of study to heads of department
Week 41	First presentation to team managers, including the study's background, planned process and general measuring approach; objective: identifying volunteers
Oct 25th	Meeting with participating team managers: • Presentation of the measuring template • Presentation of concrete examples for innovativeness within the operational setting • Presentation of the planned measures to increase innovation performance • Discussion / open questions • Decision on timing and duration of phase 1 and phase 2
Nov 13th	Meeting with participating team managers: • Sharing of lessons learned / experiences and discussing open questions • Decision on implemented measure per team • Tips for communicating the measure to the team
Nov 23rd	• Finalizing text to inform team managers about the measure and the communication approach with internal operations expert • Pre-information of department heads (including request for feedback)
Nov 27th	Information of team managers per mail, including: • A documented description of each activity • Mails to inform team about activity (standardised wording)
Nov 29th	Handing over required tools to team leaders, e.g. idea boxes, print outs for pin boards, incentives (egg surprise), etc. with team managers
Dec 11th, 15th and 21st	Standardised mails as reminder to be sent by team managers to activate teams to use the new measure / organisational practice
Jan 8th till 29th, 2018	Feedback interviews with team managers • Clarification of innovations created • Completing the tool template, in case fields were left open / were not filled • Conducting evaluation on base of interview guideline
Jan/Feb 2018	Debriefing of teams: team members were informed about background of experiment

Table 42: Communication approach

Appendix H) Template for measuring

The measuring templates also containing comprehensive descriptions and examples were handed over to the team managers in German language.

Template team innovation performance

Period: ___________________

Date of evaluation: ___________________

Team name / identifier: ___________________

	Number
Number of problems (in processes) identified	
Number of ideas generated	
Number of ideas with positive feedback	
Number of ideas with positive decisions reached	
Number of ideas tested	
Number of ideas implemented	
Additional inventive value shown by the team	

			Assessment of:		
	Name of idea/ innovation	**Short description of idea/ innovation**	Likelihood of implement-tation (scale 1-10)	Potential (please give a short description of the expected advantages)	Quality of project management (scale 1-10)
1					
2					
3					

Explanations and example

Thank you very much for supporting the test of the measuring template for innovation performance of teams. This test is part of research project on innovation which is conducted by the University of Applied Sciences Hamburg (HAW Hamburg). Without your support it would not be possible to finalize the development of the template. Confidentiality and protection of private data are important aspects of this research. The data will be anonymized after the measuring phase to make identification of single teams impossible. Individual data of single persons will not be collected. Please to not inform your team that they are taking part on an experiment. The behaviour of the team members would be changing in this case and the experimental results were not valid anymore.

Your contact person for further information: P. ter Haar, phone HH-4901

Measuring period: 2x 1 month

Measuring template part 1: Counting innovations / ideas

The described output indicators should be summarized to a total score: every time an innovation reaches the next stage 1 point is counted. The stages are described as follows:

- Number of problems (in processes) identified

  This aspect refers to all process and organisational problems which are adressed by your team members during the measuring period.

- Number of ideas generated

  All ideas should only be taken into account if they are likely to have a potential value for the team or the company and are unique relative to other ideas currently discussed in the workteam

- Number of ideas with positive feedback

  Ideas, which are regarded by the team leader or colleagues to be submitted for final implementation decision, e.g. to the manager or the steering committee.
- Number of ideas with positive decisions reached

  This stage is reached if the final decision was taken to start the implementation of the idea
- Number of ideas tested

  Some ideas are not implemented into practice at once, they require to be tested by a pilot or another kind of testing phase or implementation in only parts / sections of the final situation. This is meant by this criteria.
- Number of ideas implemented

  The idea is implemented in all relevant parts /sections and put into practice.
- Additional inventive value shown by the team

  Please add 1 point if the innovation becomes particularly popular, e.g. in case it is mentioned in internal communication, received a company award, led to publishing in external media, received a patent, etc.

Achieving a stage implies, that the stages below were also reached, e.g. generating an idea implies that a problem was observed so one should assign 2 points even though the team did not explicitly state this fact. Only progress is valued in every phase, thus every step is only counted once. In case that the decision for implementation was taken before the measured period, a maximum of 3 points (ideas tested, ideas implemented and additional value achieved) can be achieved in the current phase.

A score for additional inventive value is rare and could be granted e.g. in case the innovation is described in company’s intranet or the win of a (company-internal) price for the idea.

Example:

Measuring period 1:

- Innovation 1:

  The team realised that the programme printing customer letters does not include claim payments correctly. In contrary to the terms and conditions only 50% of the payment for stay in hospital is documented. The team is looking for a solution jointly with the IT. This takes longer than expected, thus the letters have to be corrected manually.

- Innovation 2:

  The regional sales director is currently informed by mail and asked for permission before the insurance contract is adjusted or terminated due to claims frequency. He / she then asks the local representative. If permission is granted, the claims handling team asks the local sales rep to adjust the contract. The claims handling team now proposes to waive the agreement process due to the fact that the local sales rep will get to know the plan to adjust the contract. Thus, the local rep got the chance to object anyway. The manager of the claims team agrees and asks the central sales team for a positive decision.

