Teamwork behaviour in a bystander conflict environment

van der Hoeven, N.

Award date:
2012

Eindhoven, August 2012

Teamwork behaviour in a bystander conflict environment

by
Nico van der Hoeven

BSc Industrial Engineering and Management Science — TU/e 2011
Student identity number 0616345

in partial fulfilment of the requirements for the degree of

Master of Science
in Operations Management and Logistics

Supervision:
dr. J.M.P. (Josette) Gevers, TU/e, HPM
dr. S. (Sonja) Rispens, TU/e, HPM
dr. ir. K.J.P.M. (Kim) van Erp, TU/e, HPM

TUE. School of Industrial Engineering.
Series Master Theses Operations Management and Logistics

Subject headings: training, teamwork behaviour, bystander conflict, observation.

Abstract

Paramedic teams frequently operate in demanding emergency situations. Their work in these situations is increasingly getting disturbed by undesirable behaviour by an external party. Experiencing this type of undesirable behaviour will be defined as experiencing a *bystander conflict*. In this master thesis, an analysis of the expected negative effect such a bystander conflict has on the paramedic team's performance will be discussed. But also what the role of a training programme focusing on teamwork behaviour can be for the paramedic teams to handle such a situation. The training programme that was part of this master thesis study was conducted at the 'Regionale Ambulance Voorziening' (RAV) Brabant Zuid-Oost.

Summary

The focus of this master thesis is on paramedic teams increasingly experiencing bystander conflicts in their work environment. Defining the bystander conflict as *the perception of an individual (within the primary process) that his/her task(s) or goals are frustrated by one or more persons (outside the primary process) through irrelevant behaviour that interrupts focused attention on the primary task carried out*. Such a bystander conflict is experienced as being an additional demand in a work environment that is already characterised by high demands because it has complex and intensive workflow patterns. If a paramedic team does not have enough available resources to handle the bystander conflict as an additional demand, it can cause an interference with rational thinking and focused attention. This is likely to negatively affect outcomes of the team's task such as its performance, thereby possibly having costly consequences.

For a team to maintain high performance while carrying out a task characterised by a complex and intensive work flow pattern, the critical lever is whether their teamwork behaviour is sufficient. This is especially true for the healthcare sector the paramedic teams are part of because their workflow pattern creates a high within-team interdependence. Trained teamwork behaviour is suggested to function as an additional team resource to buffer the negative effects the bystander conflict is expected to have, for paramedic teams experiencing it. Training is described in scientific literature to be the most effective way to improve teamwork behaviour. It generally has a positive effect on team performance as well. Combining these two expected outcomes of training and the expected negative effect the bystander conflict has, I suggests a research model with teamwork behaviour mediating the expected positive effect training has on team performance, for paramedic teams in a bystander conflict environment.

61 teams took part in a randomized simulation-based pre-test post-test control group design. The experimental group (n = 37) received training aimed at improving teamwork behaviour and the control group (n = 24) only received unstructured feedback. Each team took part in two different scenarios in which they were confronted with a bystander conflict. Data was acquired in two different ways. First of all, the participants filled in three surveys (before the first scenario, and after each of the two scenarios). Secondly, the scenarios the participants took part in were videotaped. Using more than one method of acquiring and analysing data is essential because using only one method can lead to deficient measurement which may beget deficient decisions about the training's effectiveness. To make use of observations, it is essential to have a robust tool for analysis, necessary to benchmark good teamwork behaviour. Because most of the observational tools within literature were not suited for the master thesis' context, one was created.

Mixed support for the research model was found when interpreting the results of the experiment. Based on the survey data, no effect of training on team performance was found. However, the interaction of the type of group with the type of scenario was significant for team performance. This suggestsed that for a more difficult scenario training could create more awareness of what can go wrong, thereby negatively affecting team member's perception of performance (and vice versa for the easier scenario).

Furthermore, no support for teamwork behaviour mediating the effect training was expected to have on team performance was found. Based on the observational data, support was found for a positive effect of training on three out of the five teamwork behaviours observed (i.e. closed-loop communication, adaptive leadership and backup behaviour). These teamwork behaviours were observed significantly more for the paramedic teams receiving training. This indicates support for an important sub-part of the research model.

It is suggested that the differences in the results found for the effect of training on the teamwork behaviours between methods can be explained by the way in which they measure teamwork behaviour. A retrospective relative score (i.e. a comparison between '*what I could have done*' and '*what did I actually do*') is used in the survey, whereas an absolute score (i.e. every occasion of observed teamwork behaviour counts) is used in the observational tool. As such, the participant's perception of their teamwork behaviour via the retrospective relative scores is different than the rater's assessment of them. This supports the suggested necessity in scientific literature of using more than one method of acquiring and analysing data.

Preface

Five years of Industrial Engineering and Management Sciences have flown by and it all comes to an end with this final version of my master thesis. The results of this study would not have been possible without the continuous support of my three supervisors. I would like to thank Josette Gevers, Sonja Rispens and Kim van Erp for their frequent and valuable feedback on all aspects of my report. But also for their general interest and involvement in the research project!

Nico van der Hoeven
August, 2012

Table of content

1. Introduction

"*Article 18 of the Geneva Convention: hospitals organized to give care to the wounded and sick, the infirm and maternity cases, may in no circumstances be the object of attack but shall at all times be respected and protected by the Parties to the conflict*".

"*Article 24 of the Geneva Convention: medical personnel exclusively engaged in the search for, or the collection, transport or treatment of the wounded or sick, …shall be respected and protected in all circumstances"*

"*If we respect these Conventions, in times of war…, then why not, in times of peace?*". This basic premise, alongside the citation of articles 18 and 24 of '*Convention IV relative to the Protection of Civilian Persons in Times of War'* of the Geneva Convention[1], are the cornerstone of the most recent advertising campaign of the *Stichting Ideële Reclame* (SIRE): '*Keep your hands of our aid workers'*[2]. The SIRE-campaign was initiated by the results of a five year longitudinal study '*Agressie en geweld tegen werknemers met een publieke taak'* (Abraham, Flight & Roorda, 2011), summarizing the progress of the *Veilige Publieke Taak* (VPT) programme. A weighted average of 66.3% of people with a public safety task and 59% of people with a public services task in general experience aggression and/or violence in 2011 was measured. These are alarming figures.

Abraham *et al.* (2011) used the more general term of people experiencing *undesirable behaviour* by an *external party* because the terms 'aggression and/or violence' are often considered to be misleading. Undesirable behaviour includes, for example[3], *verbal aggression* (e.g., shouting, humiliation, or teasing) or *physical aggression* (e.g., punching, kicking, or destroying property). People indicate that as a result of experiencing this undesirable behaviour they experience stress (35%), a decrease in job satisfaction (20%), are mentally exhausted (13%), or their job performance decreases (9%).

Based on the longitudinal results of their study, Abraham *et al.* (2011) furthermore show that even though employees with a public services task state they experience less undesirable behaviour in 2011 than in 2007 (59% instead of 66%), the intended programme's target of 51% is not yet reached. The 7% decline can be attributed to the fact that the Netherlands is one of the leading countries in Europe when it comes to having a centralized planned approach to tackle the issue (e.g., when compared to the United Kingdom, Germany and Sweden (Kerckhaert & Ruig, 2010)). This planned approach was initiated by the Balkenende II cabinet in 2005 and makes it easier to report undesirable behaviour to the appropriate authorities, redress incurred costs on the perpetrator(s), and creates an improved support system for victims. However, Abraham *et al.* (2011) indicate there are some promising areas suggested by the programme left that receive too little attention. Especially training focusing on how to handle undesirable behaviour shows

[1] http://www.icrc.org – International Committee of the Red Cross. International Humanitarian Law – Treaties and Documents.
[2] http://www.handenaf.nl – SIRE: Handen af van onze hulpverleners.
[3] Abraham *et al (*2011) describe five types of undesirable behaviour: *verbal aggression, physical aggression, intimidation, sexual harassment* and *discrimination.*

promise, but its use has decreased significantly (from 50% in 2009 to 46% in 2011). Even though most other promising areas got more attention.

