

The Use of Mixed-Methods When Analysing Work Components in Police Investigation

Lars-Martin Berglund

Master of Work- and Organizational Psychology



Department of Psychology

University of Oslo

May 2012

Acknowledgements

First and foremost, I would like to thank the research group of Work and Organizational Psychology at the University of Oslo for two years of excellent education. I will address my sincere gratitude to my main counsellor Roald Bjørklund and co-counsellor Jon Anders Lone for complementary guidance and a motivational team effort during the process of writing this thesis. Your guidance and advices have been of immense importance to me. This project had never seen the light of day without Trond Myklebust who in an outstanding manner represented the Research Department at the Police University College. His engagement and interest in the research process was admirable and fundamental to this thesis as our cooperation resulted in a unique data set. My fellow students throughout the two years as a master student have provided me with valuable friendships, thank you all. Especially, I would like to thank Martine Berg Hannevik for all the “great academic discussions”, they surely represent the silver lining. My teammates of the Investigation-Project Ane Cecilie Helland Volle, Ellen Jorunn Bergem and Vivian Therese Abildgaard Boodhun, you made a demanding process easy with bright ideas, humour and personalities that I am deeply thankful for having cooperated with.

Without my dear family, a five years education beyond upper secondary school, would have been improbable, dear Grete, Lars and Ole Magnus, thanks for the love throughout the years, and the everlasting support, it means everything to me. Of course, I have to thank my dear housemates Lars Martin Tannæs-Fjeld and Jonas Øie for providing me with a social life in times where I was totally devoted to this project.

Lars-Martin Berglund, May 2012, Oslo

Table of Contents

Acknowledgements	ii
Table of Contents	iii
Abstract	1
The Use of Mixed-Methods When Analysing Work Components in Police Investigation	2
Philosophical Approach to Data Collections	3
Measurement of Work.....	4
Prior Mixed-Methods	5
The Present Study.....	7
Aim of the Study	10
Method	11
The Research Project.....	11
Preparations	11
Sample.....	11
Measures.....	12
Procedure.....	12
Data Treatment and Statistical Analyses.....	15
Comparison of Measures.....	17
Ethical Considerations.....	17
Results	18
Descriptive Statistics	18
Inferential Statistics.....	22
Discussion	25
Summing Up Results.....	25
General Discussion.....	25
Theoretical Implications.....	30
Practical Implications	31
Limitations	32
Future Studies.....	34
Conclusion.....	34
References	36

Appendix A - The SWOT Interview Guide	41
Appendix B – OCM Questionnaire	42
Appendix C – Personal Communication with Michael West	46
Appendix D - Interview Letter of Consent.....	47
Appendix E - Coding Scheme	49

Abstract

The aim of this study was to statistically investigate the use of mixed-methods by examining open-ended interviews and closed-ended survey instruments, and their ability to generate convergent results on the same sample. The use of mixed-methods was studied by applying a general measurement of work, the Organizational Climate Measurement (OCM), on the context of police investigation. Open-ended interviews were conducted with informants from the Norwegian police. The interviews were top down coded on the OCM components and standardized. OCM surveys were sent to the interviewed participants, and the results were standardized to match the standardization of interviews. Correlations between the results from the two data collections were weak and insignificant, but correlations were found inaccurate as a measure of convergence. Paired t-tests on the interview and survey means of the 17 OCM components showed that the two measurements converged on 13 components and diverged on four. These results show that interviews can be a viable framework for the measurement of work, and show that open-ended interview and survey results can converge. The procedure in this study could therefore provide scholars and practitioners with a validation tool for both interview and survey studies on specific work contexts.

The Use of Mixed-Methods When Analysing Work Components in Police Investigation

The context of work is changing, and scholars are rapidly redesigning theories that explain and measure work (Grant & Parker, 2009). Therefore, researchers in work and organizational psychology continuously debate which features, characteristics or components to investigate (Parker, Wall, & Cordery, 2001).

A large proportion of work models are examined through closed-ended surveys. There are several reasons to why researchers choose quantitative methods. Traditional quantitative data collections are easy to administer, time-efficient, and software takes care of the analysis. Also, in regards to reliability and validity, quantitative instruments benefit from a vast collection of statistical tests that are easily conducted with large samples. However, studies have shown that closed-ended survey research is limited as the measurements often fail to account for important components experienced by workers (Keenan & Newton, 1985; Mazzola, Schonfeld, & Spector, 2011). As noted by Kuhn (1962), survey research is circular and therefore fosters no innovation or new ideas. The surveys consist of components and their respective items, and themes external to these components are not accounted for (Mazzola et al., 2011). Furthermore, participants' answers are restricted within the surveys response format (Cronbach, 1946).

Open-ended interviews, on the other hand, do not restrict participants' answers. General strengths of open-ended interview data are that they allow in depth exploration of individuals' evaluation of experiences encountered at work, and a clearer picture of relevant themes in specific contexts (Narayanan, Menon, & Spector, 1999). Further, interview research can play a role in the discovery of important work components that are not covered by structured survey instruments (Kidd, Scharf, & Veazie, 1996; Mazzola et al., 2011). However, work contexts often consist of a multitude of workers, and qualitative measures are insufficient as they most often lack representative samples (Mazzola et al., 2011). In addition, qualitative data collection and analysis is time-consuming, subjective and difficult to prove reliable and valid.

As both methods for data collection have their respective strengths and weaknesses, why constrain research to mono-methods? The use of open-ended and closed-ended data collections in conjunction elicits a method that overcomes several of the weaknesses of the two methods (Jex, Adams, Elacqua, & Lux, 1997; Mazzola, Walker, Shockley, & Spector, 2011). Advantages of mixed methods designs are that they can answer more complex research questions than isolated mono-methods, and that results from mixed data collections can

complement each other. Convergent results can increase the validity of each data collection and analysis, while divergent or contradictory findings can generate new theoretical insight, further research and revised hypothesis (Lund, 2011).

Mixed-methods may be defined as “the collection or analysis of both quantitative and qualitative data in a single study in which the data are collected concurrently or sequentially, are given a priority, and involve the integration of the data at one or more stages in the process of research (Creswell, Plano Clarke, Guttman, & Hanson, 2003, p. 212). An important aspect of this definition is the distinction between quantitative and qualitative data. Lund (2005) argues that these terms should be dropped, as qualitative and quantitative methods are corresponding variants of obtaining empirical research, and inferences depend on analyses of the collected material, rather than the method of data collection itself. Interviews and surveys are both self-reports of participants’ opinions. The difference is the process of analysis. Interviews concern the researchers rating or interpretation of interviewees statements, whereas closed-ended surveys consist of participants’ rating of the researchers items. However, open-ended data collections are not confined within a qualitative analysis, as open-ended sources of information may be quantified and match the standards of survey research (Neuendorf, 2002). Therefore, sources of survey and interview data could be used in statistical mixed-method comparisons. However, there is a gap in the existing literature regarding the practical procedure of combining different sources of data collections. So far, an attempt to statistically compare results of data collections consisting of open-ended interviews and closed-ended surveys is abundant.

Philosophical Approach to Data Collections

The differentiated philosophical approach to data collections illustrates how established traditions for scientific worldviews has hindered the use of mixed-methods (Wiggins, 2011). In recent philosophical traditions there are two distinct philosophical assumptions, the constructivist or interpretative view, and the positivist tradition. Researchers in the constructivist and the positivist traditions differ in their ontology, the way they perceive the world and the nature of it, and their epistemology, their concern with how they know what they know (Hanson, 2005). The constructivist ontology claims that reality is socially constructed by subjective interpretation. Berger and Luckmann (1967) among others, argue that there is no such thing as objective knowledge and facts. In regards to epistemology, the constructivists often utilize qualitative inductive interpretation, small samples, and research is most often conducted in non-generalizable studies. Rather, the studies retrieve depth and

nuances of context-specific complex phenomena. On the other hand, the positivist ontology assumes that objective entities are observable and describable facts. Thus, its epistemology utilizes deductive hypotheses and research designs with quantitative interpretation of large samples in statistical analyses with inferences that are broad and generalizing.

A common fallacy regarding methods of data collection is that they belong to either a positivist or constructivist assumption. Researchers within these traditions have argued and favoured methods of obtaining data, treating epistemology and methods as being synonymous (Bryman, 1984). The perspective of methods within epistemologies has led to the thesis that qualitative and quantitative data collections belong to distinct philosophical traditions, and that they are incompatible (Bednarz, 1985; Forshaw, 2007; Howe, 1988; Ogborne, 1995; Simpson & Eaves, 1985).

In this paper it is argued that methods of data collection do not belong to a certain epistemology. Rather, the appropriate methods should be used according to the research question regardless of philosophical traditions, thereby acknowledging the “third wave” within the research movement, namely pragmatism (Johnson, 2004). This movement is characterized by prioritizing the research question in empirical studies, and by combining different methods of data collections when answering such questions (Lund, 2011). In this way, the thesis of incompatibility is opposed, and a complementary use of data collections is recognized as mixed methods have unique strengths that may overlap the respective weaknesses of mono-methods (Mazzola et al., 2011). The complementary use of data collections could be beneficial to the measurement of work, as researchers have argued several challenges of work measurements.

Measurement of Work

A common problem with present work instruments is a lack of empirical support for the prevalence of the components measured (Taber, Beehr, & Walsh, 1985), and another problem is whether the items of these components actually reflect the characteristics they are intended to measure. Two examples are the Job Characteristics Model by Hackman and Oldham (1976) and the Demand-Control model by Karasek (1979). Both are general models of work design that are previously validated and claimed to be applicable across a majority of work contexts. However, these models have been criticized for their relevance in certain work contexts as they fail to capture salient work components important to employees in specific work contexts. Situation specific measures have therefore been suggested, as they uncover the

unique characteristics of specific work contexts (Bakker & Demerouti, 2007; Parker et al., 2001; Sparks & Cooper, 1999)

To improve survey instruments as the Demand-Control and the Job Characteristics models, several researchers have advocated the use of qualitative methods, as they complement the limitations of survey studies on work (Bakker & Demerouti, 2007; Sparks & Cooper, 1999). Schneider, Wheeler, and Cox (1992) established a quantitative survey for service climate through a “bottom-up” analysis of interviews utilizing prevalent concepts, i.e. frequent themes in interviewees’ reflections, as components of the measurement. However, creating “bottom up” models for measurement of specific contexts may be too exhaustive and time consuming for most work contexts. A problem with specific measures are that they increase the complexity of models intended to measure work, and the process of specifying unique components to the context does not necessarily guarantee that results of measures are valid. van Veldhoven, Taris, de Jonge, and Broersen (2005) suggests that general models are the best point of origin for specific models, and argue that components of general models may be altered according to the characteristics of the context. Spector and Jex (1991) propose that further research on job components should be conducted on independent and additional sources of data, e.g. interviews without priming the participant on the studied components, and subsequently rate interviews on the components of interest. As such prevalence of the characteristic components could be assessed. In consistence with van Veldhoven et al. (2005), such an approach would give empirical support for the specifications of the general measures. Roberts and Glick (1981) propose the use of multiple methods when defining job components, as this will give the researchers an indication of whether the components, as well as the items, reflect the employees’ experiences and opinions in the context under study. Past research has shown that survey assessments of work environment tend to converge poorly with other data sources (Cortina, Schmitt, & Whitney, 1992; Spector & Jex, 1991). However, more recent attempts have succeeded in generating convergent results of different data collections in stress research.