⇨ Team’s innovation performance in period 1: 4 (2x problem identified, 1x idea generated, 1x positive feedback achieved)

Measuring period 2:

- Innovation 1:

  The error cause could be identified and corrected within short notice.

- Innovation 2:

  Central sales team refuses a change of the current process as a matter of principle.

- Innovation 3:

  A problem while handling the system for administrating personal customer data was identified. A solution is not found yet, however, cost reductions of 70k EUR are expected. The setup of a lean six sigma team is planned but will take some more time.

⇨ Team's innovation performance in period 2: 5 (1x problem identified, 1x idea generated, 1x positive decision, 1x idea implemented (plus 1 for idea tested)

The measure should be conducted twice and should cover the period of 1 month each. Please note that it is important that you only take incidents into account which took place during the single measuring period. Thus, if an idea was created in September and the implementation is agreed on during the measuring period in November, only 1 point is achieved ("positive decision").

Measuring template part 2: Valuing innovations / ideas

The second part of the measuring template aims to get a clearer picture of the idea / innovation while analysing the results. Please name the idea and give a short description of it. Please also value the idea who you would assess it at the end of the measuring period. General guidelines are part of this description to give orientation and facilitate this process. Please assign points on a scale from 1-10.

Example: see above

Valuation measuring period 1:

Innovation	Potential	Feasibility	Quality project mgt.
1	Reducing manual work	4	7
2	Reducing manual work, improved response teime towards customer	6	8

Valuation measuring period 2:

Innovation	Potential	Feasibility	Quality project mgt.
1	Reducing manual work	10	10
2	Reducing manual work, improved response teime towards customer	0	10
3	70k EUR	2	5

Definitions / Guidelines for orientation

Feasibility:

Rank	Feasibility	Criteria: likelihood of implementation
9-10	Realization (almost) guaranteed	About 99 in 100 innovations will be implemented at this stage
7-8	High	About 95 in 100 innovations will be implemented at this stage
5-6	Moderate	About 70 in 100 innovations will be implemented at this stage
3-4	Low, failure likely	About 50 in 100 innovations will be implemented at this stage
<2	Failure (almost) guaranteed	≤ 0,5 (less than 1 in 2 innovation will be implemented in this situation)

Potential

Please describe the advantages of implementing the innovation. You might consider a financial impact as well as non-financial benefits such as improve team spirit, quick access to existing knowledge, improved response time towards customer etc.

Quality of project management

Rank	Quality of project mgt	Criteria: Accordance to plan
9-10	Better than planned	1 or more aspects have a positive deviation to plan
7-8	In (approved) plan	All 3 aspects (time, budget, quality) are in plan
5-6	Minor negative deviation	1 aspect not in plan, likelihood to catch up with plan; Project is currently being set up.
3-4	Negative deviation	2 aspects not in plan or 1 extremely out of plan
<2	Out of control	3 aspects out of plan or equivalent bad situation

Appendix I) Lessons learned interviews – template

The structured lessons learned interviews with the team leaders aimed to evaluate the following aspects:

- Helping the researcher to understand the content and status of the team's innovations and the tools' scores
- Finalizing the filling of the templates in case a team leader had open questions
- Obtaining the team leaders' subjective assessment on the effect of the measure
- Receiving feedback to the tools' ease of handling and the process of the experiment

How would you assess the impact of the measure and the usability of the template?

Thank you for supporting the experiment in the last months. In retrospect, your experiences are important for the continuous improvement of the template and the selection of effective measures. In particular, how would you assess

- the impact of the measure to improve innovativeness on your team?
- the usability / applicability of the template?
-

Team:_________

Impact of measure:

	strongly agree	agree	disagree	strongly disagree
The measure improved the innovation performance of the team:	❑	❑	❑	❑

What was the main reason, what would have increase the effect?

Assessment of innovation performance **before** the measure (scale 0 (low) up to 10 (high)): 10 8 6 4 2 0

Assessment of innovation performance **during** the phase the measure was in place (scale 0 (low) up to 10 (high)): 10 8 6 4 2 0

Usability of template:

	strongly agree	agree	disagree	strongly disagree
The template covered the aspects relevant for measuring the team's innovation performance:	❑	❑	❑	❑
The effort to quantify the ideas (upper part of template) was adequate while applying the template	❑	❑	❑	❑
The effort to describe and valuate the ideas (lower part of template was adequate while applying the template	❑	❑	❑	❑

Which alternative did you prefer? Do you want to add further comments?