Abraham *et al.* (2011) incorporate 15 different professions within six different sectors. The main focus of the SIRE-campaign is on the sectors of public safety. SIRE argues this is mainly because of the public opinion in general, as 88% of people simply will not stand for undesirable behaviour against public safety personnel anymore (i.e. it is *common sense* in a civilized society not to do this). Besides public opinion, there also appears to be only little improvement considering the percentage of people experiencing undesirable behaviour in the public safety sector, when compared to other sectors (Abraham *et al.*, 2011). More specifically, the SIRE campaign focuses on the healthcare professions. The percentage of hospital personnel experiencing undesirable behaviour has significantly increased (75% in 2011, compared to 69% in 2007) and the number of effective programme initiatives has decreased significantly for ambulance personnel. Besides these figures, Kerckhaert and Ruig (2010) mainly argue that compared to other professions within the public safety sector, the healthcare professions are especially vulnerable as they have less means to protect themselves against undesirable behaviour (e.g., when compared to police officers). Given that the number of initiatives for ambulance personnel have significantly decreased over the years, they will be the main focus of this master thesis.

1.1 Research setting

This master thesis will be conducted in collaboration with the ambulance service Regionale Ambulance Voorziening (RAV) Brabant Zuid-Oost in the Netherlands. The RAV as an organization consists of the *Meldkamer Ambulancezorg* (i.e. the emergency room for the ambulance service), the *GGD* (the *Gemeentelijke Gezondheidsdienst,* i.e. the local health authorities) and the *Veiligheidsregio Brabant Zuid-Oost* (i.e. the local safety region). As such, it is one overall organization handling local healthcare more efficiently. Within the Netherlands, there are 25 RAV-regions, employing over 5000 people. Being responsible for 17 out of the 695 ambulances and approximately 35000 out of the 1.1 million rides, the Brabant-Zuidoost region is responsible for 2,5% of the staff and 4% of the work. Each ride consists of a team of one nurse and one driver. These teams will be called *paramedic teams* in the remainder of this master thesis.

Rides are categorized in three urgency-classes: A1, A2 and B. The A1 and A2 class are sub-classes of the *emergency care class*. In the emergency care class, patients run the risk of dying or suffering permanent health risks. In these cases the ambulance should be at the requested location within 15 (A1) to 30 (A2) minutes. The B urgency care class is for the more general medical care that was requested by a made appointment. As such, the B urgency class can be planned efficiently and there are no health risks for the patient that are not known in advance. Of the 35000 rides taken care of by the RAV Brabant-Zuidoost, 80% is categorized as an emergency care class with an equal distribution of A1 and A2 rides.

1.2 Master thesis objectives- and structure

Training is an important aspect to improve medical performance in general (Salas, Diaz-Granados, Weaver & King, 2008). Salas *et al.* (2008) propose that training in the healthcare

sector should emphasize teamwork aspects, to achieve the best results in improving medical performance.

This leads to the goal of this master thesis. That is:

1. *Suggest and analyse a research model that incorporates the expected effects training focused on teamwork behaviour has on team performance, for paramedic teams experiencing undesirable behaviour*.

To satisfy the goal *of* the master thesis, there are several sub-goals *within* the thesis. They will be described below.

The training sessions at the RAV Brabant Zuid-Oost are part of a randomized simulation-based pre-test post-test experimental control group design. That is, there are two randomized groups of which only one will receive the training. Both groups take part in a pre-test (first scenario) and a post-test (second scenario). The training takes place between the two scenarios. Besides having more than one measurement moments in time, more than one method to obtain data is used. First of all, all participants fill in multiple surveys. Secondly, all the scenarios the participants take part in are videotaped. Using multiple methods of obtaining data to evaluate scenario-based training is essential because using only one method can lead to *deficient measurement* which may *beget deficient decisions* about the training's effectiveness (Rosen, Salas, Wilson, King, Salisbury, Augenstein, Robinson & Birnbach, 2008). More specific for training teamwork behaviour, Jeffcott and Mackenzie (2008) even argue that solely using the results of a survey "*cannot be relied upon in isolation to capture the complexity of teamwork in a healthcare setting"* (Jeffcott & Mackenzie, 2008, p.192). The use of observational data alongside surveys is important as the actual teamwork is analysed with it, rather than the retrospective perception of it.

Measurement of different aspects of teamwork with the use of surveys is relatively straightforward. This is less so for the use of observations as measurement. To make use of observations, it is essential to have a *robust tool*, being "*necessary to benchmark good teamwork skills"* (Hull, Arora, Kassab, Kneebone & Sevdalis, 2011, p.234). Because most *observational tools* in scientific literature are either too job specific or do not benchmark teamwork behaviour, one is created for the specific context of this master thesis. As such, the first sub-goal within this master thesis is:

1. *Create an observational tool usable for coding teamwork behaviours in the context of the experimental design the participants take part in.*

The methodology of the observational tool (i.e. more specific the considerations for why observation is relevant, a theoretical foundation on which the observational tool is based, and, its actual design), is described in Chapter 3.

Secondly, besides the creation of the observation tool, all the videotaped scenarios are analysed with it. As such, the second sub-goal within this master thesis is:

2. *Assess teamwork behaviour observed in the context of the experimental design with the observational tool, to evaluate the research model.*

Lastly, the data acquired by the surveys is used. As such, the third sub-goal within this master thesis is:

3. *Assess teamwork behaviour and team performance in the context of the experimental design with the survey data, to evaluate the research model*

The results of the second and third sub-goal are described in Chapter 4. Results of the second sub-goal are an indication about the validity of the observational tool in measuring teamwork behaviour, and which role these observed teamwork behaviours play in the research model. The second and third sub-goal combined provide an answer to the question whether training actually improves team performance for paramedic teams in a bystander situation, and what the effect of teamwork behaviour is. Conclusions related to the results of Chapter 4 are described in Chapter 5.

2. Theoretical framework

In this chapter, I describe the theoretical background on which this master thesis is based. In Paragraph 2.1, I will introduce the definition of the bystander conflict. In Paragraph 2.2, I will introduce the term *High Reliability Organisation* and how this term is applicable to paramedic teams. In Paragraph 2.3, a description is given of the expected results of experiencing a bystander conflict, based on the *resource-demand* literature. In Paragraph 2.4. I explain which teamwork behaviours the training should focus on, and why. In Paragraph 2.5. a description is given as how training is expected to improve the bystander conflict situation, and a research model based on this expectation is suggested. Lastly, hypotheses are formulated in Paragraph 2.6.