Prior Mixed-Methods

As mentioned an open-ended mono-method study could investigate the prevalence of survey components. However the items of the instruments, or the components ability to reflect participants’ experiences, will not be assessed by such an approach. Such assessments require multiple methods, or mixed methods. Studies that compare results of open-ended and closed-ended information have been conducted. Jex et al. (1997) compared the responses of an open-

ended questionnaire with closed-ended questions on the same sample, to test if open-ended measurements converged with closed-ended measures. Open-ended material was content analysed “top-down” on established stress theory components. The results showed that the components of open-ended questions displayed reasonable convergence with the similar closed-ended components. This study indicated that open-ended and closed-ended data collections were comparable and should be used in conjunction. In addition, open-ended methods were found to be a viable approach to measure work related components. In another study, Mazzola et al. (2011) mixed-methods by applying established components of stress in a survey measure that comprised closed-ended questions, as well as using the same open-ended questions as Jex et al. (1997). The qualitative data from the open-ended survey questions were both “bottom-up” and “top-down” coded, using conceptual components from findings of previous research on stress and strains, thus creating new categories when components did not account for the interview material. The open-ended questions concerned incidents of stress components and strains, or outcomes of stressful incidents. These outcomes, or strains, were analysed using frequency, i.e. prevalence and correlations with stress components, and results showed that stressor components were related to strains. Participants who responded high on incidents of open-ended stressors scored overall higher on the frequency of the same stressors in the closed-ended measure. This indicated partial evidence of a relationship between open-ended and closed-ended methods on individual levels, as well as the aggregated sample. They therefore concluded, in consistence with Jex et al. (1997), that the two data collections should be used in conjunction and that they are not interchangeable. In addition, a complementary use of open-ended and closed-ended methods was therefore recommended.

To summarize, the two studies investigated established components’ prevalence, the components validity, and the convergence, i.e. the ability of open-ended and closed ended questions to reflect agreement between measures on the same sample. In addition they found open-ended questions to be a viable framework for measurement of work, thus recommending the use of mixed-methods. However, the two data collections of the studies above were applied on the same assessment. Participants responded on a closed-ended survey before they wrote answers on the open-ended questions on the same survey sheet. This may have caused a priming effect, as well as survey fatigue, which could impact the response of participants. It is possible that attempts without this priming effect would produce the opposite of convergence, namely divergent results. This could imply that the inferences of measurements were invalid and indicate that open-ended measures are not viable measurements of work. Divergence is not exclusively negative, divergence may in fact increase the theoretical insight (Lund, 2011).

In addition, divergent results may indicate that a measurement is biased (Spector, 2006). The Multi-Trait Multi-Method by Campbell and Fiske (1959) proposed the idea of comparing mono-method measures and investigate the discriminant and construct validity of concepts i.e. the components under study. High correlation of two similar measures on a construct indicated valid constructs, whereas low correlations indicated invalid constructs or “apparatus factors” also known as common method variance. As the Campbell and Fiske (1959) method concerned comparisons of similar data collections, and not two different methods of data collections, the terminology of mixed-methods does not comply with their procedure. However, mixed-methods should have the potential of discovering biases, which is often related to common method variance. According to Spector (2006) the term common method variance should be replaced by an emphasis of specific biases, as the reason to measures variance may be more nuanced than the mere method of data collection and that there may be plausible alternative explanations for the observed phenomena than common method variance.

Lund (2011) points out that a study may be validated, however a test in itself may not be valid. This notion is aimed towards earlier validated work instruments, as their prior validity may not apply in the next study. The easy way of explaining validity is “are we measuring what we want to measure” (Neuendorf, 2002 p.112). As such, the mixed methods may play a role in confirming the two measures validity, as a convergence between measures could indicate that they generate equal results and that the measures reflect workers opinions regarding the studied work components. A disagreement could indicate that either of the two measures taps something else than the intention of the component. This is important as invalid studies of contexts may create fallible inferences. One of the strengths in the use of mixed methods is its ability in investigating the validity of measures (Lund, 2011). In this study, validity is important as the studied context, police investigation is understudied in regards to salient work components and there are few studies to compare findings and inferences with. In addition, former literature on the convergence of open-ended interviews and surveys is minimal, if not absent.

The Present Study

Narayanan et al. (1999) found that salient components of work depended on occupation. Some components tended to be prevalent in several contexts, but specific components were relevant in different occupations. It is therefore probable to assume that the context of police investigation has common characteristics of other work contexts, and some characteristics that

are specific to this context. Luen and Al-Hawamdeh (2001) regard policing, especially police investigation, as knowledge-intensive work. Prior studies of investigative work have been measured as knowledge-intensive by transferring instruments from other domains onto the investigation context. For example, Glomseth, Gottschalk, and Solli-Sæther (2007) concluded that a measure intended for the context of law firms, the value shop instrument by Stabell and Fjeldstad (1998), was applicable to the measure of police investigation as “policing institutions are experiencing higher demands on performance while working within tight resource constraints” (Home Office, 2005, p. 97). The assumption of an instrument as face-valid implies that the instrument components cover the salient characteristics of investigation. The value shop measure has not been recognized within the work design literature as a general measure, to the contrary it appears situation-specific. As mentioned above, a common problem with instruments of work measurement are that they lack empirical support for being relevant for the context under study. As such it is natural to question the applicability of a lawyer context intended measure for police investigation.

Based on the inferences of van Veldhoven et al. (2005), a general measure should be applied and made specific according to the analysis of employees in the work context. A rather new, general instrument that intends to capture salient components of several work contexts is the Organizational Climate Measurement (Patterson, West, Shackleton, Dawson, Lawthorn, Maitlis, Robinson, & Wallace, 2005). Prior research on the prevalence of the general Organizational Climate Measurement (OCM) components in the Norwegian knowledge-intensive public domain have proved the measure to be of some relevance (Hønsen, 2010). Thus, the OCM components could be relevant in police investigation, which is a public domain argued by other researchers to be knowledge-intensive. However, whether the components match interviewees’ perception of police investigation cannot be determined until interviews “top-down” coding of the general model has been conducted.

The OCM. The OCM comprises 17 general work components (see method section for components and description) that Patterson et al. (2005) found to be frequently used as characteristics of employees’ experiences of work from research studies on work climate between 1960 and 2000. In addition, the OCM is previously empirically validated (Patterson et al., 2005). The creators claim the OCM to be a global construct that provide an overall snapshot of the operational functioning of studied organizations, and that it is a valid basis for the investigations of work environment perceptions across most work contexts. An important aspect regarding OCM is the composition of component items, which refer to the organizational level of the organization. The concept of the component may address leaders,

development, autonomy etc. in the organization, yet in a wording that taps the rater's organizational view on the components presence in the organization.

Patterson et al. (2005) have shown that the OCM discriminates effectively between organizations, and generate agreement within the organizations on the components. As such, it fulfils the requirements of a general work model. Furthermore, the instrument can be refined by selecting those components that are relevant to the research question when conducting survey research (Patterson et al., 2005). This is fortunate in consideration to the van Veldhoven et al. (2005) argument of customizing general measures in order to assess unique characteristics of contexts.

The mixed-methods. This study compares open-ended interviews with closed-ended survey results of the OCM components in a mixed method design. Only two known studies have previously compared components on open-ended and closed-ended questions (see Jex et al., 1997; Mazzola et al., 2011). This study differs from the two known studies, as the source of open-ended questions is interviews, which encourage free reflections cued on the SWOT interview framework (see method section) regarding the organizing of police investigation. Thus, emergent concepts from interviews are unbiased of researchers preconceptions of work characteristics. Further, interviews are conducted prior to surveys, and separated in time, so that the suggested priming and fatigue effects from previous studies is avoided.

The interviews are a-priori content analysed, according to Neuendorf (2002), by "top-down" coding the OCM components. This is in correspondence with the Jex and Spector's (1991) proposal for future studies, and will analyse if the components of the OCM are prevalent to police investigation. The results of interview coding are standardized to generate comparable results with standardized survey results. This elicits statistical tests of the two measurements' convergence on components. Mazzola et al. (2011) and Jex et al. (1997) concluded that open-ended questions were a viable framework for measurement of work. This study elaborates on these earlier conclusions and assesses open-ended interview questions ability to measure work. The comparison of measures will generate information of both measurements independent of convergent or divergent results. The use of mixed methods can reveal possible biases of measures, general differences of methods and misinterpretations in coding of open-ended data. Mixed-methods is often associated with validity studies, a convergence between open-ended interviews and closed-ended surveys could validate the interview data and the interpretation. Furthermore, a convergence between survey and interview data could imply that the OCM survey instrument reflect participants' opinions. As the sample consists of three different occupations within the police, it is predicted, in

consistence with Narayanan et al. (1999), that individual values will differ in regards to how participants perceive the work components. In addition, the OCM discriminate between organizations and show that participants agree on components within the organization, (Patterson et al., 2005). This study's sample consists of participants from different districts in the Norwegian police organization, who were asked to reflect on the investigation in their district, as well as rate their district in surveys. It is therefore predicted that individual values will differ on the components, resulting in weak correlations. Paired t-tests between measures on components sample mean are therefore predicted to be more accurate analysis as they capture the aggregated variation of scores between open-ended interviews and surveys.

Aim of the Study

The aim of this study is to statistically investigate the use of mixed methods by examining open-ended interviews and closed-ended survey instruments, and their ability to generate convergent results on the same sample. The general work environment measure, the OCM, is tested on the specific context of Norwegian police investigation. This study differs from other studies as the “top-down” measure of interviews has yet to be compared statistically with equal components of a closed-ended survey instrument. This study could add information about the capabilities of open-ended interviews and the survey, possible biases of measures and the validation properties of mixed-methods. Moreover, it is reasonable to assume that findings will indicate how mixed-methods could enrich work and organizational research of specific contexts and the applicability of general measures as the OCM.

The research question of this study is:

Are open-ended interviews and surveys capable of generating convergent results?

Method

The Research Project

This study is part of a long-term project between the department of work and organizational psychology at the University of Oslo and the research department at the Norwegian Police University College. Data collection began 28 of April 2009. This project is known as the Investigation-Project and is funded by the Norwegian Ministry of Justice. The purpose of the project is to increase scientific knowledge regarding police investigation.

Preparations

Before the author entered the project 39 interview records were transcribed, unitized and coded onto SWOT and IGLO categories (see below) by research assistants. The author read articles and theses based on the methodological framework to assure competence in the different proceedings. To understand the organisation and its context, conditioning documents on system of agreement were revised before empirical data was approached.

Sample

Interview sample. The Norwegian police organization encompasses 27 districts. The original sample consisted of 51 strategically selected participants from 16 districts. Geographical dispersion was a criterion for pseudo-random selection of districts. The police commissioners in selected districts were personally addressed through a letter that contained a note of informed consent, a request for participation in the project by the commissioner, and his or hers compliance in arranging interviews with a principle investigator and an investigator. Hence, selections of the commissioners were pseudo-random whilst the commissioners on behalf of the strategic prerequisites, randomly selected other participants.

Survey sample. When the interviews were completed, the 51 interviewed participants received personal letters by mail with the survey. Of the 51 invited participants, 33 responded, 65% of the total sample.

Mixed-method sample. The 33 survey responding participants with their respective interviews were included in this study. Participants that did not respond to the survey were removed from the analyses.