Innovation performance of a team – guidelines for assessment:

Value	Innovation performance	Criteria: generate, promote and implement ideas
9-10	Very creative and get ideas implemented	All members of the team burst of ideas to improve the situation and implement them successfully
7-8	Pushing continuous improvement	(Nearly) all members of the team are continuously aiming to identify and implement new ideas, processes, products, …
5-6	Supporting continuous improvement	Single members of the team are continuously aiming to identify and implement new ideas, processes, products, … The team is supporting these.
3-4	Occasional innovative	The team is only searching for improvements when requested and has little interest in implementing them. Single innovative team members are unable to assert themselves.
<2	„Work to rule“	The team keeps the status quo or works according to existing specifications, even though these do not describe the optimal way of acting. The request to be innovative does not lead to change, innovative team members are hindered by the group.

Reference List

Alge, B. J., Ballinger, G. A., Tangirala, S., & Oakley, J. L. (2006). Information privacy in organizations: Empowering creative and extrarole performance. *The Journal of applied psychology, 91*(1), 221–232.

Amabile, T. M., Conti, R., Coon, H., Lazenby, J., & Herron, M. (1996). Assessing the Work Environment for Creativity. *Academy of Management Journal, 39*(5), 1154–1184.

Anderson, N., Potocnik, K., & Zhou, J. (2014). Innovation and Creativity in Organizations. *Journal of Management, 40*(5), 1297–1333.

Anderson, N. R., & West, M. A. (1998). Measuring climate for work group innovation: Development and validation of the team climate inventory. *Journal of organizational behaviour, 19*, 235–258.

Astebro, T. (2004). Key Success Factors for Technological Entrepreneurs' R&D Projects. *IEEE Transactions on Engineering Management, 51*(3), 314–321.

Axtell, C. M., Holman, D. J., Unsworth, K. L., Wall, T. D., Waterson, P. E., & Harrington, E. (2000). Shopfloor innovation: Facilitating the suggestion and implementation of ideas. *Journal of Occupational and Organizational Psychology, 73*(3), 265–285.

Basu, R., & Green, S. G. (1997). Leader-Member Exchange and Transformational Leadership: An Empirical Examination of Innovative Behaviors in Leader-Member Dyads. *Journal of Applied Social Psychology, 27*(6), 477–499.

Bell, S. T. (2007). Deep-level composition variables as predictors of team performance: A meta-analysis. *The Journal of applied psychology, 92*(3), 595–615.

Bergmann, G., & Daub, J. (2008). *Systemisches Innovations- und Kompetenzmanagement: Grundlagen - Prozesse - Perspektiven ; [Bachelor geeignet!]* (2., aktualisierte Aufl.). *Lehrbuch.* Wiesbaden: Gabler.

Bieck, C., Cornelius, D. A., Patel, S., & Uramatsu, H. (2016). Rethinking insurance, 1–21. Retrieved April 27, 2019, from https://www.ibm.com/downloads/cas/9DLZ2LNL.

Black, S. (2016). *How the 15% Rule Became a Stepping Stone for 3M's Innovation.* Retrieved September 19, 2018, from Market Realist: https://marketrealist.com/2016/06/15-rule-became-stepping-stone-3ms-innovation.

Bogner, A., & Menz, W. (2009). Expertenwissen und Forschungspraxis: Die modernisierungstheoretische und methodes Debatte um die Experten: Zur Einführin in ein unübersichtliches Problemfeld. In A. Bogner, B. Littig, & W. Menz (Eds.), *Experteninterviews. Theorien, Methoden, Anwendungsfelder* (3rd ed.). Wiesbaden: VS Verlag für Sozialwissenschaften.

Bowles, J. B. (2003). An assessment of RPN prioritization in a failure modes effects and criticality analysis. In *Annual Reliability and Maintainability Symposium* (pp. 380–386). Piscataway, N.J: Institute of Electrical and Electronics Engineers.

Bowles, J. B., & Peláez, C.E. (1995). Fuzzy logic prioritization of failures in a system failure mode, effects and criticality analysis. *Reliability Engineering & System Safety, 50*(2), 203–213.

Brolmann, H. A.M., Vervest, H. A.M., & Heineman, M. J. (2001). Declining trend in major gynaecological surgery in The Netherlands during 1991-1998. Is there an impact on surgical skills and innovative ability? *BJOG: An International Journal of Obstetrics and Gynaecology, 108*(7), 743–748.

Bryman, A., & Bell, E. (2019). *Business research methods* (Fifth edition). New York: Oxford University Press.

Bunce, D., & West, M. A. (1995). Self perceptions and perceptions of group climate as predictors of individual innovation at work. *Applied Psychology, 44*(3), 199–215.

Bunduchi, B. (2009). Implementing best practices to support creativity in NPD cross-functional teams. *International Journal of Innovation Management, 13*(04), 537–554.

Burgess, S., Propper, C., Ratto, M., & Tominey, E. (2012). *Incentives in the public sector: Evidence from a government agency. Discussion paper series / Forschungsinstitut zur Zukunft der Arbeit: Vol. 6738.* Bonn: IZA.

Butter, R. & van Beest, W. (2017). *Psychometric validation of a tool for Innovation Competencies Development and Assessment using a mixed-method design.* Retrieved July 05, 2017, from www.fincoda.eu/fincoda-blog/2017.