2.1 A phenomenon in search of a definition

As mentioned in the introduction, Abraham *et al.* (2011) distinguish between five types of *undesirable behaviour* by an *external party*. They also describe several types of external parties, one of them being the *bystander*[4]. Abraham *et al.*'s (2011) classification furthermore mainly focuses on *aggression*. Even though aggression is described to be an increasingly important topic within the healthcare sector, non-aggressive behaviour should also be taken into account (Hershcovis & Barling, 2010). Furthermore, the described behaviour is stated as *being expressed on purpose* (Abraham *et al.*, 2011). This is also not the complete picture: one could imagine that behaviour not being expressed on purpose (e.g., a worried family member) is still experienced as *undesirable behaviour*. The *bystander conflict* focuses on all aspects of behaviour by bystanders.

'*Conflict*' has been a popular topic for researchers the last few decades, debating effect size and direction (de Dreu & Weingart, 2003) for different types of conflict (e.g., task, process or relationship conflict) on several related outcomes (e.g., job satisfaction or task performance). However, *conflict literature* mostly investigates an artificial or workplace setting, thereby often neglecting aspects of a more realistic situation. De Dreu (2008, p.6) defines conflict as when "*one party –be it an individual or a group of individuals– perceives its goals, values, or opinions being thwarted by an interdependent counterparty*". *Impeding behaviour* (e.g., worried individuals distracting ambulance personnel during an emergency), *intruding* and *aggressive behaviour* (e.g., a patient verbally threatening a doctor in his/her practice during working hours), and *extreme forms of aggressive behaviour* (e.g., a group of over 250 individuals destroying ambulances and assaulting the responding police officers[5]) partly fit this definition. However, paramedic teams experiencing a bystander conflict are no interdependent counterparts of the conflict party. As such, one aspect of the bystander conflict is that the persons experiencing the conflict are not in the same *primary process.*

Besides people not being in the same primary process, they are generally not familiar with each other. The interaction between parties is therefore characterised by being *unanticipated* and *unpredictable*. It can also be classified as an *interruption* (Jett & George, 2003). Jeff and George (2003) describe several kinds of interruptions. The bystander conflict resembles what they

[4] The others types included by Abraham *et al.* (2011) as being an external party are, for example, customers, patients, or inmates.
[5] Even though these examples range from small impeding distractions to extreme forms of aggression, all three examples are based upon news published in various newspapers and magazines (Dutch media agency *ANP* as a source).

describe as an *intrusion*; defined as "*unexpected encounters initiated by another person that interrupts the flow and continuity of an individual's work and brings that work to a temporary halt*" (Jeff & George, 2003, p.495), as well as a *distraction*, defined as "*psychological reactions triggered by external stimuli or secondary activities that interrupt focused concentration on a primary task*" (Jett & George, 2003, p.500). Besides these two types of intrusions, the bystander conflict has three more additions (van Erp, Demerouti, Gevers & Rispens, 2011). It is: (1) *not part of the job*; (2) *irrelevant to the job*; and (3) *not schedulable*.

Combining these aspects of *conflict* and *interruption* literature leads to the following definition for the bystander conflict:

The bystander conflict is *the perception of an individual (within the primary process) that his/her task(s) or goals are frustrated by one or more persons (outside the primary process) through irrelevant behaviour that interrupts focused concentration on the primary task being carried out.*

2.2 Paramedic teams as 'High Reliability Organization'

Within the scientific literature, the healthcare sector is classified as a High Reliability Organization (HRO). HROs are defined as "*organizations that exist in such hazardous environments where the consequences of errors are high, but the occurrence of error is extremely low*" (Baker, Day & Salas, 2006, p.1576). For example, in the Netherlands there was an estimate of 1735 patients dying that could be related to preventable circumstances, out of 1.3 million take-ins in 2004 (Wagner & De Bruijne, 2007). Only being approximately 0.1% of the total number of take-ins emphasizes the sector's relative low occurrence of errors, but severity of the situation if they occur. Roberts and Rousseau (1989) created a list of eight criteria to classify an organization as an HRO: (1) *hyper-complexity*; (2) *tight coupling*; (3) *extreme hierarchical differentiation*; (4) *large number of decision makers in complex communication networks*; (5) *degree of accountability that does not exist in most organizations*; (6) *high frequency of immediate feedback about decisions*; (7) *compressed time factors*; and (8) *more than one critical outcome that must happen simultaneously*. Based on the HRO definition and the eight criteria, Roberts (1990) describes examples of traditional HROs in scientific literature to be Air Traffic Control (ATC) and nuclear power plants. However, more recently the healthcare sector was characterised as an HRO as well, as all criteria's described by Roberts and Rousseau (1989) are applicable to it (Hines, Luna & Lofthus, 2004).

For most HROs the work environment is characterised by *teams* having an *intensive pattern of workflow* (Tesluk, Mathieu, Zaccaro & Marks, 1997). This workflow pattern indicates that "*work and activities come into the team and members must diagnose, problem solve, and/or collaborate as a team in order to accomplish the team's task*" (Tesluk *et al.*, 1997, p.201). For this intensive work flow pattern to lead to sufficient performance, *especially* in the healthcare sector, the critical lever is whether teamwork behaviour is sufficient (Salas, Rosen & King, 2007). Teamwork behaviour is critical because the workflow pattern creates a high within-team interdependence. Before discussing which teamwork behaviours are important, and why, it is first necessary to define what *teams* and *teamwork behaviour* are precisely. As described in the introduction, the focus of this master thesis is on paramedic teams consisting of one nurse and one driver.

A *team* is defined as "*two or more individuals who: (1) socially interact; (2) possess one or more common goals; (3) are brought together to perform organizationally relevant tasks; (4) exhibit interdependencies with respect to workflow, goals, and outcomes; (5) have different roles and responsibilities; and (6) are together embedded in an encompassing organizational system, with boundaries and linkages to the broader system context and task environment*" (Kozlowski & Ilgen, 2006, p.79). Within the *team literature,* the view of teamwork behaviour shifted from a static *input-process-output* (IPO) model (McGrath, 1964), to a dynamic view of teamwork behaviour and its effectiveness, the *input-mediator-output-input* (IMOI) model (Ilgen, Hollenbeck, Johnson and Jundt, 2005). The IMOI-model chooses a mediator perspective rather than a basic process perspective to include a broader range of variables that might be important to explain performance. It furthermore acknowledges that not all inputs result in output necessarily by only a process. Within this IMOI-model, a distinction between *teamwork processes* and *emergent states* is made (Marks, Mathieu & Zaccaro, 2001). Teamwork is defined as "*a set of interrelated thoughts, actions, and feelings of each team member that are needed to function as a team and the combine to facilitate coordinated, adaptive performance and task objectives resulting in value-added outcomes"* (Salas, Sims and Burke, 2005, p. 562). Emergent states are defined as "*properties of the team's that are typically dynamic in nature and vary as a function of team context, inputs, processes and outcomes* (Marks *et al.,* 2001, p.357). As such, emergent states describe states of a team (i.e. *cognitive*, *affective*, and *motivational*), that can be considered both input and/or proximal outcome, but not the actual teamwork process. For example, members of a team experiencing low mutual trust (an emergent state) may start to communicate less (teamwork process), which could lead to even less mutual trust experienced (proximal outcome as well as possible new input). Both teamwork processes and emergent states are included whenever the term *teamwork behaviour* is mentioned in this remainder of this master thesis.