Measures

The interviews. Interviews were conducted according to the structure of the PEACE model, which refers to Planning and Preparation, Engage and Explain, Account, Clarification and Challenge, Closure and Evaluation (Clarke & Milne, 2001). Interview data was gathered through four open-ended semi-structured questions based on the acronymic SWOT framework: Strengths, Weaknesses, Opportunities and Threats. SWOT is originally a strategic tool for organizational development, but can also be applied to individuals (Helms & Nixon, 2010). The four questions prompts free associations and reflections by the interviewee on the general topic “organizing of investigation”. The questions asked consist of few cues, and previous research has shown that SWOT is a viable framework for individual interviews (Hoff, Straumsheim, Bjørkli, & Bjørklund, 2009). Follow-up questions such as paraphrasing on themes mentioned by the subject, and probes such as “other elements you want to mention” and “other strengths you remember” were consciously applied by the interviewer. For the interview guide see Appendix A.

Survey. The questionnaire (see Appendix B) was a wording modified version of the Norwegian OCM (Bernstrøm, 2009), which is a direct translation of the English OCM (Patterson et al., 2005). Words as “market”, “company”, “the boss” and “client” were predicted as incomprehensible in the context of the police. Changes were proposed and made in cooperation with associate professor Cato Bjørkli, professor Roald Bjørklund, PhD candidate Jon Anders Lone and PhD Trond Myklebust. Wording was changed after e-mail correspondence with OCM researcher M.A. West (personal communication, November 17 & 21, 2011) (see Appendix C). West argued that the changes are not supposed to reduce the reliability of the measurement. To the contrary, it maintains the precision of the measurement due to the customization.

Procedure

Interviews. Informants received an e-mail in advance that contained the date of the interview, a brief explanation of the project and the four interview questions. Interviews were conducted in Norwegian and performed by an expert on cognitive interviewing and the context under study. This was a beneficial circumstance as interviewing rests on the skills and the judgment of the interviewer (Kvale & Brinkmann, 2009). The author participated as co-interviewer in four interviews. Interviews were conducted at different police stations depending on the geographical placement of the police commissioner. The interviews took place in offices or meeting rooms that were quiet and private and there was no time limit, so closure of the

interview occurred when the informant had nothing further to add. Interviews were recorded with two digital recorders to prevent loss of data. Average duration of interviews was 65.43 minutes, where the shortest interview was 31.27 minutes and the longest interview was 131.40 minutes.

Surveys. Surveys were sent to participants from the interviews. Thus, 51 surveys were sent out by the police postal system the 2-12-2011. The author received a control letter the 5-12-2011, which served as an indicator of questionnaire arrival at the offices. The questionnaires were sent to employees at their place of work in personally addressed envelopes, which contained a personal letter, an instruction, the survey, and an addressed return envelope. A reminder was sent by mail the 5-1-2012 to increase response. Data collection was terminated the 2-2-2012.

Transcription of recordings. Research assistants transcribed 39 interviews. Therefore, a quadruple transcription reliability analysis was conducted on three randomly selected interviews. Differences that altered semantics were not discovered. The author and fellow students transcribed recordings where they co-interviewed. Of the total transcription sample the author transcribed four interviews. Before transcribing the author and fellow students developed a shared understanding for transcribing procedure. The procedure was formalized and approved by the Research group. Recordings were transcribed verbatim in Norwegian Bokmal. Thus, local expressions and dialect were translated. In consent we decided to leave out contextual matters as coughs, pauses or other elements that had no effect on the manifest meaning. Repetition was included despite its lack of impact on semantics (Krippendorff, 2004).

Content analysis. Interview transcripts were systematically analysed, governed by the requirements of an a priori content analysis research design (Neuendorf, 2002) Content analysis was approached as a procedural research technique for an explicit, replicable, systematic multistep quantification process. Collected data were manifest classified in a priori categories, eligible for statistical analysis (see e.g. Berelson, 1952; Schneider et al., 1992; Neuendorf, 2002; Krippendorff, 2004). Further, this method allows drawing of valid inferences from texts when ensuring a reliable process (Krippendorff, 2004).

Coder training. A crucial step in the coding process is coder training (Neuendorf, 2002). Therefore, the author and fellow students received extensive training prior to coding of the included data.

Coding scheme. The first step in the content analysis process was development of coding schemas. Procedures for Unitizing, SWOT, IGLO and OCM were established in

consensus meetings to create a formalized and shared understanding among the multiple coders. See appendix E for coding schemas.

Unitizing. Secondly, the author unitized five transcriptions. Unitizing concerns the process of breaking text into statements or units of text. Transcripts were divided into meaningful statements, where units were defined by membership in a category or a class. This is known as unitizing by categorical distinctions (Krippendorff, 2004). A statement was defined as the smallest meaning-bearing unit that expresses a logical, consistent and separate view. A statement could be a sentence, a piece of a sentence or several sentences. A statement should be small enough to contain only one coherent description of a theme, yet comprehensive enough to be meaningful (Hoff, 2009).

Coding. Data were coded in three models. Two of these, were explanatory the models SWOT and IGLO whilst the OCM was a normative and validated research model for work climate. The author coded five interviews in SWOT and IGLO and the 51 interviews (12429 statements) in companionship with a fellow student, though the coding was done independently. The description below explains the codes.

The SWOT: *Strengths:* Positive statements that describes the present “here and now” status in the organization. *Weaknesses:* Negative statements that describes the present “here and now” status in the organization. *Opportunities:* Positive statements regarding the future of the organization. *Threats:* Negative statements regarding the future of the organization. The SWOT statements were coded based on the dimensions: Positive/negative, present/future. First and foremost the coders emphasized positive/negative, then made decisions regarding the time perspective present/future.

The IGLO: These codes comprise the level in organization the interviewee reflected on and referred to. *The individual level:* Individual perceptions, feelings and opinions regarding oneself or another individual. *The group level:* Interaction and cooperation in work groups, teams and departments. *The leadership level:* Behaviour of immediate supervisors, other leaders or the top management. *The organizational level:* Management practices, organizational culture, strategies, organizational goals and values, and the physical environment of the organization. IGLO External: Statements directed towards external matters, as letters from the director of public prosecutions.

The OCM: Data accounted for by SWOT and IGLO were further analysed through the normative and validated model OCM. The 17 OCM categories were extracted from Patterson et al. (2005, p. 385). The subsequent description was included in the codebook to illuminate the concept and specify what it comprised. 1. *Autonomy:* Designing jobs in ways,

which give employees wide scope to enact work. 2. *Integration*: The extent of interdepartmental trust and cooperation. 3. *Involvement*: employees have considerable influence over decision-making. 4. *Supervisory Support*: The extent to which employees experience support and understanding from their immediate supervisor. 5. *Training*: a concern with developing employee skills. 6. *Welfare*: The extent to which the organization values and cares for employees. 7. *Formalization*: A concern with formal rules and procedures. 8. *Tradition*: The extent to which established ways of doing things are valued. 9. *Innovation & Flexibility*: flexibility—an orientation toward change; innovation—the extent of encouragement and support for new ideas and innovative approaches. 10. *Outward Focus*: The extent to which the organization is responsive to the needs of the customer and the marketplace in general. 11. *Reflexivity*: A concern with reviewing and reflecting upon objectives, strategies, and work processes, in order to adapt to the wider environment. 12. *Clarity of Organizational Goals*: A concern with clearly defining the goals of the organization. 13. *Efficiency*: The degree of importance placed on employee efficiency and productivity at work. 14. *Effort*: How hard people in organizations work towards achieving goals. 15. *Performance Feedback*: The measurement and feedback of job performance. 16. *Pressure to Produce*: The extent of pressure for employees to meet targets. 17. *Quality*: The emphasis given to quality procedures.

Data Treatment and Statistical Analyses

Interviews. The statements were transferred to the statistical computer program PASW 18 (Predictor Analysis Software). Each statement was counted as a case. Further, these cases were “top-down” or “a priori” analysed i.e. interpretation were based on formalized categories. The statements were coded on SWOT, IGLO and OCM variables. All of these variables included a category called residuals. This category accounted for redundant information, that did not fit or concern the codes e.g. contextual matters from the interview, history, neutral information or themes not covered by the conceptualization of the codes. The majority (39) of the SWOT and IGLO coding were conducted by research assistants.

Intercoder reliability. Coding reliability was assessed to ensure that proceedings in, unitizing, coding of SWOT, IGLO and OCM was reliable (Krippendorff, 2004). A segment of the unitized data set was compared with a segment of the same recording unitized by the author, the P.A.P (Zharghooni, 2011) was estimated to 58%. Further, the SWOT and IGLO coding of the student group and the dataset were compared on three different interviews. A total of five coders assessed SWOT and IGLO coding reliability independently on the same

source. The SWOT rating of agreement was 0,70 on Krippendorff's alpha (77,68%) and the range for these interviews was 0,66 – 0,78 (74,36% - 83,03%) in three interviews. The IGLO scores were lower than the SWOT with an average of 0,43 in Krippendorff's Alpha (73,58 %) the range was: 0,31 – 0,58 (72,02 % - 75,33). The OCM was rated by 2 coders and measured in the midst of the coding process and at the end. The average agreement on categories was: 0,41 in Krippendorff's Alpha (50,83 %) and the range between the different estimates was: 0,37 – 0,46 (47,75% - 52,53%). The agreement in match on OCM or residuals was on an average 0,29 in Krippendorff's Alpha (72,03 %) and the range was: 0,21- 0,35 (67,26 % - 76,77 %). As such, the explanatory and normative models differed in reliability although the coders received the same amount of training, consensual meetings and time to establish coding schemas. Inter-rater reliability was calculated in the Freelon intercoder reliability web service (Freelon, 2010).

Survey. As the questionnaires were returned, the author manually plotted the response Values, along with their unique reference number into a PASW 18 file. Further, the 32 reversed items were recoded into new compatible variables. The 8 missing values were replaced by series of mean. Then, the 17 OCM factors of 82 items were computed as in Patterson et al. (2005)

Survey reliability. Cronbach's Alpha was not estimated, as the sample was too small. However, prior studies of the Norwegian OCM have found sufficient Chronbach's Alpha (Bernstrøm, 2009).

Interview data. Statements accounted for by the SWOT, the Organizational level of IGLO and OCM categories of the 33 participants who responded on the survey were brought into further proceedings. The statements on individual, group and leadership level of IGLO were removed from the comparison as the OCM survey taps the organizational level of work. Residuals of the coded categories were also removed. The reason was to compare prevalent OCM data from the same sample of participants on both measures. Therefore, the applied survey results in the comparative analyses depended on thematic coverage by reflections in the interviews. Consequently, if 18 people spoke about Autonomy, the respective surveys of these participants were preceded on the characteristic component to avoid biased comparisons

Standardization of interview data. Open-ended interview data was standardized through a transformation process in order to be comparable of survey data. Strengths and opportunities of the SWOT codes were computed as positivity variables, whereas weaknesses and threats were computed as negativity variables. Then, individual indexes were calculated for participants on the respective OCM components. The equation for this calculation was:

OCM index= (Componentpos – Componentneg) / (Componentpos + Componentneg). This calculation gave an estimate of whether the participant was mostly negative or positive regarding the component in a -1to1 format. Then individual scores were standardized by adding 1, and multiplied by 1.5. As such standardized scores ranged from 0 to 3.

The individual score on each component were aggregated, so that sample means for the 17 OCM components could be estimated.

Standardization of survey data. The response format of the OCM ranged from 1 to 4. Values from 1 to 2 were considered as negative, 3 to 4 as positive. The standardized interview data ranged from 0 to 3, therefore survey means on the OCM components were subtracted by 1, giving values from 0 to 3 on both survey and interview data. The individual scores of standardized survey scores were also aggregated so that a comparable sample to the interview data mean was established.