Campbell, D. T., & Stanley, J. C. (1967). *Experimental and quasi-experimental designs for research* (2. print). Boston: Houghton Mifflin Comp.

Carmeli, A., & Spreitzer, G. M. (2009). Trust, Connectivity, and Thriving: Implications for Innovative Behaviors at Work. *The Journal of Creative Behavior, 43*(3), 169–191.

Catlin, T., & Lorenz, J. T. (2017). Digital disruption in insurance: Cutting through the noise, 1–59, from https://www.mckinsey.com/~/media/McKinsey/Industries/Financial%20Services/Our%20Insights/Time%20for%20insurance%20companies%20to%20face%20digital%20reality/Digital-disruption-in-Insurance.ashx.

Chen, C.-J. (2009). Strategic human resource practices and innovation performance: The mediating role of knowledge management capacity. *Strategic Direction, 25*(6).

Choi, J. N., & Chang, J. Y. (2009). Innovation implementation in the public sector: An integration of institutional and collective dynamics. *The Journal of applied psychology, 94*(1), 245–253.

Churchill, G. A. (1979). A Paradigm for Developing Better Measures of Marketing Constructs. *Journal of Marketing Research, 16*(1), 64.

Cornell University, INSEAD, & WIPO (2014). *The Global innovation index 2014: The human factor in innovation* (2nd printing). Fontainebleau: INSEAD.

Cornell University, INSEAD, & WIPO (2017). *The Global Innovation Index 2017: Innovation Feeding the World.* Ithaca, Fontainebleau, Geneva.

Coy, P. (2015). *The Bloomberg Innovation Index.* Retrieved June 30, 2015, from http://www.bloomberg.com/graphics/2015-innovative-countries/.

Deacon, D., Bryman, A., & Fenton, N. (2014). Collision or collusion?: A discussion and case study of the unplanned triangulation of quantitative and qualitative research methods. *International Journal of Social Research Methodology, 1*(1), 47–63.

Delgado Piña, M. I., María Romero Martínez, A., & Gómez Martínez, L. (2008). Teams in organizations: A review on team effectiveness. *Team Performance Management: An International Journal, 14*(1/2), 7–21.

DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J., Lazzara, D. J., et al. (2007). A psychometric toolbox for testing validity and reliability. *Journal of nursing scholarship : an official publication of Sigma Theta Tau International Honor Society of Nursing, 39*(2), 155–164.

Dey, P. K. (2006). Integrated project evaluation and selection using multiple-attribute decision-making technique. *International Journal of Production Economics, 103*(1), 90–103.

DoD (1980). *Military standard: Procedures for performing a failure mode, effects and criticality analysis.* [Washington, D.C.], Irvine, CA: U.S. Dept. of Defense; Obtained from Global Engineering Documents.

Dorenbosch, L., van Engen, Marloes L., & Verhagen, M. (2005). On-the-job Innovation: The Impact of Job Design and Human Resource Management through Production Ownership. *Creativity and Innovation Management, 14*(2), 129–141.

Fincoda Project (2018). *How it works.* Retrieved April 13, 2019, from https://www.fincoda.eu/how-it-works.

Fischer, T. M., Möller, K., & Schultze, W. (2015). *Controlling: Grundlagen, Instrumente und Entwicklungsperspektiven* (2., Šuberarbeitete Auflage). Stuttgart, [Germany]: Schäffer-Poeschel Verlag.

Fuchs, W. (2014). *Innovation und Motivation - das Gewinner-Tandem: Ideenfindung als Unternehmensphilosophie.* München: mi-Wirtschaftsbuch.

Furrer, O., Thomas, H., & Goussevskaia, A. (2008). The structure and evolution of the strategic management field: A content analysis of 26 years of strategic management research. *International Journal of Management Reviews, 10*(1), 1–23.

Gale, S. F. (2009). The Real Deal. *PM Network, 23*(12), 30–35, from https://www.pmi.org/learning/library/real-deal-concept-sustainability-2941.

Geissdoerfer, M., Bocken, N., Steingrímsson, J. G., & Evans, S. (2015). Incorporating design thinking into sustainable business modelling. *The Journal of Innovation Impact, 8*.

George, J. M. (2007). 9 Creativity in Organizations. *The Academy of Management Annals, 1*(1), 439–477.

Gerbing, D. W., & Anderson, J. C. (1988). An Updated Paradigm for Scale Development Incorporating Unidimensionality and Its Assessment. *Journal of Marketing Research, 25*(2), 186.

Goffin, K., & Mitchell, R. (2017). *Innovation management: Effective strategy and implementation* (3rd ed.). Basingstoke: Palgrave Macmillan.

Grant, A. M., & Wall, T. D. (2008). The Neglected Science and Art of Quasi-Experimentation. *Organizational Research Methods, 12*(4), 653–686.

Hair, J. F., Babin, B. J., Anderson, R. E., & Black, W. C. (2014). *Multivariate data analysis* (Seventh edition, Pearson new international edition). *Pearson custom library.* Harlow, Essex: Pearson.