2.3 Consequences of the bystander conflict: a demand-resource approach

Conflicts are generally negatively related to job satisfaction, job performance or personal well-being (de Dreu, 2008; Rispens & Jehn, 2010; Diskell & Salas, 1991). In the particular job-setting in which teams experience intensive patterns of workflow and teamwork behaviour is critical, the effect of a conflict interfering with rational thinking is of particular importance (Pruitt & Rubin, 1986). The interference with rational thinking is caused by a disparity between an additional demand (e.g., the bystander conflict) and the available resources the team has (e.g., teamwork behaviour). The effects of this disparity are described as the *demand-resource perspective*. The demand-resource perspective is well described in literature, by multiple models (e.g., Hobföll, 1989; Karasek, 1979; Siegrist, 1996; Demerouti, Bakker, Nachreiner & Schaufeli, 2001). These models have in common that there should be a *stable balance between available resources* and the *demand that is placed on them*. Even though some research suggests that adding job demands can sometimes improve performance due to making the work more challenging (LePine, Podsakoff & LePine, 2005), this is not the case for job demands of an HRO as they are inherently characterized by having high job demands.

The oldest model is the *Conservation of Resources* model (COR; Hobföll, 1989). The COR-model is a stress-related model, defining resources as "*… those objects, personal characteristics, conditions, or energies that are valued by the individual or that serve as a means for attainment*

of these objects, personal characteristics, conditions or energies" (Hobföll, 1989, p.516). The COR-model suggests that the main goal for human motivation is maintaining its current state of resources and try to accumulate additional ones. However, the COR-model is fairly general, and other, more modern models are better applicable to a job-related environment (Bakker & Demerouti, 2007).

The most basic of those more job-related models, is the *Demand Control Model* (DCM; Karasek, 1979, 1988). According to the DCM-model, there is a relation between *job decisional latitude* and *job demands*. Karasek (1979, 1988) suggests that when the job decisional latitude a person experiences is low and his/her job demands are high, the job can be described as a *high strain job*. Such a high strain job results in a state of unresolved strain. This premise is supported by scientific literature (e.g., Karasek, 1979; Schnall, Landsbergis & Baker, 1994). The DCM-model is often criticized for its simplistic nature (De Jonge, Mulder & Siegrist, 1999), even though this does not always have to be a drawback per se.

Alongside the DCM is the *Effort-Reward Imbalance model* (ERI; Siegrist, 1996). Instead of focussing on control aspects of the job, as the COR-model emphasizes, the reward structure of the job is assumed to be important. Siegrist (1996) suggests that strain results when an imbalance between *effort* (job demands) and *rewards* (intrinsic motivation) is experienced, specifically when effort is particularly high and reward particularly low. The main critique on the ERI-model is the same as the DCM: its simplicity (Van Vegchel, de Jonge, Bakker & Schuafeli, 2002).

More recently developed, and more complete, is the *Job-Demand Resources* model (JD-R) (e.g., Bakker, Demerouti, De Boer, & Schaufeli, 2003; Demerouti *et al.,* 2001). It will therefore be used as main input for the research model in this master thesis. The demand-resource perspective of the JD-R model is similar to what Karasek (1998) described for the DCM. However, it takes a more narrow view on *resources* while taking a broader definition of *demands*. It incorporates *job resources*, defining them as "*physical, psychological, or organizational aspects of the job that are either: (1) functional in achieving work goals; (2) reduce job demands and the associated physiological and psychological costs;* or *(3) stimulate personal growth, learning and development*" (Bakker & Demerouti, 2007, p.312). Given this definition, and that of teamwork behaviour in the previous paragraph, it is suggested that teamwork behaviours are part of *job resources*. Being an important asset on their own, as well as to cope with demands in the JD-R model.

The JD-R model is depicted in Figure 2.1. There is a negative reciprocal relation between job demands and job resources. This means that the higher one's job demands are, the less resources one will have at his/her disposal, and vice versa. Furthermore, two underlying different psychological processes are supported, thereby being a *dual process model* (Bakker & Demerouti, 2007). First of all, there is a relation between an increase in job demands and an increase in experienced strain (which can be buffered by an increase in job resources) (e.g., Bakker *et al.,* 2003; Diener & Fujita, 1995; Kahn & Byosserie, 1992). However, this effect can be different for different kinds of resources (Bakker & Demerouti, 2007). The increase in experienced strain is negatively related to outcomes (Leiter, 1993). Secondly, an increase in available job resources increases motivation, which can be buffered by an increase in job demands. An increase in motivation is positively related to outcomes. Both processes are

supported by literature (e.g., Bakker *et al.*, 2003; Bakker *et al.*, 2004). For example, Bakker *et al.* (2003) tested the model's predictive capabilities on self-papered absenteeism and turnover intentions within a Dutch telecom company. They report that job demands predicted health problems (i.e. *strain*) best, which was in turn related to absenteeism (i.e. *outcomes*). Job resources were found to be the only predictors of dedication and commitment (i.e. *motivation*), in turn being related to turnover intentions (i.e. *outcomes*). Similarly, Bakker *et al.* (2004) reported that job demands were the best predictors of a burnout (i.e. *strain*), which in turn was the best predictor for in-role performance (i.e. *outcomes*). Whereas job resources were the best predictors engagement (i.e. motivation), in turn related to extra-role performance (i.e. *outcomes*).

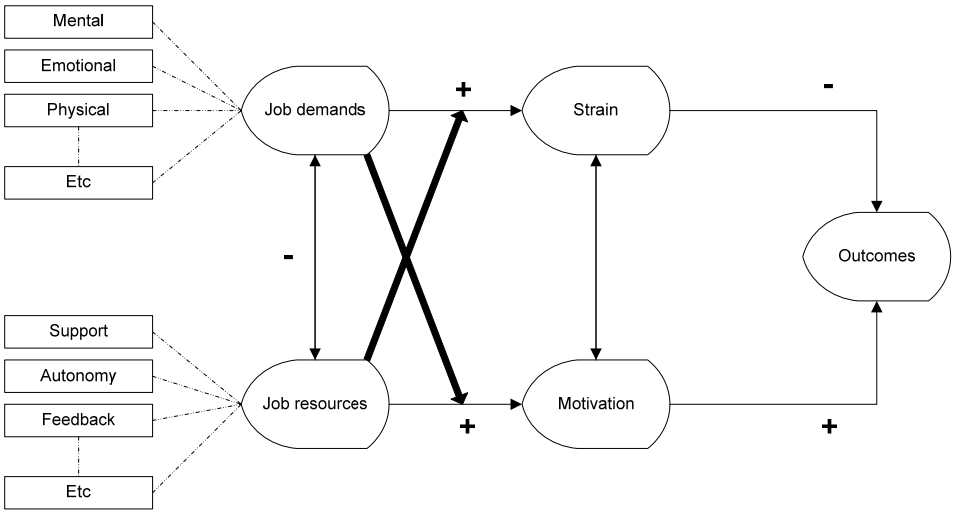

Figure 2.1: The Job Demand-Resources model (JD-R), derived from Bakker & Demerouti (2007).

An important implication of the JD-R model regarding the consequences for paramedic teams of experiencing a bystander conflict, is the *resource-demand quadrant* for predictive purposes (Figure 2.2). Paramedic teams already function in a high demand situation and the bystander conflict acts as an additional demand depleting available resources. This will cause high strain and low motivation, *if* the demand is not handled properly (Figure 2.2, lower right).