Comparison of Measures

Relationship between open-ended interview and closed ended survey measures were tested through Pearson's product moment coefficients. To compare the measures aggregated sample means, paired t-tests corrected by Bonferroni were executed.

Ethical Considerations

Informed consent. Participants were informed of the procedure, their rights of confidentiality and the possibility to withdraw from the project in a letter acquired prior to the interview (see Appendix D).

Pairing survey and interview. The interviewer paired interviews and surveys, as he already knew the identity of the participants. The pairing list is numerical and do not provide any information regarding participants identity.

REK and NSD. The retrieved data are not found in an official register, nor does it concern critical information regarding informants or their environment. Hence, REK and NSD were not affiliated.

Results

Descriptive Statistics

Table 1 presents means of OCM aggregates from the 33 returned surveys (N=33). The displayed means are not transformed, i.e. made compatible with indexes by subtracting 1. In the OCM response format 1–2 are regarded as negative values, whereas 3-4 as positive values. The OCM means expressed whether the participants agreed or disagreed with the posed statements in the survey concerning the presence of the respective components in their district. Formalization (3.29) was the highest mean. The lowest mean was Efficiency (2.04). Accordingly, the displayed means express rather neutral values as means range between (2.04) and (3.29).

Table 1

Means of OCM Survey Components

Factor	Mean	St.Dev	N
Autonomy	2,87	0,37	33
Integration	2,84	0,47	33
Involvement	2,68	0,34	33
Support	2,91	0,32	33
Training	2,59	0,55	33
Welfare	3,04	0,56	33
Formalization	3,29	0,46	33
Tradition	2,61	0,73	33
Innovation	2,36	0,55	33
Outward	2,80	0,64	33
Reflexivity	2,62	0,41	33
Clarity	2,57	0,58	33
Efficiency	2,04	0,52	33
Effort	3,02	0,42	33
Performance	2,49	0,49	33
Pressure	2,39	0,43	33
Quality	2,95	0,47	33

Table 2 shows the distribution of OCM statements in IGLO and SWOT categories. The total (7925) display the coded amount of statements from the 33 interviews before

Individual, group and leader statements were removed. Residuals regard statements that matched SWOT and IGLO, but not the OCM. Counts on the “Organization” row explain number of statements that were used in the comparison of survey and interview data. The “organization” category covered the most statements (4278) of the IGLO indicating that the organization level was often referred to by interviewees. The (322) residuals were SWOT statements on “organization” that did not match any OCM components and were therefore removed from the comparison.

Table 2

Distributed OCM Statements on IGLO and SWOT

	Individual	Group	Leader	Organization	Total
Strength	431	533	482	1518	
Weakness	408	178	503	1294	
Opportunity	188	139	287	844	
Threat	103	27	82	300	
Residual	130	48	108	322	
Total	1260	925	1462	4278	7925

Table 3 shows the prevalence of OCM components on the Organizational level of IGLO statements distributed on the SWOT categories. The 33 participants (N=33) shared 3956 statements that accounted for the Organization level, SWOT and OCM codes. The most frequent category of statements was Strengths (1518). Weaknesses accounted for second most (1294) of the statements, opportunities (844) and the least statements were coded as threats (300). “Integration” represents the component from the OCM that was most frequently discussed during interviews with a total of 698 statements. The second most discussed component was “Training”, counting 483 statements. The least discussed component was “Autonomy”, by 58 statements in total. The rationale for mapping statements onto research models was to test the OCM components relevance in the specific context. Consequently, frequent themes as integration and training should apply as important for interviewees when describing how they perceive the organizing of investigation. In general, the components of OCM seem relevant for the shared reflections of the informants as each factor were addressed. Further, the descriptive indicated a favour in positive attitudes regarding amount of statements. These results showed that there were more strengths than weaknesses coded on

the OCM dimensions except in the “Quality of procedures”, “Efficiency”, “Reflexivity”, “Tradition” an “Welfare” factors, In the future perspective i.e. opportunities and weaknesses, the pattern between positive and negative attitudes were more distinct, the only factors where negativity accounted for more statements were “Tradition” and “Pressure to produce”.

Table 3

Frequency of SWOT and Organization Level Coded Statements on the OCM Components

Factor	Strengths	Weaknesses	Opportunities	Threats	Total
Autonomy	28	21	8	1	58
Integration	335	213	123	27	698
Involvement	50	33	13	5	101
Support	44	34	21	6	105
Training	185	138	130	30	483
Welfare	35	82	41	39	197
Formalization	113	60	29	10	212
Tradition	49	83	27	32	191
Innovation	49	27	72	9	157
Outward	51	22	15	9	97
Reflexivity	85	108	151	39	383
Clarity	61	42	30	3	136
Efficiency	86	89	42	5	222
Effort	128	58	17	15	218
Performance	108	74	39	25	246
Pressure	23	76	14	24	137
Quality	88	134	72	21	315
Total	1518	1294	844	300	3956

Table 4 displays paired samples statistics of transformed and standardized variables. The statistics represent an aggregate of the total sample. Values ranged between 0-3. Mean values beneath 1.0 were negative, whereas scores above 2.0 were considered positive. Values in the midst 1.5 were neutral.

Interview and survey means appear fairly similar in 12 out of the 17 components. However, mean difference in the “Welfare”, “Innovation” “Efficiency”, “Pressure to produce”, and “Quality in procedures” dimensions spanned above 0,58. Hence, scores may

have indicated divergence in measures on these dimensions. The low values in interview means, as in the “Welfare” component (0,98), showed that participants statements regarding the concept of the component has been found more negative than positive. The welfare component from survey measures (2.00) may indicate that participants were more prone to respond positive on the factor in the survey, than when they reflected on it in interviews. The “Quality in procedures” component coincide with this tendency, whereas the survey mean (1.96) is regarded positive, whilst the interview mean (1.38) appear negative. Conversely, the “Innovation” component displayed a positive mean (2.32) from the interviews, and a more neutral mean (1.44) in the surveys. The “Efficiency” component mean also appeared more positive in interviews (1.73) than in surveys (1.05). This could imply that the measures were not consistently producing results in favour of a specific attitude, or that the survey in general yielded more positive or negative results than the interview or opposite. Additionally, the distinguishing means may have indicated that the two approaches generated different results. An important aspect of this table (N) accounted for number of participants that talked about the specific component in interviews and therefore were eligible for comparisons with survey results.” “Autonomy” was the least discussed topic (N=18) whereas, “integration”, and “training” was mentioned by every participant (N=33).

In general, the standard deviation was larger in components from interview data than in survey data. Thus, results showed a greater spread in the distribution of statements and the rated evaluation of these than the participants’ self-reports from surveys. Mapped statements on the OCM components from the interviews differed more between participants than the survey response regarding positive/negative.

Table 4

Paired Samples Statistics of Standardized Interview and Survey Mean

Factor	Interview		Survey		N
	Mean	Std. Dev.	Mean	Std. Dev.	
Autonomy	1,96	1,11	1,91	0,40	18
Integration	1,97	0,66	1,84	0,47	33
Involvement	1,79	1,04	1,95	0,50	28
Support	2,07	1,05	1,86	0,35	23
Training	1,91	0,60	1,59	0,55	33
Welfare	0,98	0,95	2,00	0,57	29
Formalization	1,78	0,93	2,30	0,46	32
Tradition	0,96	0,87	1,34	0,70	20
Innovation	2,32	0,80	1,44	0,54	25
Outward	1,98	0,98	1,80	0,64	24
Reflexivity	1,75	0,83	1,63	0,42	32
Clarity	2,05	0,95	1,58	0,57	26
Efficiency	1,73	0,98	1,05	0,54	31
Effort	1,99	0,93	2,03	0,42	32
Performance	1,72	0,98	1,44	0,48	29
Pressure	0,72	0,78	1,44	0,39	29
Quality	1,38	0,75	1,96	0,48	30

Inferential Statistics

Table 5 present correlation tests on the two standardized measurements variables. The Pearson product moment coefficients were executed in order to test whether the two measurements were covariant. The correlations that appeared were weak. The strongest correlation was the “Performance” component that displayed a correlation of (.360). None of the measurement comparisons on the respective factors were significant ($p < .05$). This indicated a difference between the measures on individual level or individual differences.

Table 5

Pearson R Correlation Test Between Standardized Interview and Survey Variables

Factor	Correlation	Sig.	N
Autonomy	-,015	,954	18
Integration	-,050	,781	33
Involvement	,014	,942	28
Support	-,311	,148	23
Training	,091	,615	33
Welfare	,195	,310	29
Formalization	-,136	,459	32
Tradition	-,326	,161	20
Innovation	-,178	,396	25
Outward Focus	-,115	,594	24
Reflexivity	,175	,339	32
Clarity of Org. Goals	,111	,591	26
Effort	-,226	,213	32
Efficiency	-,224	,225	31
Performance Feedback	,360	,055	29
Pressure	-,254	,184	29
Quality Procedures	,156	,410	30

To examine whether the differences between the means of the OCM components measured by survey and interviews were significantly different, we executed paired t-tests. Table 6 presents paired sampled t-tests of the two measures on the 17 OCM components. The significance level was corrected by Bonferroni, according to the number of compared pairs, as number of comparisons increases the chance of finding significant differences. The column “means” express the difference in means between the measurements on the respective factors. The least different component means were: “Autonomy” (0.05), “Effort” (-0.05), “Integration” (0.13) and “Reflexivity” (0.13). Accompanied by nine other scales these means were not significantly different. Statistical significant differences occurred on the 4 subsequent factors: “Welfare”, “Innovation”, “Pressure to produce” and “Quality procedures”. These measured means were the most divergent of the 17 OCM components distinct by (0.58) or more. By one exception (Efficiency -0.68) these findings support the descriptive statistics from table 3, as the means with greatest range were statistical different in

the paired sampled t-test. The estimation of means indicated whether measures generated equal results or not. As means of 13 scales matched, one may argue that the two measures caused similar results and converged and that the four statistically different means diverged.

Table 6

Paired Samples T-Test of Standardized Interview and Survey Indexes on OCM Components

Factor	Mean	Std. Dev	t	Df	Sig.
Autonomy	0,05	1,18	0,17	17	,866
Integration	0,13	0,83	0,87	32	,389
Involvement	-0,16	1,15	-0,73	27	,474
Support	0,21	1,21	0,82	22	,423
Training	0,32	0,78	2,32	32	,027
Welfare	-1,02	1,01	-5,43	28	,000**
Formalization	-0,51	1,09	-2,65	31	,012
Tradition	-0,38	1,28	-1,31	19	,205
Innovation	0,88	1,04	4,26	24	,000**
Focus	0,18	1,23	0,73	23	,472
Reflexivity	0,13	0,86	0,82	31	,418
Clarity	0,46	1,06	2,23	25	,035
Efficiency	-0,68	1,22	-3,10	30	,004
Performance	0,28	0,92	1,65	28	,110
Effort	-0,05	1,10	-0,23	31	,819
Pressure	-0,72	0,96	-4,07	28	,000**
Quality	-0,58	0,87	-3,82	29	,001**

Bonferroni corrected $0.05/17=0.0029\approx 0.003$.

**Significant at 0.001

Discussion

The aim of this study was to statistically investigate the use of mixed methods by examining open-ended interviews and closed-ended survey instruments, and their ability to generate convergent results on the same sample. This was done by applying a general measurement of work, the OCM, on the specific context of police investigation. The research questions was, are open-ended interviews and surveys capable of generating convergent results?