Hambrick, D. C. (2007). The Field of Management's Devotion to Theory: Too Much of a Good Thing? *Academy of Management Journal, 50*(6), 1346–1352.

Hauschildt, J., & Salomo, S. (2016). *Innovationsmanagement* (6., überarb., erg. und aktualisierte Aufl.). *Vahlens Handbücher der Wirtschafts- und Sozialwissenschaften.* München: Vahlen.

Heathfield, S. M. (2018). *What is a team?: Form teams to pursue a purpose or a goal with interdependent employees.* Retrieved December 04, 2018, from www.thebalancecareers.com/what-is-a-team-1919226.

Himme, A. (2007). Gütekriterien der Messung: Reliabilität, Validität und Generalisierbarkeit. In S. @Albers, D. Klapper, U. Konradt, A. Walter, & J. Wolf (Eds.), *Methodik der empirischen Forschung* (pp. 485–500). Wiesbaden: Gabler.

Hollanders, H., & Es-Sadki, N. (2017). *European Innovation Scoreboard 2017. Research & innovation policy.* Luxembourg: Publications Office of the European Union.

Houghton, J. (2014). *How to complete your ethics Self-Assessment: Horizon 2020:* European Commission.

Hu, M.-C., & Mathews, J. A. (2005). National innovative capacity in East Asia. *Research Policy, 34*(9), 1322–1349.

Hülsheger, U. R., Anderson, N., & Salgado, J. F. (2009). Team-level predictors of innovation at work: A comprehensive meta-analysis spanning three decades of research. *Journal of Applied Psychology, 94*(5), 1128–1145.

Iacono, W. G., & McGue, M. (2002). Minnesota Twin Family Study. *Twin Research, 5*(05), 482–487.

Innovety (2014). *Innovety LLC - Innovation Management and Operations Excellence > Services > Innovation Readiness Assessments.* Retrieved May 12, 2016, from http://www.innovety.com/site/Services/InnovationReadinessAssessments.aspx.

Janssen, O. (2001). Fairness perceptions as a moderator in the curvilinear relationships between job demands, and job performance and job satisfactions. *Academy of Management Journal, 44*(5), 1039–1050.

Jong, J. d., & den Hartog, D. (2010). Measuring Innovative Work Behaviour. *Creativity and Innovation Management, 19*(1), 23–36.

Jong, J. d., Kemp, R., & Snel, C. (2001). *Determinants of innovative ability: An empirical test of a causal model. Research report / [EIM, Small Business Research and Consultancy]: 0010/A.* Zoetermeer: EIM, Business & Policy Research.

Joppe, M. (2006). *The Research Process.* Retrieved June 04, 2018, from University of Guelph: https://www.uoguelph.ca/hftm/book/export/html/2046.

Kahlfuss, B. (2013). *Partizipative Innovation: Ein Konzept zur Schaffung nachhaltiger Wettbewerbsvorteile.* Hamburg: [Selbstverl.].

Kairisto-Mertanen, L., & Mertanen, O. (2011). Defining innovation competence - the learning outcomes of innovation pedagogy. *INNOVATIONS FOR COMPETENCE MANAGEMENT - Conference Proceedings-Articles, reports and other current publications*, 25–33.

Kanter, R. M. (1996). When a Thousand Flowers Bloom: Structural, Collective, and Social Conditions for Innovation in Organizations. In P. S. Myers (Ed.), *Resources for the knowledge-based economy. Knowledge management and organizational design* (pp. 93–131). Boston, Mass: Butterworth-Heinemann.

Kanter, R. M. (2006). Innovation: The Classic Traps. *Harvard Business Review, 84*(11), 73–83.

Kaplan, R. S., & Norton, D. P. (1996). Linking the Balanced Scorecard to Strategy. *California Management Review, 39*(1), 53–79.

Kauffeld, S., Jonas, E., Grote, S., Frey, D., & Frieling, E. (2004). Innovationsklima - Konstruktion und erste psychometrische Überprüfung eines Messinstrumentes. *Diagnostica, 50*(3), 153–164.

Kleysen, R. F., & Street, C. T. (2001). Toward a multi-dimensional measure of individual innovative behavior. *Journal of Intellectual Capital, 2*(3), 284–296.

Koepsell, T. D. (2005). *Quasi-experimental Design.* Retrieved August 28, 2018, from John Wiley & Sons, Ltd.: https://onlinelibrary.wiley.com/doi/full/10.1002/0470011815.b2a4a037.

Kubicek, H. (1975). *Empirische Organisationsforschung: Konzeption und Methodik. Sammlung Poeschel: Vol. 78.* Stuttgart: Poeschel.

Kurz, V. (2013). *Employee Driven Innovation in Germany: Eine Untersuchung intra-organisatorischer Einflussfaktoren auf das Verhalten von Mitarbeitern* (1. Aufl.). *Stone's publishing Cologne: Bd. 13.* Köln: WiKu-Verlag Verlag für Wissenschaft und Kultur.