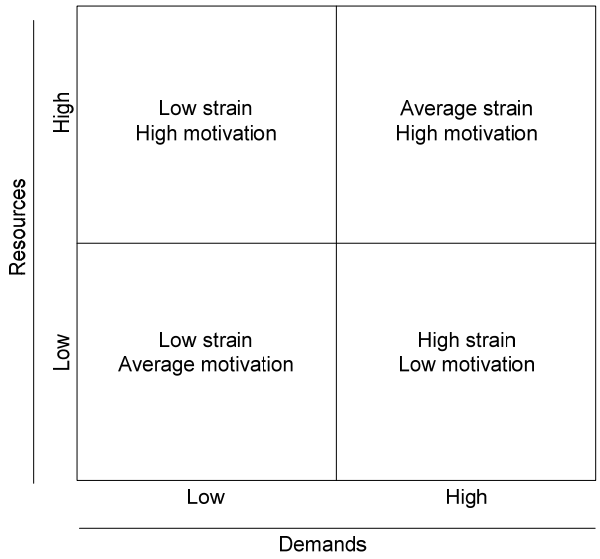

Figure 2.2: Resource-demand quadrant for predictive purposes, derived from Bakker & Demerouti (2007)

Furthermore, paramedic teams often deal with an additional demand: time pressure. Karau and Kelly (1992) describe the interacting relation time pressure can have as a limited resource with other variables (i.e. task features, group structure, individual differences, and situational factors) and how this can narrow the focus of team members (the Attentional Focus Model; AFM, Figure 2.3). This narrowing focus can decrease the amount of interacting behaviour within a team (Gevers, van Erven, de Jonge, Maas & De Jong, 2010), being one of the most crucial factors of interdisciplinary teamwork (Baker *et al.*, 2006).

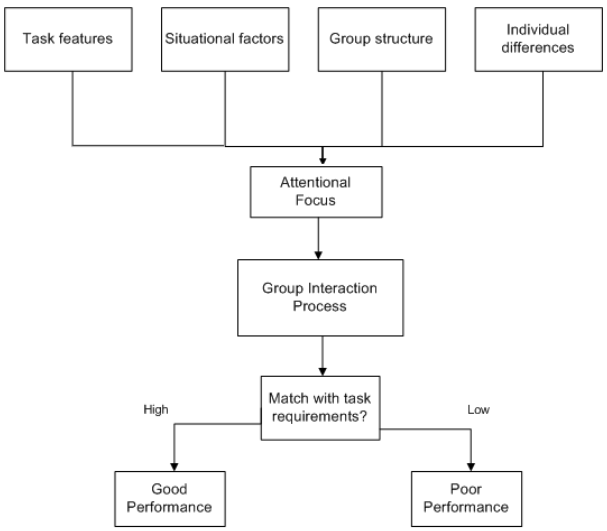

Figure 2.3: the Attentional Focus Model (AFM), derived from Karau & Kelly (1992)

Because of a decreasing amount of interacting behaviour, more attention is put on individual activities. This occurs because in a time pressure situation individual activities appear to be more

important and thereby teamwork behaviour becomes less salient (Karau & Kelly, 1992). The AFM is furthermore consistent with several earlier theories, stating that factors such as *arousal*, *stress*, and *distractions* can also cause a narrowed focus (Hirokawa, 1988).

Karau and Kelly (2003) argue that time pressure can both enhance or reduce performance based on the requirements of the task. However, in complex situations, time pressure reduces performance. A complex environment needs a high degree of coordination (e.g., information processing) which will be disturbed by a lack of focus on the task created by time pressure. More specific to the healthcare setting is recent research by Gevers *et al.* (2010). They investigated the effects of acute and chronic job demands and their effect on strain, within the Emergency Department of a Dutch general hospital. They describe that high acute job demands cause teamwork behaviour to be less effective and results in acute strain. Similarly, emotional demands can result in acute job strain if chronic emotional job demands are already high. This suggests that acute job strains are detrimental to effective teamwork behaviour in situations in which it is most required. Whenever paramedic teams would experience a bystander conflict, the conflict is an additional acute demand in the situation. It thereby would be detrimental to effective teamwork behaviour and create a high strain situation, as also predicted by the Resource-demand quadrant.

2.4 Teamwork behaviour

Even though *teamwork behaviour* appears to be a popular topic among researchers, there was not always much consensus among them until recently. To cover all different frameworks created in the last 25 years of research (*over 140 frameworks*), Salas *et al.* (2005) suggest the *Big Five model of teamwork*.

The Big Five model is a taxonomy consisting of five main teamwork behaviours being subject to three coordination mechanisms. The three coordination mechanisms are: (1) *shared mental models*; (2) *closed-loop communication*; and (3) *mutual trust*. They ensure information is constantly updated and properly distributed within the team. The five main teamwork behaviours are: (1) *team leadership*; (2) *adaptability*; (3) *mutual performance monitoring*; (4); *backup behaviour* and (5) *team orientation*. These teamwork behaviours are especially relevant for teams operating in highly interdependent environments, having an intense pattern of workflow. As such, they are appropriate for the paramedic teams which are the focus of this master thesis.

I will elaborate on each teamwork behaviour of the Big Five model below, and explain why they could be relevant for a paramedic team facing a bystander conflict. I start with the three coordination mechanisms (Paragraph 2.4.1 to 2.4.3) followed by the main teamwork behaviours of the Big Five model (Paragraph 2.4.4 to 2.4.8). Furthermore, I suggest that situational awareness is an important teamwork behaviour (Salas, Prince, Baker & Shrestha, 1995), complementing the Big Five model. Situational awareness is described in Paragraph 2.4.9.

2.4.1 Shared mental models

Rouse and Morris (1986, p.7) define mental models as "*mechanisms whereby humans generate descriptions of system purpose and form, explanations of system functioning and observed system states, and prediction of future system states*". When these mental models are shared

within a team, individuals can use them to coordinate information about the dynamics of the team to attain a common goal, required for sufficient team performance (Zaccaro, Rittman & Marks, 2001). Team-level mental models focus on the functioning of the team and the expected behaviours within the team. Kleinman and Serfaty (1989) suggest that shared mental models are of more importance for teams that function in a stressful and interdependent situation. In such a situation explicit communication decreases and there is a need to shift onto implicit communication.

For paramedic teams experiencing a bystander conflict, the amount of stress and pressure experienced by the team members is expected to increase. The amount of explicit communication used during teamwork behaviour is thereby decreased. When shared mental models are improved within the team, team members' expectations of needs and behaviours are more accurate. This improves the explicit communication process, required for sufficient team performance.

2.4.2 Closed-loop communication

For interdependent teams in a complex situation, mechanisms that decrease miscommunication or increase the use of more implicit communication are invaluable for other teamwork behaviours (Salas & Cannon-Bowers, 2000). Miscommunication might occur in two ways: (1) *misinterpretation due to an external factor*; or (2) simply *not receiving the communicated message*. To improve communication within a team, Salas *et al.* (2005) suggest using *closed-loop communication*. Team members should: (1) *initiate;* (2) *acknowledge receipt;* and (3) *follow-up* communications, to prevent miscommunication. Besides the transfer of information, *verification of accuracy*, *timeliness* and *clarity* are similarly important aspects of sent communications (Temkin-Greener, Gross, Kunitz & Mukamel, 2004). Seigel and Federman (1973) reported that teams trained on dimensions intended to increase closed-loop communication performed better, compared to untrained teams.

Paramedic teams experiencing a bystander conflict are likely to lose their focus on communication processes by being distracted by the bystander. This could cause misinterpretation, as the communicated message could be received only party (or *not at all*). Therefore, teams trained on using closed-loop communication will be more effective in "*combating information exchange difficulties and ensuring that sent communications are heard and accurately understood*" (Salas *et al.,* 2005, p.568). This decreases the likelihood of errors and thereby improves overall performance.