Summing Up Results

The distribution of interview statements on the OCM, SWOT and Organization level of IGLO show that the general measure is prevalent to the interviewees as each component are reflected on in interviews. As components are prevalent in interviews, the OCM account for participants' experiences of investigation. As such the measure is applicable to investigation, but also comparable between data-collections. Prevalent components of a survey do not prove the interview and survey to reflect convergent opinions of the participants. However, the paired sampled t-tests show that the two measures mean match in 13 comparisons, whereas four out of the 17 OCM components were statistically different. This indicates that the top down coding of interviewees free reflections converge with participants ratings of the OCM survey on 13 components and that four are diverging. In regards to individual values, Pearson correlation analyses presented weak insignificant correlations as predicted.

General Discussion

The research question addressed whether results of open-ended interviews and surveys could produce convergent results. Results of empirical analyses showed that the two measurements might generate similar results. The derived results from the paired t-tests support this statement, as the two measurements converge in 13 occasions. This is in alignment with existing research on mixed methods between open-ended and closed-ended data collections (Jex et al., 1997; Mazzola et al., 2011). Furthermore, this study empirically shows that interviews and surveys data collections are comparable, in opposition to theories that claim these methods to be incompatible (Bednarz, 1985; Forshaw 2007; Howe, 1988; Ogborne, 1995; Simpson and Eaves, 1985).

It was suggested that former research might have found convergent results between open-ended and closed-ended, as open-ended measures were thematically cued on work components (Jex et al., 1997; Mazzola et al., 2011) an. By posing interview questions without cues of work components, and measure interviews by subsequent top down coding as

suggested by Jex and Spector (1991) results still converged. This supports earlier studies conclusions, as thematic cues were probably not the reason of convergence between the data collections.

Furthermore, in this study interviews were conducted before the surveys and separated in time. These factors were assumed to influence the results as prior studies conducted open-ended and closed-ended questions at the same time of assessment, exposing participants of the same themes twice (Jex et al., 1997; Mazzola et al., 2011). Therefore, previous studies may have primed the participants. In addition, survey fatigue could have impacted results of the open-ended questions, which were answered after the closed-ended questions. However, the current design of mixed-methods prevents the priming effect, and show that open-ended and closed ended measures may still converge. Most importantly this indicates that open-ended interviews that are not cued on work components may be a viable framework for measurement of work. This is in consistence with prior studies, which concluded that open-ended data collections were a viable measure of work (Jex et al., 1997; Mazzola et al., 2011).

The responses from the OCM survey display an agreement with the coded interviews on thirteen components. This in consistence with the notion of Neuendorf (2002) that content analysis is probable to match any survey research. Participants' free associations of investigation were coded in SWOT, IGLO and OCM codes. Participants were primed by the SWOT interview framework, but not IGLO and OCM. Unaware of the subsequent coding process interviewees reflected upon characteristics of OCM components. This enabled that opinions could be assessed by trained coders who categorically placed statements in regards to the relevance of OCM components. The respective statements of interviews are interviewees expressed perceptions of the organizing of investigation, their terminology may not in words match the operationalization of the OCM codes, but still refer to the concept of the OCM components. As such their answers are not restricted or biased by a response format. If participants say: "We struggle with quality of investigation" they conceptually refer to "quality of procedures". A statement as this would be coded as a weakness i.e. a negative statement concerning the component. The interview and interpretation process retrieve information based on free reflections and association in opposition to the closed ended survey that collect data within the limits of the item composition. Interview values may transpire through expressions that surveys are not capable to capture. In addition, answers from surveys may reflect what the participant thinks is appropriate to answer based on the wording of items. This may occur in interviews as well, however the amount of cues are limited and the interviewee are encouraged to associate on the aspects he or she finds important.

Four of the components diverge on the paired t-test between measures. This may indicate that interviews and surveys reflect different opinions regarding the components. This divergence may in correspondence with Lund (2011) increase the theoretical insight in measures. As both data collections concern self-reports, they are predisposed to common method variance (Howell, 2013) or, as Spector (2006) argues, reveal specific biases. The two measures do not appear to consistently generate values in favour of a positive or negative value on the 17 components. Yet, three of four divergent means display that components are reported more negative in interviews than in surveys. This may indicate that opinions reflected in the interviews are more negative in regards to the component, than what participants rate in surveys, prior studies have argued negative affect as a bias in self report data (Brief, Burke, George, Robinson, & Webster, 1988; Chen & Spector, 1991; Jex et al., 1997). If negative affect influences the results, the interview data in this study seem more prone to this bias. However, there might be other biases that cause the divergence and these might be explainable as suggested by Spector (2006).

As three survey components are reported significantly more positively than interview scores on the “Welfare”, “Pressure to produce” and “Quality of procedures” components, a tendency to answer more positive on the closed-ended instrument is evident. A well-known bias within the common method variance literature refers to items that are written in such a way that they reflect more socially desirable attitudes, behaviours or perceptions, namely social desirability (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). More than 50 percent of the sample consists of leaders. In the survey instructions participants were asked to rate the district as an organization. As more than half of the participants are responsible of employee welfare, it would seem odd if the leaders answered that they did not take care of employees or tried to be fair. This could explain why the “Welfare” component is rated more positively in the surveys than in the interviews.

Another example of social desirability bias might have occurred on the “Quality in procedures” component. Bearing in mind that the police procedures govern the Norwegian law, and that results of the investigations are used in court (Politiloven, 1995) it could be unfortunate to report that the quality of procedures was absent. This is shown in the results, as the survey score is higher than in interviews. In opposition to the survey score that indicates a presence of quality procedures, interviewees said that the emphasis on quality was abundant, as they lacked competence and time to stress such matters. This corresponds with Spector (1992) who found self-reports of work environment to be influenced by both biases and the environment. If this proves to be correct, the inference of OCM as a general and global

measure (Patterson et al., 2005) may be revised, as variations between organizations may be caused by context specific social desirability. In accordance with Roberts and Glick (1981), the discovery of context specific biases shows that survey measures need to be assessed by multiple methods to avoid that prevalent components do not measure the concept of interest or reflect wrong answers.

The divergence on the “pressure to produce” component might indicate other elements than biased instruments. The Norwegian police is regulated by the Ministry of Justice, and the demands toward solving or investigating reported offences of the public emphasize that a large quantity of cases are investigated. The districts are continually evaluated and the performance is measured based on the total number of cases that are reported, processed and solved (Riksrevisjonen, 2012). However, this example illustrates that divergent means provide valuable insight in the distinction of information from measures. The pressure mentioned in interviews was rarely referred to as positive. The participants said they were too few employees, and that there were too many cases to handle. The divergence on this component illustrates that the participants agreed on a presence of pressure to produce in their districts. However, the survey does not show how participants feel about the presence of the reported pressure. As the interview mean presents a negative score, it appears that the participants refer to this pressure as negative. This is in correspondence with Schonfeld and Farrel (2010) who argue that qualitative data enrich surveys. Furthermore, the divergence illustrates the benefit of mixed methods, as the insight into participants' experiences increases.

“Innovation and Flexibility” is the only component that displays a more positive value in interviews than in surveys. A plausible reason is the coding schema and the subsequent coding. The description of this code is: “Flexibility- An orientation toward change; Innovation – The extent of encouragement and support for new ideas and innovative approaches”. Both coders perceived this code as being positive towards change. The “Tradition” code, on the other hand, is described as: “The extent to which established ways of doing things are valued”, i.e. the code is negative towards change. Consequently, negative statements regarding change were coded as “Tradition”, whereas positive statements concerning change were coded as “Innovation and Flexibility”. This explanation is supported by the negative interview mean on “Tradition” (0.96), and the positive “Innovation & Flexibility” (2.32) score. This example illustrates that subjective interpretation affects the comparisons between qualitative and quantitative data, and that the use of mixed-methods depend on researchers interpretation and familiarity with the data. This also addresses the framework of the OCM as both components may measure change. The survey scores of “Tradition” (1.34) and

“Innovation and Flexibility” (1.44) are neutral, or slightly negative, and may support the suggestion that the two components measure the same concept, as values are somewhat similar in the surveys.

These notions of divergence show how the mixed methods measurement of OCM components in interviews and surveys are advantageous regarding their ability to reflect participants’ experiences. The literature of mixed-methods often refers to its capacity of validating studies. Critics of open-ended methods frequently question the reliability of interpreted data. The subjective ratings of statements are addressed in content analysis through inter-rater reliability. However, existing literature is not in agreement of criteria for sufficient inter-rater reliability or how to measure it (Neuendorf, 2002). Krippendorff (2004) argues that a reliable process ensures that valid inferences may be made, and suggest a rather high consensual agreement of raters, measured by inter-rater reliability as a criterion of valid inferences. Other theorists claim that through systematic approaches and craftsmanship, you can validate qualitative data (Kvale & Brinkmann, 2009). The author and fellow students in this study received extensive training on coding, collaborated in design of codes, and held consensus meetings to elaborate the shared understanding of codes. Yet, the inter-rater reliability of OCM was only 0.41 on Krippendorff’s A, indicating a low reliability score. Conversely, SWOT codes were as high as 0.71. This could indicate that the OCM codes are too broad to become reliable, as some statements could refer to several OCM components. Another alternative may be that the assessment of inter-rater reliability in this study does not account for the actual reliability. As measures converge in 13 occasions and the divergence is explainable, a more substantial explanation may be that the measurement of inter-rater, which measure agreement between coders on a certain amount of statements, is not a proper criterion of valid inferences. This is inconsistent with Krippendorff (2004), but Zeller and Carmine (1980) argue that measurements may be valid without being reliable, and reliable without being valid. However, the low inter-rater could serve as an explanation of the divergent measures, or more importantly indicate that components converge by chance. Yet, the consensus between coders’ interpretation of interviews and participant ratings of surveys should be a more accurate estimate of valid inferences than the agreement between coders. As such the convergence between measures can indicate that the open-ended interview data collection and interpretation in this study is valid.

Prior studies have validated measures based on high correlations between measurements (Jex et al., 1997; Mazzola et al., 2011). As such this study’s low correlations could indicate that measures diverged. However, studies have shown that employee's

perceptions and experiences of work constructs differ according to occupation (Mazzola et al., 2011; Narayanan et al., 1999). The sample in this study encompasses three different occupations within the police. These professions vary in responsibility and work tasks concerning investigation. Even though the participants refer to the same process or product, it is reasonable to expect that an investigator experience the investigation process and the organization differently than a police commissioner. Accordingly, it was assumed that individual values on components would differ with regards to occupation, and therefore cause weak correlations. An agreement on the perception of the components could display high correlations, as participants' values would be somewhat equal. Howell (2013) points out that weak correlations indicate inconsistent scattering of participants' values around the mean. Patterson et al. (2005) claim that OCM components scores within organizations are similar, but discriminate between different organizations. In the interviews participants were asked to reflect on their district, and in the survey instructions participants were encouraged to rate their districts. Therefore each district was considered an organization and discriminance was probable. As correlations consider the agreement of individual values, this test was predicted to be an inaccurate analysis of convergence between measures. Rather, a weak correlation could indicate that investigators, chief officers and the commissioners perceive work regarding investigation differently.