Kvale, S., & Brinkmann, S. (2015). *InterViews: Learning the craft of qualitative research interviewing* (Third edition). Los Angeles: Sage Publications.

Lamnek, S. (2002). Qualitative Interviews. In E. König & P. Zedler (Eds.), *UTB Pädagogik, Erziehungswissenschaft: Vol. 8218. Qualitative Forschung. Grundlagen und Methoden* (2nd ed., pp. 157–193). Weinheim, Basel: Beltz.

Laursen, K. (2003). New human resource management practices, complementarities and the impact on innovation performance. *Cambridge journal of economics, 27*(2), 243–263.

Leiponen, A., & Helfat, C. E. (2006). When does distributed innovation activity make sense? Location, decentralization, and innovation success. *ETLA Discussion Papers, The Research Institute of the Finnish Economy (ETLA), No. 1063*.

Leonidou, L. C., Katsikeas, C. S., & Coudounaris, D. N. (2010). Five decades of business research into exporting: A bibliographic analysis. *Journal of International Management, 16*(1), 78–91.

Littkemann, J., & Derfuß, K. (2011). Innovationscontrolling. In S. Albers & O. Gassmann (Eds.), *Technologie- und Innovationsmanagement. Handbuch* (2nd ed., pp. 587–602). Wiesbaden: Gabler.

López-Claros, A., & Mata, Y. N. (2010). Policies and Institutions Underpinning Country Innovation: Results from the Innovation Capacity Index. In A. López-Claros (Ed.), *The innovation for development report 2010-2011. Innovation as a driver of productivity and economic growth* (pp. 3–63). Basingstoke: Palgrave Macmillan.

Lukjanska, R. (2010). Innovation Capacity–Problems and Solutions for Successful Development., *2*, 42–48. Retrieved April 01, 2016.

Madill, A., Jordan, A., & Shirley, C. (2000). Objectivity and reliability in qualitative analysis: Realist, contextualist and radical constructionist epistemologies. *British Journal of Psychology, 91*(1), 1–20.

Maier, G. W., Streicher, B., Jonas, E., & Frey, D. (2007). Innovation und Kreativität. In *Enzyklopädie der Psychologie* (pp. 809–855). Göttingen [u.a.]: Hogrefe, Verl. für Psychologie.

Marin-Garcia, J. A., Perez-Peñalver, M. J., Vidal-Carreras, P. I., & Maheut, J. (2012). How to assess the innovation competence of higher educations students. In *XVII Congreso de Ingenieria de Organizacion* (pp. 920–929).

Marin-Garcia, J. A., Andreu Andres, M. A., Atares-Huerta, L., Aznar-Mas, L. E., Garcia-Carbonell, A., González-Ladrón-de-Guevara, F., et al. (2016). Proposal of a Framework for Innovation Competencies Development and Assessment (FINCODA). *WPOM-Working Papers on Operations Management, 7*(2), 119. Retrieved April 10, 2019.

Marin-Garcia, J. A., Aznar-Mas, L., & Ladrón de Guevara, F. G. (2011). Innovation Types and Talent Management for Innovation. *Working Papers on Operations Management, 2*(2), 25–31.

Martínez-Román, J. A., Gamero, J., & Tamayo, J. A. (2011). Analysis of innovation in SMEs using an innovative capability-based non-linear model: A study in the province of Seville (Spain). *Technovation, 31*(9), 459–475.

Mazars (2017). *Challenges Facing the Insurance Sector: An Interview with AXA Group's CFO Gérald Harlin.* Retrieved April 16, 2019, from Mazars Financial Services blog: http://financialservices.mazars.com/challenges-facing-insurance-sector-interview-axa-group-chief-financial-officer-gerald-harlin/.

Meran, R., John, A., Staudter, C., Roenpage, O., & Lunau, S. (2013). *Six Sigma+Lean Toolset: Mindset zur erfolgreichen Umsetzung von Verbesserungsprojekten* (4., vollst. überarb. u. erw. Aufl. 2013). *SpringerLink.* Berlin, Heidelberg, s.l.: Springer Berlin Heidelberg.

Meredith, J. R., & Mantel, S. J. (2012). *Project management: A managerial approach* (8. ed., internat. student version). Singapore: Wiley.

Michie, J., & Sheehan, M. (2003). Labour market deregulation, "flexibility" and innovation. *Cambridge journal of economics, 27*(1), 123–143.

Möller, K., Menninger, J., & Robers, D. (2011). *Innovationscontrolling: Erfolgreiche Steuerung und Bewertung von Innovationen.* Stuttgart: Schäffer-Poeschel.

Morgan, D. L. (1998). Practical strategies for combining qualitative and quantitative methods: Applications to health research. *Qualitative health research, 8*(3), 362–376.

Mudrak, T., van Wagenberg, A., & Wubben, E. (2004). Assessing the innovative ability of FM teams: A review. *Facilities, 22*(11/12), 290–295.