2.4.3 Mutual trust

Mutual trust is related to the amount of trust team members experience when interacting as a team. Trust is defined as "*the shared perception that individuals in the team will perform particular actions important to its members and will recognize and protect the rights and interest of all the team members engaged in their joint endeavour*" (Webber, 2002, p.205). As such, it is a mechanism by which team members rely more on each other. Team members will share more information and cooperate better (Jones & George, 1998). This occurs because they feel that their input is valued more (Bandow, 2001). Furthermore, it influences how team members interpret each other's behaviour (Simons & Peterson, 2000), similar to the effect of shared

mental models. When mutual trust is high, other teamwork components' s effectiveness is enhanced further because they can be interpreted better.

Paramedic teams experiencing a bystander conflict will handle a situation less effective if they experience low mutual trust. For example, if the team has a capable team leader with a good strategy, but mutual trust is low, team members perhaps wilfully will not follow the team leader's strategy (or, do follow it even when they know the strategy it is flawed). To prevent such a situation from happening, sufficient mutual trust is a necessary prerequisite.

2.4.4 Adaptability

Adaptability is defined as "*the ability to recognize deviations from expected action and readjust actions accordingly*" (Salas *et al.*, 2005). A team *culture of adaptability* is the need to understand the team's task and the possibility and recognition of possible changes in roles to accomplish that task. Adaptability is important for all kinds of teams, but is particularly driven by a complex task having a dynamic nature. Salas *et al.* (2005) suggest the importance of *quality of adaptation* and not simply the occurrence of it.

Paramedic teams experience the bystander conflict as a performance decreasing deviation from what is expected. As such, it is necessary to cope with this by effectively adapting a strategy for this new situation. This requires shifting roles whenever required. Teams more capable in adapting to a new situation, are expected to better be able to handle such a situation.

2.4.5 Team leadership

Most of what is known about team leadership is derived from individual leadership theory (Fleishman, Mumford, Zaccaro, Levin, Korotkin, & Hein, 1991), and should be interpreted with some caution. Team leaders use different roles to enable: (1) *effective teamwork behaviour*; and (2) *interdependent action*. Both aspects are used to effectively diagnose problems and to generate and implement solutions for them (Fleishman *et al.*, 1991). The first role is related to shared mental models. Team leadership requires the *creation*, *maintenance*, and *accuracy* of the mental models the team should use (Salas *et al.*, 2005). The second role, *monitoring the internal and external environment*, is related to the constant enhancement of adaptability. Enhancing adaptability is necessary in a changing environment by the proper use of *information*, *coordination of behaviour* and *general development of skills*. The third role is establishing a climate that encourages *interaction patterns*, *mutual performance monitoring*, and *backup behaviour* (Burke, Stagl, Klein, Goodwin, Salas & Halpin, 1996). These three roles should enhance team effectiveness by a general understanding of the existing interdependence within the team and the advantages of working together, to overcome certain situations (Zaccaro *et al.*, 2001).

Paramedic team are often characterised by a small team size and an interdependence mainly due to differences in roles, rather than hierarchy. As such, there often is no real *team leader*. Therefore, *shared leadership* should be considered. Shared leadership is an important form of leadership for cross-functional teams that are not hierarchically differentiated. Shared leadership is defined as "*a collaborative, emergent process of group interaction in which members engage in peer leadership while working together*" (Pearce & Conger, 2003, p.53). In contrast to normal

team leadership, it is a reciprocal concept of leadership, rather than being unidirectional. Even though most researchers suggest an overall positive effect of shared leadership on *team performance* (Pearce & Sims, 2002), the amount of research remains limited. Paramedic teams can use the three roles of leadership to diagnose the problems created by the bystander conflict and implement a solution to them. Besides the expected positive effect leadership generally has on team performance in other situations.

2.4.6 Mutual performance monitoring

Mutual performance monitoring is defined as "*the ability to keep track of fellow team members' work, while carrying out their own. To ensure that everything is running as expected and to ensure that they are following procedures correctly*" (McIntyre & Salas, 1995, p.23). Salas *et al.* (2005) suggest that effective teams consist of team members monitoring other members to prevent mistakes, or other aspects possibly decreasing performance as soon as possible before or after their occurrence. Baker, Day, and Salas (2006) indicate that mutual performance monitoring is a prominent necessary teamwork behaviour to overcome conflict. In a conflict situation, if there is high performance monitoring intended to improve performance and its specific intent is clear, individuals not 'correctly' adhering to what is necessary to solve the conflict can be corrected. In the same situation however, but experiencing low performance monitoring, individuals not sufficiently doing what is necessary to solve the conflict will keep on doing so.

Research by Doten, Cockrell and Sadacca (1966) shows that a possible compounding factor to making mistakes is *not being aware of one's own performance deficiencies*. Using mutual performance monitoring to identify performance deficiencies and cope with them via feedback and backup behaviour, is expected to increase team performance to more than just the sum of individual behaviours. Furthermore, it is suggested that shared mental models and mutual trust are prerequisites for mutual performance monitoring to have a significant effect, due to their nature of creating a shared, open, and trusting team climate (McIntyre & Salas, 1995).

For paramedic teams, there is a necessity to follow procedures correctly and receive feedback while performing their task. They should also be able to react on errors as quickly as possible, due to their high impact whenever they are not solved quickly or adequately enough. Paramedic teams trained on mutual performance monitoring are expected to provide this feedback more efficiently (i.e. *more frequent, real-time updated,* and *understandable*). Team members that effectively monitor each other will therefore be more effective than the sum of all member's individual behaviour. Mutual performance monitoring is furthermore expected to increase the handling of the conflict without team members neglecting other task aspects.

2.4.7 Backup behaviour

Backup behaviour is defined as "*the discretionary provision of resources and task-related effort to another, when there is recognition by potential backup providers that there is a workload distribution problem in their team*" (Porter, Hollenbeck, Ilgen, Ellis, West & Moon, 2003, p.391-392). It is split up into three aspects: (1) *provide feedback and coach to improve performance;* (2) *assisting team members when performing a task*; and (3) *help team members to complete their task if that member is overloaded* (Marks *et al.*, 2001). Based on these three aspects, backup behaviour is generally assumed to increase team performance by increasing the flexibility of

teamwork behaviour by which work can be completed within the team (Campion, Medsker & Higgs, 1993). Shared mental models and mutual performance monitoring are often required as antecedents for backup behaviour (Salas *et al.,* 1997).

Backup behaviour differs from simply *helping*, as backup behaviour is a "*recognition of a genuine need for assistance*" (Salas *et al*., 2005, p.580). As such, backup behaviour is a legitimate request for assistance that should not lead to decreasing performance whenever other tasks receive less attention. This is something that might occur if the request for help is not legitimate (Porter *et al.*, 2003).

Paramedic teams experiencing a bystander conflict have to cope with a situation that has the potential to cause a workload distribution problem. Using backup behaviour is useful because it enables team members to recognize the need for legitimate assistance. It thereby decreases the effects of the workload distribution problem because team members are able to take over a task, or help overcome the conflict *as a team* by combining teamwork behaviour. Furthermore, important tasks receiving less effort due to the conflict can be continued by another team member.