An additional cause of low correlations could be that participants express other opinions in interviews than in surveys. As such, another estimate of convergence was needed. The paired t-tests cover the samples value mean between measures. As occupations in the police probably affect the small sample of 33 participants, it was predicted that the paired t-tests were more accurate measures, as it accounts for the total variation and the deviates of values on both measures. If the interviews consistently generated other values than the survey, it is probable that means would diverge to a greater extent than the four significantly different component measures in this study.

Theoretical Implications

This study elaborates on Mazzola et al. (2011) and Jex et al. (1997), and shows that open-ended interviews and closed-ended surveys of data collections should be used in conjunction to complement the respective strengths and weaknesses of each method.

Matched pairs of measures may serve as validation of qualitative material, as well as the composition of survey components. The procedure of mixing methods in this study offers one such tool for validating qualitative material. In opposition to Krippendorff (2004), valid

inferences do not seem to depend on high reliability. Rather, in consistence with Kvale and Brinkmann (2009), the different steps of qualitative collection and interpretation were systematic and thorough, and thus an example of empirically proven craftsmanship.

The convergence on thirteen components indicates that open-ended interviews are a viable framework for measurement of work. In consistence with Lund (2011) mixed methods increase the theoretical insight, this study's use of standardized data and paired t-tests obtain insight on possible biases, meanings of data collections and analyses.

The convergence of open-ended interviews and surveys illustrate that methods of data collections do not necessarily belong to a philosophical tradition like positivism or constructivism. For instance, open-ended interviews may be quantified, standardized and analysed statistically on the same basis as closed-ended measures. This is in opposition to the thesis of incompatible methods (Bednarz, 1985; Forshaw 2007; Howe, 1988; Ogborne, 1995; Simpson and Eaves, 1985). Rather, the assumptions of Neuendorf (2002), who claim that content analysis match the requirements and results of surveys, are concurrent with the empirical results of convergent measures from this study.

The mixed-method comparison supports van Veldhoven (2005) in the argument of using general instruments on specific domains, as the OCM components are seemingly prevalent to the specific context under study. The convergent and divergent components indicate how to specify general measures to the context. By using mixed-methods you retrieve information on how to improve and interpret surveys. In consistence with Spector (1992), this study implies that survey instruments may be prone to context specific biases according to that which is socially desirable. Therefore, mixed-methods serve as an opportunity to discover context specific influences on general measures as the OCM.

Practical Implications

Organizational analysts may use coding of interviews as a measure of standardized scores on abstract components. This elicits a huge variety of possibilities within selection and development, and organizational learning among others as tests of people's personality and values might be conducted through interviews.

Appraisal interviews could be coded top down by analysts and give important feedback to organizations. This would increase the standards of such interviews and benefit employers and employees as the information could be substantial to the work environment and possible outcomes.

Organizations that wish to apply a survey instrument on their context may use interviews as empirical support for which work model to use, and specify the instrument according to top down coding of interviews. Participants can fill out the original survey measure after analysts have found the most relevant instrument, and then they can assess the instruments ability to reflect participants opinions by comparison of standardized results.

For the specific context, this study implies that police investigation may be measured through qualitative SWOT interviews. This elicits a variety of possible interpretations that can generate information and knowledge on the understudied investigation domain.

Further, the 17 OCM components account for a large variety of statements from interviews, implying that the OCM survey is applicable for measurement of investigation in the Norwegian police. Furthermore, the least prevalent components may be removed when measuring investigation, instead adding outcome components of the respective components. Some considerations should be done when interpreting the results of the “welfare” and “quality in procedures” components, as they seem apt to biases.

Limitations

There are four main limitations of this study. The first limitation concerns the quantification and subsequent statistical proceedings of open-ended data material. The assumption that these categories are distinct, and thereby may be coded, aggregated, and used in descriptive and inferential statistics is of course experimental. The analysed data in this study consists of categories on a nominal level. The Pearson correlation requires categories on the interval level to be a precise measure of linear dependency (Field, 2009). As nominal categories are not intended for Pearson's R, a linear independency might be a natural effect of an inappropriate test. Further, the positivity and negativity variables consist of values in two different time perspectives, the present (Strength & Weaknesses) and the future (Opportunities & Threats). The OCM survey is composed to measure the presence. If the future values of informants were removed, results may have differed. Additionally, as the OCM measures the organizational level, statements that referred to the individual, group and leadership level were removed from the comparison. Surely, this limits the information and nuances of qualitative data, as important meanings get lost in the methodical proceeding. However, these data are still found in the PASW 18 file and the transcriptions of interviews, which make further investigation in residuals of the OCM possible.

A second limitation is the low inter-rater reliability. Even if Neuendorf (2002) argues that a universal standard of content analysis coding reliability requirements are not present in

the literature a Krippendorff's A of 0.41 on the OCM is low. As mentioned, this may cause convergence by chance and the qualitative data to reflect wrong work components. In addition unitizing reliability was not properly assessed. The reliability of this procedure might be critical. By unitizing you risk fragmenting participant's opinions as the structure of language and thoughts may be communicated in larger units than the statements in this study. Krippendorff (2004) argue that too small statements may cause lack of reliability. Indeed, the statements of this approach are small, as some content analytics consider an interview as a unit. The approach of categorical distinctions must be emphasized, as psychological phenomenon may be frequent and complex when people talk about perceptions of organized work. It is reasonable to assume that interviewees reflect on several psychological components, and that larger units would neglect some of the concepts.

The third limitation concerns the sample. The original sample consisted of 51 strategically selected participants, which reduce the possibility of generalization. The sample that was statistically analyzed comprised those 33 participants that responded on the survey. Thus, the sample may have been response biased. In addition, the sample size affect the significance of correlations as the 33 participants may be too few for sufficient statistical power, large samples increase the chance of statistical significance (Howell, 2012). The sample size is not representative for general statistical inferences of investigation due to a small sample. However, the inferences of this study concern surveys and interviews ability to reflect agreement, not a general understanding of the organizational climate in police investigation

The fourth and last limitation is related to the collection of qualitative data. The restrictive rules of follow-up questions may have been violated, as the interviewers own experience and interests occasionally may have affected probing questions. Some follow-up questions were thematically cued regarding education and formal transfer of education. This could have skewed the data and thereby generated a larger distribution on the OCM component "Training". The theme was brought up unmotivated by informants, so the components relevance is present but the frequency of the components distribution may have been biased. Paraphrasing is an ordinary technique in PEACE structured SWOT interviews, but precautions must be made. Consequences of thematic follow up questions are that they increase the studied research models relevance, according to the assumption that interview referrals indicate the prevalence of components for the specific context.

Future Studies

This study did not perform a “bottom-up” or inductive analysis of the dataset or residuals. Such a procedure might discover important components of investigation beyond the categories of the OCM. If the original statements of this dataset were kept, an inductive coding of every statement could later be compared with the coding of this study. This would show codes that are misinterpreted, or nuance the conceptual understanding of components. Additionally, one could establish a specific instrument for measurement of investigation of emergent components from interviews similar to the Schneider et al. (1992) “Passion for service” measurement.

Top-down coding of the 82 OCM items is extensive and difficult, however this could assess the most relevant questions of a component, and generate data that are more specific in the comparison of survey and interview data. Statements match on items could be rated in positive or negative, as such acquiring interview data that are not affected by the two time-perspectives of SWOT.

Instead of coding on the SWOT framework and transforming SWOT values on components to variables in positive and negative, one may assess statements of interviews that have been matched in the OCM components and use the four categories from the OCM response format as codes. This could create comparable data that do not depend on transformation.

Conclusion

This study advocates the use of open-ended interview and closed-ended survey components in conjunction. The main finding is that open-ended interviews can produce similar results as surveys, as the t-tests showed convergence in 13 of 17 OCM components. Furthermore, the study found that the use of standardized open-ended data compared to surveys, could provide research with a tool for validation. The study also shows that open-ended SWOT interviews could be a viable framework for the measurement of work. Based on the use of mixed-methods, the OCM components seem to be prevalent to investigation and applicable for the measurement work characteristics in investigation.

This study's weak correlations concur with findings of earlier research, indicating that occupation affects participants expressed experiences. The four divergent means provided valuable insight into both measures.” Welfare” and “Quality in procedures” are prevalent components in investigation. However, a tendency of social desirability might have been discovered among participants in the closed-ended OCM components, or an example of

negative affect in the interviews. Additionally, divergence can indicate fallible coding of open-ended data, or conceptual inconsistencies in the investigated measurements framework. Based on information from divergence one may specify the general OCM measure.

This paper demonstrates the advantageous complementary abilities of mixed methods in the measurement of specific work contexts. Consequently, if organizations wish to construct a situation-specific measure, mixed-methods can play a crucial role in the development and validation of such a measure. In sum, mono-methods have their respective weaknesses and strengths, while the use of mixed methods overlap the weaknesses of different data collections.

References

- Bakker, A. B., & Demerouti, E. (2007). The Job Demands-Resources model: state of the art. *Journal of Managerial Psychology*, 22(3), 309-328.
- Bednarz, D. (1985). Quantity and quality in evaluation research: A divergent view. *Evaluation and Program Planning*, 8(4), 289.
- Berelson, B. (1952). *Content analysis in communication research*: New York, NY, US: Free Press.
- Berger, P. L., & Luckmann, T. (1967). *The social construction of reality: a treatise in the sociology of knowledge*. Garden City, N. Y.: Doubleday.
- Bernstrøm, V. (2009). *Investigating the Organizational Climate Measure's Generalizability*. (Master Work and Organizational Psychology), University of Oslo. Oslo.
- Brief, A. P., Burke, M. J., George, J. M., Robinson, B. S., & Webster, J. (1988). Should negative affectivity remain an unmeasured variable in the study of job stress? *Journal of applied psychology*, 73(2), 193-198. doi: 10.1037/0021-9010.73.2.193
- Bryman, A. (1984). The Debate about Quantitative and Qualitative Research: A Question of Method or Epistemology? *The British Journal of Sociology*, 35(1), 75-92.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56(2), 81-105. doi: 10.1037/h0046016
- Chen, P. Y., & Spector, P. E. (1991). Negative affectivity as the underlying cause of correlations between stressors and strains. *Journal of applied psychology*, 76(3), 398-407. doi: 10.1037/0021-9010.76.3.398
- Clarke, C., & Milne, R. (2001). *A National Evaluation of the PEACE Investigative Interviewing Course*. London: Home Office
- Cortina, J., Schmitt, N., & Whitney, D. J. (1992). The linkage between objective and subjective assessments of job dimensions. *Paper presented at the 52 annual meeting of the Academy of management. August 1992, Las Vegas.*
- Creswell, J. W., Plano Clark, V.L., Gutmann, M.L., & Hanson, W.E. (2003). Advanced mixed methods and research design. In A. Tashakkori, & Teddlie, C., (Ed.), *Handbook of mixed methods in social and behavioral research* (pp. 209-240). London: Sage.
- Cronbach, L. J. (1946). Response sets and test validity. *Educational and psychological measurement*, 6(4), 475.
- Field, A. P. (2009). *Discovering statistics using SPSS*. London: SAGE publications Ltd.