Muindi, M. (2018). *Making innovation culture a way of life for business.* Retrieved April 16, 2019, from The New Times: https://www.newtimes.co.rw/opinions/making-innovation-culture-way-life-business.

OECD; Europäische Kommission (2005). *Oslo manual: Guidelines for collecting and interpreting innovation data* (3. ed.). *The measurement of scientific and technological activities.* Paris.

Olson, R., Verley, J., Santos, L., & Salas, C. (2004). What we teach students aboth the Hawthorne studies: A review of content within a sample of introductory I-O and OB textbooks. *The Industrial Organizational Psychologist, 41*(3), 23–39. Retrieved March 15, 2018.

Owens, J. A., Dearth-Wesley, T., Herman, A. N., Oakes, J. M., & Whitaker, R. C. (2017). A quasi-experimental study of the impact of school start time changes on adolescent sleep. *Sleep health, 3*(6), 437–443. Retrieved April 27, 2019.

Pappas, R. A., & Remer, D. S. (1985). Measuring R&D Productivity. *Research Management, 28*(3), 15–22.

Persaud, A. (2005). Enhancing Synergistic Innovative Capability in Multinational Corporations: An Empirical Investigation. *Journal of Product Innovation Management, 22*(5), 412–429.

Peter, J. P. (1979). Reliability: A Review of Psychometric Basics and Recent Marketing Practices. *Journal of Marketing Research, 16*(1), 6.

Pew Research Center (2018). *Question Order.* Retrieved April 14, 2019, from https://www.people-press.org/methodology%20/questionnaire-design/question-order/.

Pieterse, A. N., van Knippenberg, D., Schippers, M., & Stam, D. (2010). Transformational and transactional leadership and innovative behavior: The moderating role of psychological empowerment. *Journal of Organizational Behavior, 31*(4), 609–623.

Pittaway, L., Robertson, M., Munir, K., Denyer, D., & Neely, A. (2004). Networking and innovation: A systematic review of the evidence. *International Journal of Management Reviews, 5-6*(3-4), 137–168.

Quadbeck-Seeger, H.-J. (1998). *Faszination Innovation: Wichtiges und Wissenswertes von A bis Z.* Weinheim: Wiley-VCH.

Rack, O., & Christophersen, T. (2007). Experimente. In S. @Albers, D. Klapper, U. Konradt, A. Walter, & J. Wolf (Eds.), *Methodik der empirischen Forschung* (pp. 17–32). Wiesbaden: Gabler.

Ramamoorthy, N., Flood, P. C., Slattery, T., & Sardessai, R. (2005). Determinants of Innovative Work Behaviour: Development and Test of an Integrated Model. *Creativity and Innovation Management, 14*(2), 142–150.

Rammer, C., Crass, D., Doherr, T., & et al (2016). *Innovationsverhaltender deutschen Wirtschaft: Indikatorenbericht zurInnovationserhebung 2015.* Mannheim: ZEW. Retrieved June 09, 2016.

Reuvers, M., van Engen, Marloes L., Vinkenburg, C. J., & Wilson-Evered, E. (2008). Transformational Leadership and Innovative Work Behaviour: Exploring the Relevance of Gender Differences. *Creativity and Innovation Management, 17*(3), 227–244.

Ritchie, J., Lewis, J., McNaughton Nicholls, C., & Ormston, R. (2013). *Qualitative research practice: A guide for social science students and researchers* (Second edition).

Roethlisberger, F. J. (1939). *Management and the worker.* Cambridge Mass.: Harvard University Press.

Rousseau, D. M., Manning, J., & Denyer, D. (2008). Evidence in Management and Organizational Science: Assembling the Field's Full Weight of Scientific Knowledge Through Syntheses. *The Academy of Management Annals, 2*(1), 475–515.

Scandura, T. A., & Williams, E. A. (2000). Research methodology in management: Current pracitices, trends, and implications for future research. *Academy of Management Journal, 43*(6), 1248–1264.

Schentler, P., Lindner, F., & Gleich, R. (2010). Innovation Performance Measurement. In A. Gerybadze, U. Hommel, H. W. Reiners, & D. Thomaschewski (Eds.), *Innovation and international corporate growth* (pp. 299–317). Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg.

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review. *The Journal of Educational Research, 99*(6), 323–338.

Schulz, T. (2014). "Yeah, lasst es uns versuchen": Innovationsmgt bei Google. *Der Spiegel*. (51), 68–70.

Schumpeter, J. A. (1939). *Business cycles.* New York, London: McGraw-Hill.

Scott, S. G., & Bruce, R. A. (1994). Determinants of innovative behavior: A path model of individual innovation in the workplace. *Academy of Management Journal, 37*(3), 580–607.

Sekaran, U., & Bougie, R. (2016). *Research methods for business: A skill-building approach* (Seventh edition). Chichester, West Sussex, United Kingdom: John Wiley & Sons.

Shalley, C. E., & Gilson, L. L. (2004). What leaders need to know: A review of social and contextual factors that can foster or hinder creativity. *The Leadership Quarterly, 15*(1), 33–53.