2.4.8 Team orientation

Team orientation is an attitudinal aspect rather than a behavioural aspect. It is defined as "*the tendency to coordinate, evaluate, and utilize task inputs from other group members in an interdependent manner in performing a group task*" (Driskell & Salas, 1992, p.278). Whenever team orientation is present in the team, team members take a team attitude, rather than an individual attitude. There will be an increasing attitude towards mutual performance monitoring and backup behaviour. In prior research, it has also been found to increase overall team performance (Driskell & Salas, 1992) via multiple aspects (e.g., increased cooperation and coordination (Wagner, 1995)).

Teams taking other people's attitudes into account are more effective than teams characterised by only individual attitudes. Paramedic team experiencing a bystander conflict are expected to handle the conflict more effectively if they have a better team orientation because of increased cooperation and coordination between team members.

2.4.9 Situational awareness

Situational awareness is defined as "*the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future"* (Endsley, 1995, p.36). Therefore, situational awareness is not static but a reciprocal behavioural process. It focuses on *cognitive* aspects (i.e. the integration and understanding of information), *temporal* aspect (i.e. taking into account relevance of past and future events) and *synchronization aspect* (i.e. process and actual state of situational awareness). It supports an overall continuous comprehension of the situation, leading to better decisions (Salas *et al.*, 1995). This implicitly leads to increased performance regarding those decisions, but also in general.

Situational awareness has been researched mainly in the aviation sector. For example, Hartel, Smith and Prince (1991) investigated over 200 aviation mishaps and concluded that a lack of

situational awareness was one of the most important causal factors for errors. Similarly, Endsley (1988) in fact stated that situational awareness was *the* most important aviation performance factor. Besides its role in aviation, Salas *et al.* (1995) propose that it has an important role for other types of teams operating in a dynamic and complex environment in which errors lead to serious consequences. This is similar to the environment a paramedic team operates in. With increased situational awareness, paramedic teams facing a bystander conflict are better able to assess a situation, project possible cause-and-effect scenarios, and minimize escalation and medical errors. This is expected because of better decision making and overall being more knowledgeable about the situation.

2.5 Training

The report '*To err is human: building a safer health system*' by the Institute of Medicine (IOM) (Kohn, Corrigan & Donaldson, 1999) was the initiation of a rapid increase in attention and development of training methodologies applicable to the healthcare sector. The report's main premise was quite simple and as old as ancient philosophy[6]: people will always make mistakes, but persevering in making them is the actual problem. Based on the IOM report, more focus was put on how to prevent these mistakes from occurring repeatedly, thereby increasing medical performance. Kohn *et al.* (1999) suggest that the critical development for teams is the development of the team's teamwork behaviour (e.g., Cannon-Bowers, Tannenbaum, Salas & Volpe, 1995; Stout, Salas & Fowlkes, 1997). The teamwork behaviours described in Paragraph 2.4.1 to 2.4.9 can be classified as non-technical skills. The easiest and most efficient way to develop non-technical skills is with the implementation of training programmes (e.g., McGaghie, Issenberg, Petrusa & Scalese, 2006; Rosen, Salas, Wilson, King, Salisbury, Augenstein, Robinson & Birnbach, 2008; Bowers, Salas & Rhodenizer, 1998). As such, it is being used increasingly for the development of non-technical skills in a medical team-setting (e.g., Salas, Burke, Bowers & Wilson, 2001; Shapiro, Morey, Small, Langford, Kaylor, Jagminas, Suner, Salisbury, Simon & Jay, 2004; Baker, Beaubien & Holtzman, 2003).

Original training methods training non-technical skills were developed for the military and aviation sector (Helmreich, Merritt & Willhelm, 1999). The most commonly used training method in these sectors is the Crew Resource Management (CRM) programme (or, sometimes *cockpit*, or *crisis*). This methodology has recently been applied to the healthcare setting, known as the Anaesthesia Crisis Resource Management (ACRM; Gaba, Howard, Fish, Smith & Sowb, 2001). A distinction is made between a crew (similar to a *discipline*, in healthcare) and a team, as teams can have several different crews. An integration of technical, cognitive and behavioural skills as part of a team's combined disciplinary team-domain is trained, to overcome troubling situations. Shapiro *et al.* (2004) report a trend of improvement in teamwork behaviour for teams working in the emergency department taking part in a simulation based ACRM-training, whereas the control group showed no change in teamwork behaviour. They indicate that it is a promosing training method for the healthcare sector.

Within the ACRM programme, there is a focus on increasing teamwork collaboration within multidisciplinary teams in the complex and dynamic environments the teams will be in. The focus on collaboration distinguishes between two key points: (1) *cognition and decision making*;

[6] *Errare humanum est, perseverare diabolicum* (Seneca, 4 BCE – 65 CE) – To err is human, to persevere is diabolic.

and (2) *teamwork behaviour and resource management*. The focus on cognition and decision making is to gain knowledge about the structure and problems of the environment (as it expected to be *ill-structured* and *full of uncertainty*). That is: (1) *knowing the environment*; (2) *anticipation and planning*; (3) *use of all available information and cross checking all used information;* (4) *preventing or managing errors*; and (5) *use of cognitive aids, whenever available*. The focus on *teamwork behaviour* and *resource management* is to shift between, and, complete roles and decrease time pressure. That is: (1) *effective leadership and followership*; (2) *asking for help, early-on*; (3) *effective communication*; (4) *distribution of workload*; and (5) *availability and use of necessary resources.* The goal related to these key points is minimizing the occurrence of errors, and, *if* an error occurs to minimize the consequences. The training programme that is used in the experimental design of this master thesis is based on the '*teamwork behaviour and resource management*'-aspect of the ACRM programme. Its form is described in Appendix A. Besides focusing on improving teamwork behaviour and resource management, it also focuses on how to deal directly with the specific situation of a bystander conflict. However, conflict management is outside the scope of this master thesis. The reader is referred to the external sources used for the training programme, as described in Appendix A.

Besides the development of teamwork behaviours, it is expected that training also has a direct effect on team performance. Salas, Diaz-Granados, Klein, Burke, Stagle, Goodwin and Halpin (2008) meta-analytically investigated the effect of team training on team outcomes, consisting of team functioning and team performance as a measure. They describe that there is a moderate, positive effect of team training on team outcomes. Combining the expected effect training will have on the development of teamwork behaviour and on team performance, the following research model is suggested (Figure 2.4):

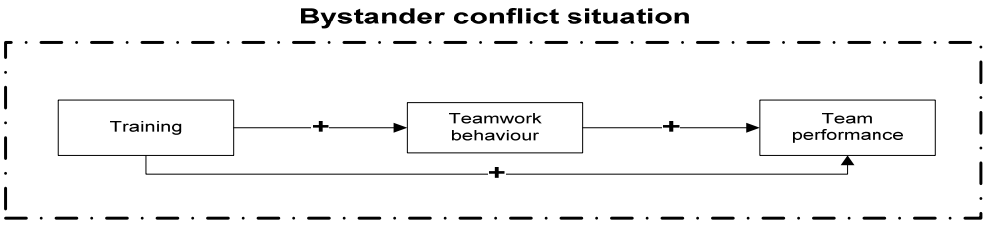

Figure 2.4: Research model

Teamwork behaviour has a mediating role in the research model. It can be argued that even though teamwork behaviour is a necessary prerequisite for sufficient team performance in the complex situation the paramedic teams are in, it can be an additional demand as well when the teamwork behaviour is not *managed* behaviour for the team members. It is suggested that training is necessary to make it less salient as a demand and more usable as an actual manageable resource the team members can use.