- Forshaw, M. J. (2007). Free qualitative research from the shackles of method. *The Psychologist*, 20(8), 478-479.
- Freelon, D. (2010). ReCal: Intercoder Reliability Calculation as a Web Service. *International Journal of Internet Science*. 5 (1), 20–33.
- Glomseth, R., Gottschalk, P., & Solli-Sæther, H. (2007). Occupational culture as determinant of knowledge sharing and performance in police investigations. *International journal of the sociology of law*, 35(2), 96.
- Grant, A. M., & Parker, S. K. (2009). 7 Redesigning Work Design Theories: The Rise of Relational and Proactive Perspectives. *The Academy of Management Annals*, 3(1), 317-375. doi: 10.1080/19416520903047327
- Hackman, J. R., & Oldham, G. (1976). Motivation through the design of work: Test of a theory. *Organizational Behavior and Human Performance*, 16(2), 250.
- Hanson, W. E. (2005). Mixed Methods Research Designs in Counseling Psychology. *Journal of counseling psychology*, 52(2), 224.
- Helms, M. M., & Nixon, J. (2010). Exploring SWOT analysis – where are we now?: A review of academic research from the last decade. *Journal of Strategy and Management*, 3(3), 215-251.
- Hoff, T., Flakke, E., Larsen, A.K., Lone, J.A., Bjørkli, C.A., Bjørklund, R. (2009). On the Validity of M-SWOT for Innovation Climate Development. *Scandinavian Journal of Organizational Psychology*.
- Hoff, T., Straumsheim, P., Bjørkli, C. A., & Bjørklund, R. A. (2009). An External Validation of Two Psychosocial Work Environment Surveys—A SWOT Approach. *Scandinavian Journal of Organizational Psychology*, 1(1).
- Home Office. (2005). Guidance on Statutory Performance Indicators for Policing 2005/2006. Police Standards Unit. Home office of the UK Government. (www.policereform.gov.uk)
- Howe, K. R. (1988). Against the Quantitative-Qualitative Incompatibility Thesis or Dogmas Die Hard. *Educational Researcher*, 17(8), 10-16. doi: 10.3102/0013189x017008010
- Howell, D. C. (2013). *Statistical methods for psychology*. Belmont, Calif.: Wadsworth Cengage Learning.
- Hønsen, L. (2010). *Arbeidsmiljø i en kunnskapsintensiv organisasjon*. (Master Work and Organizational Psychology), University of Oslo, Oslo.

- Jex, S. M., Adams, G. A., Elacqua, T. C., & Lux, D. J. (1997). A comparison of incident-based and scale measures of work stressors. *Work & Stress, 11*(3), 229-238. doi: 10.1080/02678379708256837
- Johnson, R. B. (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher, 33*(7), 14.
- Karasek, R. A., Jr. (1979). Job Demands, Job Decision Latitude, and Mental Strain: Implications for Job Redesign. *Administrative Science Quarterly, 24*(2), 285-308.
- Keenan, A., & Newton, T. J. (1985). Stressful events, stressors and psychological strains in young professional engineers. *Journal of organizational behavior, 6*(2), 151-156. doi: 10.1002/job.4030060206
- Kidd, P., Scharf, T., & Veazie, M. (1996). Linking Stress and Injury in the Farming Environment: A Secondary Analysis of Qualitative Data. *Health Education & Behavior, 23*(2), 224-237. doi: 10.1177/109019819602300207
- Krippendorff, K. (2004). *Content analysis: an introduction to its methodology*. Thousand Oaks, Calif.: Sage.
- Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
- Kvale, S., & Brinkmann, S. (2009). *Interviews: learning the craft of qualitative research interviewing*. Los Angeles, Calif.: Sage.
- Luen, T. W., & Al-Hawamdeh, S. (2001). Knowledge management in the public sector: principles and practices in police work. *Journal of information science, 27*(5), 311-318. doi: 10.1177/016555150102700502
- Lund, T. (2005). The Qualitative–Quantitative Distinction: Some comments. *Scandinavian Journal of Educational Research, 49*(2), 115-132. doi: 10.1080/00313830500048790
- Lund, T. (2011). Combining Qualitative and Quantitative Approaches: Some Arguments for Mixed Methods Research. *Scandinavian Journal of Educational Research, 56*(2), 155-165. doi: 10.1080/00313831.2011.568674
- Mazzola, J. J., Walker, E.J., Shockley, K.M., Spector, P.E. (2011). Examining Stress in Graduate Assistants. *Journal of mixed methods research, 5*(3), 198-2011.
- Mazzola, J. J., Schonfeld, I. S., & Spector, P. E. (2011). What qualitative research has taught us about occupational stress. *Stress and Health, 27*(2), 93-110. doi: 10.1002/smi.1386
- Narayanan, L., Menon, S., & Spector, P. E. (1999). Stress in the workplace: A comparison of gender and occupations. *Journal of organizational behavior, 20*(1), 63-73. doi: 10.1002/(sici)1099-1379(199901)20:1<63::aid-job873>3.0.co;2-j

- Neuendorf, K. A. (2002). *The content analysis guidebook: Kimberly A. Neuendorf*. Thousand Oaks, Calif.: Sage.
- Ogborne, A. C. (1995). Unhelpful divide but important distinction. *Addiction*, 90(6), 755-757. doi: 10.1111/j.1360-0443.1995.tb02216.x
- Parker, S. K., Wall, T. D., & Cordery, J. L. (2001). Future work design research and practice: Towards an elaborated model of work design. *Journal of Occupational and Organizational Psychology*, 74(4), 413-440. doi: 10.1348/096317901167460
- Patterson, M. G., West, M. A., Shackleton, V. J., Dawson, J. F., Lawthom, R., Maitlis, S., . . . Wallace, A. M. (2005). Validating the organizational climate measure: links to managerial practices, productivity and innovation. *Journal of organizational behavior*, 26(4), 379-408. doi: 10.1002/job.312
- Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of applied psychology*, 88(5), 879-903. doi: 10.1037/0021-9010.88.5.879
- Politi-loven. (1995). *Lov om politiet. LOV 1995-08-04 nr 53. (politi-loven)*
- Riksrevisjonen. (2012). *Riksrevisjonens undersøkelse av politiets arbeid med vinningskriminalitet. Dokument 3:6.*
- Roberts, K. H., & Glick, W. (1981). The job characteristics approach to task design: A critical review. *Journal of applied psychology*, 66(2), 193-217. doi: 10.1037/0021-9010.66.2.193
- Schneider, B., Wheeler, J. K., & Cox, J. F. (1992). A passion for service: Using content analysis to explicate service climate themes. *Journal of applied psychology*, 77(5), 705-716. doi: 10.1037/0021-9010.77.5.705
- Schonfeld, I. S., & Farrel, E. (2010). Qualitative methods can enrich quantitative research on occupational stress: An example from one occupational group. *Research in occupational stress and wellbeing series*.
- Simpson, R. G., & Eaves, R. C. (1985). Do we need more qualitative research or more good research? A reaction to Stainback and Stainback. *Exceptional Children*, 51(4), 325-329.
- Sparks, K., & Cooper, C. L. (1999). Occupational differences in the work-strain relationship: Towards the use of situation-specific models. *Journal of Occupational and Organizational Psychology*, 72(2), 219-229. doi: 10.1348/096317999166617

- Spector, P. E. (1992). A consideration of the validity and meaning of self reports of job characteristics. In C. L. Cooper & I. T. Robertson (Eds.), *International Review of Industrial and Organizational Psychology*. Chichester: Wiley.
- Spector, P. E. (2006). Method variance in organizational research: Truth or urban legend? *Organizational Research Methods*, 9(2), 221-232.
- Spector, P. E., & Jex, S. M. (1991). Relations of job characteristics from multiple data sources with employee affect, absence, turnover intentions, and health. *Journal of applied psychology*, 76(1), 46-53. doi: 10.1037/0021-9010.76.1.46
- Stabell, C. B., & Fjeldstad, Ø. D. (1998). Configuring value for competitive advantage: on chains, shops, and networks. *Strategic Management Journal*, 19(5), 413-437. doi: 10.1002/(sici)1097-0266(199805)19:5<413::aid-smj946>3.0.co;2-c
- Taber, T. D., Beehr, T.A., Walsh, J.T. (1985). Relationships between job evaluation ratings and self-ratings of job characteristics. *Organizational behavior and human decision processes*, 35(1), 27.
- van Veldhoven, M., Taris, T. W., de Jonge, J., & Broersen, S. (2005). The relationship between work characteristics and employee health and well-being: How much complexity do we really need? *International Journal of Stress Management*, 12(1), 3.
- Wiggins, B. J. (2011). Confronting the dilemma of mixed methods. *Journal of Theoretical and Philosophical Psychology*, 31(1), 44-60. doi: 10.1037/a0022612
- Zeller, R. A., Carmine, E. G. (1980). *Measurement in the Social Sciences*. London: Cambridge University Press.
- Zharghooni, S. (2011) *Applying models for safety climate on rail traffic management*. (Master Work and Organizational Psychology), University of Oslo, Oslo.

Appendix A - The SWOT Interview Guide

1.

Please tell us what you think works well regarding the investigative work here in the police district - we call this the strength of the investigative work.

2.

Please tell us what you think does not work well regarding the investigative work here in the police district – we call this the weakness of the investigative work.

3.

Please tell us what you consider to be opportunities for improving the investigative work here in the police district – we call this the opportunities in the investigative work.

4.

Please tell us what you consider to be threats against improving the investigative quality here in the police district - we call this the threats in the investigative work.

Appendix B – OCM Questionnaire

Undersøkelse av organisasjonsklimaet i Politiet

Prosjektet «Kvalitet i etterforskningen»

Spørsmål	Helt Feil	Ganske feil	Ganske riktig	Helt riktig
1. Ledelsen lar stort sett ansatte ta sine egne beslutninger				
2. Ledelsen har tillit til at folk kan ta arbeidsrelaterte beslutninger uten å innhente tillatelse først				
3. Ledelsen holder streng kontroll med arbeidet til de ansatte				
4. Ledelsen har for strengt regime over måten ting blir gjort på				
5. Det er viktig å dobbeltsjekke med nærmeste leder før man tar en beslutning				
6. Folk er mistenksomme overfor andre avdelinger				
7. Det er svært lite konflikt mellom avdelingene her				
8. Folk er innstilt på å dele informasjon på tvers av avdelinger				
9. Det er svært effektivt samarbeid mellom avdelingene				
10. Det er lite respekt mellom noen av avdelingene her				
11. Ledelsen lar de ansatte medvirke i beslutninger som angår dem				
12. Endringer blir gjort uten å snakke med de involverte				
13. Folk har ingen innvirkning i avgjørelser som påvirker arbeidet deres				
14. Folk føler at beslutninger ofte tas uten at de blir hørt				
15. Informasjon deles i stor grad				
16. Det er ofte kommunikasjonssvikt her				
17. Overordnede er svært dyktige til å forstå folks problemer				
18. Overordnede viser at de har tiltro til sine ansatte				
19. Overordnede hos oss er vennlige og lette å henvende seg til				
20. Folk kan stole på at overordnede gir god veiledning				