Shebl, N. A., Franklin, B. D., & Barber, N. (2012). Failure mode and effects analysis outputs: Are they valid? *BMC health services research, 12*, 150–160.

Smith, J. K., & Heshusius, L. *Closing Down the Conversation: The End of the Quantitative-Qualitative Debate Among Educational Inquirers.*

Sobral, J., Teixeira, D., Morais, H., & Neves, M. (2017). Methodology to assess medical processes based on a Failure Mode and Effects Analysis (FMEA). In *2017 IEEE 5th Portuguese Meeting on Bioengineering (ENBENG)* (pp. 1–4). IEEE.

Social Policy Association (2009). Guidelines on Research Ethics.: SPA. 1–8.

Spender, J.-C., Corvello, V., Grimaldi, M., & Rippa, P. (2017). Startups and open innovation: A review of the literature. *European Journal of Innovation Management, 20*(1), 4–30.

Spicker, P. (2011). Ethical Covert Research. *Sociology, 45*(1), 118–133.

Stashevsky, S., Carmeli, A., Meitar, R., & Weisberg, J. (2006). Self-leadership skills and innovative behavior at work. *International Journal of Manpower, 27*(1), 75–90.

Stern, S., Porter, M. E., & Furman, J. L. (2000). *The determinants of national innovative capacity. Working paper series / National Bureau of Economic Research: Vol. 7876.* Cambridge Mass.

Stewart, G. L. (2006). A Meta-Analytic Review of Relationships Between Team Design Features and Team Performance. *Journal of Management, 32*(1), 29–55.

Taylor, J. (2007). In praise of the feasibility study. *Journal of clinical nursing, 16*(10), 1789–1791.

Teng, S.-H., & Ho, S.-Y. (1996). Failure mode and effects analysis. *International Journal of Quality & Reliability Management, 13*(5), 8–26.

ter Haar, P. J. (2017). Measuring Innovation Performance of Work Teams: a State of the Art Review of Approaches and Guiding Framework for Selecting Indicators. In P. Vrgovic, A. Pilov-Velic, & C. Dreher (Eds.), *1st Doctoral Studens Conference on Innovation Management* (1st ed., pp. 48–57). Novi Sad: Faculty of Technical Science.

The Economist (2014). *Business Environment Rankings: Which country is best to do business in?* Retrieved April 11, 2016, from The Economist: http://www.eiu.com/public/thankyou_download.aspx?activity=download&campaignid=bizenviro2014.

Thompson, L. L., & Choi, H.-S. (2006). *Creativity and innovation in organizational teams. LEA's organization and management series.* Mahwah, N.J.: Lawrence Erlbaum Associates.

Tidd, J., & Bessant, J. R. (2014). *Strategic innovation management.* Chichester, West Sussex: Wiley.

Trott, P. (2017). *Innovation management and new product development.* Harlow, Enlgand: Pearson.

Vahs, D., & Brem, A. (2015). *Innovationsmanagement: Von der Idee zur erfolgreichen Vermarktung* (5., überarbeitete Auflage). Stuttgart: Schäffer-Poeschel.

van Deventer, J. P. (2009). Ethical considerations during human centred overt and covert research. *Quality & Quantity, 43*(1), 45–57.

Villa, L. S. (1990). Invention, inventive learning, and innovative capacity. *Behavioral Science, 35*(4), 290–310.

Warschat, J. (2005). *Der Weg zur Innovationsexzellenz: Vorlesungsunterlage vom 15.02.2005.* Stuttgart.

Wegmann, C., & Winklbauer, H. (2006). *Projektmanagement für Unternehmensberatungen: Mit Beispielen aus dem Inhouse Consulting von Deutsche Post World Net* (1. Aufl.). Wiesbaden: Betriebswirtschaftlicher Verlag Dr. Th. Gabler / GWV Fachverlage GmbH Wiesbaden.

Weisstein, E. W. (2018). *Gaussian Function.* Retrieved January 28, 2018, from MathWorld--A Wolfram Web Resource: http://mathworld.wolfram.com/GaussianFunction.html.

Yuan, F., & Woodman, R. W. (2010). Innovative Behavior in the Workplace: The Role of Performance and Image Outcome Expectations. *Academy of Management Journal, 53*(2), 323–342.

Zhou, J., & George, J. M. (2001). When job dissatisfaction leads to creativity: Encouraging the expression of voice. *Academy of Management Journal, 44*(4), 682–696.

Zhou, J., & Shalley, C. E. (2004). Research on emploayee creativity: a critical review and directions for future research. In J. J. Martocchio (Ed.), *Research in personnel and human resources management: v. 23. Research in personnel and human resources management* (pp. 165–217). Amsterdam, London: Elsevier JAI.

Žižlavský, O. (2016). The Use of Finanical and Nonfinancial Measures within Innovation Mangament Controll: Experience and Research. *Economics and Sociology, 9*(4), 41–65.