21. Overordnede viser forståelse for sine ansatte				
22. Folk får ikke tilstrekkelig opplæring i nye systemer eller nytt utstyr				
23. Folk får tilstrekkelig opplæring i bruk av nye systemer og utstyr				
24. Organisasjonen gir kun et minimum av den opplæringen folk trenger for å gjøre jobben sin				
25. Folk blir sterkt oppmuntret til å utvikle sine ferdigheter				
26. Denne organisasjonen vier lite oppmerksomhet til ansattes interesser				
27. Denne organisasjonen forsøker å ta vare på sine ansatte				
28. Denne organisasjonen bryr seg om sine ansatte				
29. Denne organisasjonen prøver å handle rettferdig overfor sine ansatte				
30. Hos oss blir det oppfattet som svært viktig å følge reglene				
31. Folk kan ignorere formelle prosedyrer og regler hvis det bidrar til å få jobben gjort				
32. Hos oss må alt gjøres etter reglene				
33. Hos oss er det ikke nødvendig å følge alle prosedyrer til punkt og prikke				
34. Hos oss blir ingen særlig opprørt hvis reglene brytes				
35. Toppledelsen foretrekker å holde seg til de etablerte, tradisjonelle måtene å gjøre ting på				
36. Måten denne organisasjonen gjør ting på har aldri forandret seg særlig mye				
37. Ledelsen er ikke interessert i å prøve ut nye ideer				
38. Hos oss skjer endringer i måten ting gjøres på svært langsomt				
39. Hos oss blir nye ideer gjerne akseptert				
40. Organisasjonen reagerer raskt når endringer er nødvendig				
41. Behov for å gjøre ting annerledes fanges raskt opp av ledelsen				
42. Denne organisasjonen er svært fleksibel; den kan raskt endre prosedyrer for å møte nye vilkår, og problemer løses når de oppstår				
43. Støtte til utvikling av nye ideer er lett tilgjengelig				

44. Folk i denne organisasjonen er alltid ute etter å se problemer fra nye vinkler				
45. Denne organisasjonen er ganske innadrettet; man bryr seg ikke om hva som skjer i samfunnet				
46. Det legges ikke mye vekt på måter å bedre service til publikum				
47. Publikum sitt behov er ikke ansett som topp prioritet hos oss				
48. Denne organisasjonen er treg til å reagere på publikums behov				
49. Denne organisasjonen ser stadig etter nye muligheter i samfunnet				
50. Måten de ansatte jobber sammen på i denne organisasjonen endres gjerne hvis det bedrer prestasjonen				
51. Arbeidsmetodene brukt i denne organisasjonen blir ofte diskutert				
52. Hvorvidt de ansatte jobber effektivt sammen, blir regelmessig diskutert				
53. Denne organisasjonens målsetninger endres i takt med forandringer i miljøet				
54. I denne organisasjonen tar man seg tid til å evaluere organisasjonens målsetninger				
55. Ansatte har en god forståelse av organisasjonens formål				
56. Organisasjonen sin fremtidige retning blir klart og tydelig kommunisert til alle				
57. Ansatte har ikke en klar forståelse av hva som er organisasjonens mål				
58. Alle som jobber her er bevisst på organisasjonens fremtidsplaner og retning				
59. Det finnes en klar oppfatning av hvor organisasjonen går				
60. Tid og penger kunne blitt spart dersom arbeidet var bedre organisert				
61. Ting kunne blitt gjort mye mer effektivt hvis folk tok seg tid til å tenke seg om				
62. Dårlig planlegging resulterer ofte i at man ikke når sine målsetninger				
63. Produktiviteten kunne blitt forbedret om arbeidet ble bedre organisert og planlagt				
64. Hos oss ønsker folk alltid å prestere så godt de kan				
65. Folk er entusiastiske i forhold til jobben sin				
66. Her slipper folk unna med å gjøre så lite som mulig				

67. Folk er innstilt på å gjøre en ekstra innsats for å utføre en god jobb				
68. Her legger ikke folk mer innsats i arbeidet sitt enn det de må				
69. Folk får som regel tilbakemelding i forhold til kvaliteten på det arbeidet de gjør				
70. Folk har ingen anelse om hvorvidt de gjør en god jobb				
71. Det er generelt vanskelig for ansatte å vurdere kvaliteten på det de presterer				
72. Folks prestasjoner måles regelmessig				
73. Måten folk gjør jobben sin på blir sjelden evaluert				
74. Det forventes for mye av folk i løpet av en dag				
75. Vanligvis er ikke folks arbeidsbelastning spesielt krevende				
76. Ledelsen krever at folk jobber ekstremt hardt				
77. Folk er under sterkt press for å nå målsetninger				
78. Arbeidstempoet her er ganske avslappet				
79. Denne organisasjonen forsøker alltid å oppnå de høyeste kvalitetsstandardene				
80. Hos oss blir kvalitet tatt svært seriøst				
81. Folks oppfatning er at organisasjonens suksess avhenger av høy kvalitet på arbeidet				
82. Denne organisasjonen har ikke noe særlig rykte for å levere tjenester av topp kvalitet				

Appendix C – Personal Communication with Michael West

Michael West:

I think you should make the changes in wording you suggest. They will make no material difference to the reliability and validity of the scales.

Yes, in pure terms one should not change wording but comprehensibility from the perspective of those completing the questionnaire is far more important.

Use the words that will make most sense to those completing therefore. Of course one should change only minimally if possible but what you propose seem to me to be absolutely fine.

Best wishes

Michael West

Professor of Organizational Psychology

Lancaster University Management School

Lancaster University

LANCASTER

LA1 4YX

Tel: 01524 510907

Fax: 01524 594720

Mobile: 0777 833 2424

m.a.west@lancaster.ac.uk

<http://www.lums.lancs.ac.uk/profiles/mike-west/>

<http://www.lums.lancs.ac.uk/news/22839/nhs-teamwork-critical/>

Appendix D - Interview Letter of Consent

Project police investigation

We hereby refer to the previous presentation at the national meeting for the Chief of Police. As we explained, the National Police Directorate has given the Norwegian Police University College the task of carrying out a project aiming to assess the organization of police investigative work in Norway.

The project group from the Norwegian Police University College consist of:

- Professor Tor-Geir Myhrer.
- Professor Johannes Knutsson.
- Police inspector Trond Myklebust.

In addition we have a formal cooperation with the professional group at Centre for Applied Positive Work-psychology at the University of Oslo.

We will contact the Chief of Police in each police district requesting participation in the project.

Data will be gathered through interviewing:

- i) Chief of police / Deputy Chief of Police
- ii) Senior Investigating Officer
- iii) Detective/Investigator

Participation in the project is voluntary. Interviewees may withdraw their participation at any given time without providing any explanation for their choice. If a participant withdraws, all responses from him/her will be made anonymous. The data will be treated with confidentiality, and personal identifying information will not be included in the written outputs from the project (i.e. reports/articles)

Signatory will in the upcoming days contact you to schedule the interview with yourself or the Deputy Chief Police. Additionally, I request that you select one Senior Investigating Officer and one Detective/Investigator for the project. I will contact them directly, providing them with information about the project and scheduling a time- and place for the potential interviews.

The interviews will be structured according to the so-called SWOT approach and will include

four themes / questions:

- I. Please tell us what you think works well regarding the investigative work here in the police district - we call this the strength of the investigative work.
- II. Please tell us what you think does not work well regarding the investigative work here in the police district – we call this the weakness of the investigative work.
- III. Please tell us what you consider to be opportunities for improving the investigative work here in the police district – we call this the opportunities in the investigative work.
- IV. Please tell us what you consider to be threats against improving the investigative quality here in the police district - we call this the threats in the investigative work.

We primarily request to tape- and video record the interviews, but if for practical reasons only tape recording is accomplishable, we wish to record interviews in the MP3 format. The duration of interviews will vary according to the amount of information given by the respondent. Experiences with this type of interviews from other projects indicate an average duration of approximately 90 minutes.

Interviews will be made anonymous such that names and personal information will not be transcribed or included in the analyses.

We want to express our appreciation for your police district conveying a positive attitude towards the current project.

Questions or comments to the project may be addressed to the signator:

trond.myklebust@phs.no, tlf direct 23 19 98 55,

Appendix E - Coding Scheme

Unitizing

The definition of a unit:

- In the content analysis, a unit is an identifiable message or message component (Neuendorf, 2002, p. 71)
- Units can be words, characters, themes, time periods, interactions, or any other result of “breaking up a ‘communication’ into bits” (Carney, 1971, p 52, cited in Neuendorf, 2002)
- Generally, units are wholes that analysts distinguish and treat as independent elements. For example, in the operation of counting, the objects that are counted must be distinct- conceptually or logically, if not physically- otherwise the numerical outcome would not make sense. The counting of meanings is problematic unless it is possible to distinguish among meanings and ensure that one does not depend on another (Krippendorff, 2004 p. 97).

The definition of a statement:

A statement was defined as the smallest meaningful unit that reflects the informant’s experience and understanding of the topic of interest, namely innovation climate (Hoff, Flakke et al., 2009, p. 7) A statement is a part of a sentence, a whole sentence, or several sentences expressed by the interviewee, that constitute a coherent, meaningful point of view that describe an aspect of the work environment (Hoff, Straumsheim et al., 2009, p 14). A change from positive to negative or a change in topic may indicate a new statement.

SWOT categories

Statements derived from the transcription will be coded on SWOT, i.e. strengths weaknesses, opportunities and threats. Statements that do not fit the SWOT categories will be coded as residuals.

The SWOT categories

- **Strengths:** Positive aspects of the work environment in the present situation
- **Weaknesses:** Negative aspects of the work environment in the present situation
- **Opportunities:** Future opportunities for a good working environment
- **Threats:** Future threats towards a good working environment
- **SWOT residuals:** Statements that do not fit the presented categories

The context in which each statement appears is taken into consideration during the coding procedure. If the context does not provide enough information regarding appropriate code, the SWOT question preceding the statement is taken into consideration.

Example: “We have many **training days.**” The preceding question and context may provide information regarding appropriate code for this statement.

Examples of statements coded on the four categories + residual:

- **Strengths:** *"There is competition for every vacancy, which results in a highly competent staff"*
- **Weaknesses:** *"We get more cases than we can handle"*
- **Opportunities:** *"A higher degree of flexibility would enhance the organizational performance"*
- **Threats:** *"We face the threat that there is a tendency to choose operative work over investigative work, because it provides a higher income"*
- **Residual:** *"Sorry, I have to take this phone call"*

The IGLO categories

- **The individual level:** Individual perceptions, feelings and opinions
- **The group level:** Interaction and cooperation in work groups, teams and departments
- **The leadership level:** Behaviour of immediate supervisors, other leaders or the top management
- **The organizational level:** Management practices, organizational culture, strategies, organizational goals and values, and the physical environment of the organization
- **IGLO External:** Statements directed towards external matter, e.g. circular letter from the director of public prosecutions
- **The IGLO residuals:** Statements that did not fit the categories above

Procedure:

1. Unitize statements without predispositions
2. Statements are coded on the SWOT categories + residuals
3. Statements are coded on the IGLO categories + the external category and residual

The OCM Codes:

Autonomy: designing jobs in ways which give employees wide scope to enact work.

Integration: the extent of interdepartmental trust and cooperation .

Involvement: employees have considerable influence over decision-making

Supervisory Support: the extent to which employees experience support and understanding from their immediate supervisor

Training: a concern with developing employee skills

Welfare: he extent to which the organization values and cares for employees

Formalization: a concern with formal rules and procedures

Tradition: the extent to which established ways of doing things are valued

Innovation & Flexibility: **flexibility**—an orientation toward change; **innovation**—the extent of encouragement and support for new ideas and innovative approaches

Outward Focus: the extent to which the organization is responsive to the needs of the customer and the marketplace in general

Reflexivity: a concern with reviewing and reflecting upon objectives, strategies, and work processes, in order to adapt to the wider environment

Clarity of Organizational Goals: a concern with clearly defining the goals of the organization

Efficiency: the degree of importance placed on employee efficiency and productivity at work

Effort: how hard people in organizations work towards achieving goals

Performance Feedback: the measurement and feedback of job performance

Pressure to Produce: the extent of pressure for employees to meet targets

Quality in procedures: the emphasis given to quality procedures

Residual: statements that do not fit in any of the 17 other categories