In Paragraph 2.4.1 to 2.4.9, I described the nine teamwork behaviours that are most prominently mentioned in scientific literature. However, not all of them are used in the training programme of this master thesis. Shared mental models (*cognitive*), mutual trust (*affective*), and team orientation (*attitudinal*) are considered to be *black box constructs* (Wickens, 1984). Black box constructs are within-person processes and are therefore more difficult to train and observe. As such, they will be kept out of the training programme.

Furthermore, adaptability and team leadership will be combined into one teamwork behaviour: *adaptive leadership*. Similar to shared mental models, mutual trust and team orientation, adaptability is hard to observe as teamwork behaviour. Combining it with leadership makes it an observable teamwork behaviour via the expression of *adaptive delegation of leadership*. Given the rapidly changing environmental characteristics of a paramedic team experiencing a bystander conflict, adaptive delegation of leadership will be necessary to adequately make use of information and to coordinate behaviour. Via the coordination of behaviour the situation can be stabilized (*emergency phase*) whereas making use of information can identify and clarify the stabilized situation (*adaptive phase*) (Heifetz, Grashow & Linskey, 2009). This is similar to what was defined as adaptability (Paragraph 2.4.4), as it suggests a clear understanding of the team's task, but also the possibility and recognition of changes in roles via leadership, to accomplish the team's task (Salas *et al.,* 2005).

2.6 Hypotheses

Based on the goal of the master thesis (Paragraph 1.2) and the theoretical framework (Paragraph 2.1 to 2.5), the following hypotheses are stated (hypothesis 1 to 11) to evaluate the theoretical framework (Figure 2.4, p.17):

Hypothesis 1: Training is positively related to the observed amount of situational awareness.

Hypothesis 2: Training is positively related to the observed amount of closed-loop communication.

Hypothesis 3: Training is positively related to the observed amount of adaptive leadership.

Hypothesis 4: Training is positively related to the observed amount of mutual performance monitoring.

Hypothesis 5: Training is positively related to the observed amount of backup behaviour.

Hypothesis 6: Training is positively related to team performance.

Hypothesis 7: The effect of training on team performance is mediated by situational awareness.

Hypothesis 8: The effect of training on team performance is mediated by closed-loop communication.

Hypothesis 9: The effect of training on team performance is mediated by adaptive leadership.

Hypothesis 10: The effect of training on team performance is mediated by mutual performance monitoring.

Hypothesis 11: The effect of training on team performance is mediated by backup behaviour.

3 Methodology

In this chapter, I will describe the experimental design, and, the main methods of research of this master thesis.

First of all, in Paragraph 3.1 a description is given of the experimental design. Secondly, as mentioned in the introduction, within this thesis there are three sub-goals: (1) *create an observational tool usable for coding teamwork behaviours in the context of the experimental design the participants take part in;* (2) *assess teamwork behaviour observed in the context of the experimental design with the observational tool, to evaluate the research model*; and (3) *assess teamwork behaviour and team performance in the context of the experimental design with the survey, to evaluate the research model.* In Paragraph 3.2, I will describe the methodology that is used for the first and second sub-goal. In Paragraph 3.3, I will describe the methodology for the survey that is used for the third sub-goal.

3.1 Experimental design

The experimental design is a randomized simulation-based pre-test post-test control group design (Figure 3.1)

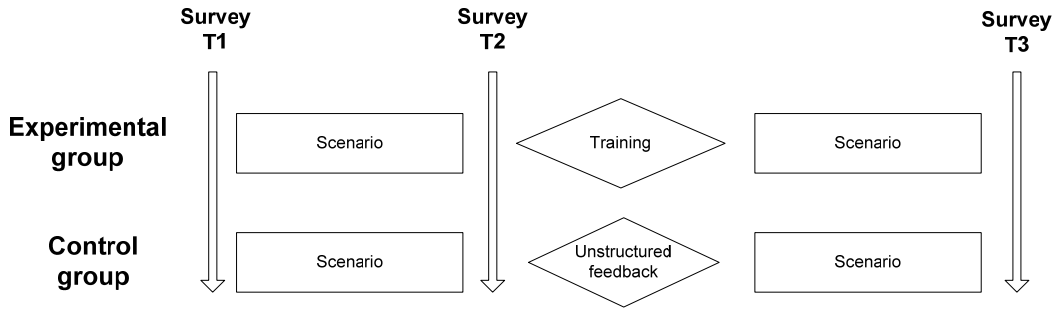

Figure 3.1: experimental design

A total of 119 participants took part in the experimental design, divided over 61 teams (37 teams in the experimental group, 24 teams in the control group). Of these 61 teams, three teams had a team size of three participants (all three teams were in the control group), the other 58 teams had a team size of two participants. Six participants took part in two teams.

The 61 teams were divided over nine days of which each day had two sessions. Each session had three or four teams, of which each team took part in two scenarios (the *heroin* and *diabetes* scenario; five minutes per team, for a total of ten minutes per scenario). All the scenarios the participants took part in were videotaped.

If the session was part of the experimental group, the participants took part in a 1.5 hour structured feedback and training session between the two scenarios. The control group only got an unstructured feedback session between the two scenarios. The time-schedule of a session and a description of the two scenarios is given in Appendix A.

3.2 Observational tool

In this paragraph, a description is given of the design process for the observational tool. The considerations for why observation is relevant are discussed in Paragraph 3.2.1. A summary of the already available observational tools is given in Paragraph 3.2.2 . In Paragraph 3.2.3 a description is given of the theoretical foundation on which the newly designed observational tool is based. The actual design of the observational tool is described in Paragraph 3.2.4. Lastly, in Paragraph 3.2.5 a description is given of the types of analysis that will be used to analyse the data acquired by the observational tool.

3.2.1 Considerations for teamwork behaviour

Measurement of teamwork behaviour is generally done by measuring retrospective *impressions* (i.e. self-reports via a survey), rather than observing the teamwork behaviour itself. This provides a *perception* of the teamwork behaviour. However, this perception does not necessarily have to be an accurate reflection of the *actual* teamwork behaviour (Weingart, 1997). While team member perception of the teamwork behaviour remains an important aspect, inaccurate recall (*measurement invalidity*) and subjective assessment (*inter-rater invalidity*) can introduce a measurement bias. Regarding inaccurate recall, teamwork behaviour could occur at a complexly aggregated level thereby possibly not being easily discernible by group members given the complexity of the situation they are in. Or differently, some teamwork behaviours could occur subconsciously. Regarding subjective assessment, past experience, knowledge about the task and performance outcome, and individual differences can influence retrospective impressions of teamwork behaviour (Staw, 1975). Besides the possibility of bias, it suggests that teamwork behaviour is static and measurable by a snapshot approach rather than acknowledging that certain teamwork behaviours can evolve over a given timeframe. Direct observation enables the possibility of a more dynamic and continuous insight into the possible mediation of teamwork on the transformation of input into output variables (Hackman & Morris, 1975; Goodman, Ravlin, & Schminke, 1987).

Direct observation has some practical and methodological issues to be answered as well. The first decision to be made is whether to directly observe, or, to observe video data in hindsight. Observing recorded video data is always preferable to directly observing in real time (Jeffcott & Mackenzie, 2008), *if* the nature of the situation allows it. This has several reasons: (1) coding will be based on the "*coder's rather than the group's pace*" (Weingart, 1997, p.212); (2) real time observation requires an extremely experienced coder; and (3) it enables observation by multiple coders more easily. The recording of data, however, does have some problems as well, even though most of them are mentioned in more dated literature and are becoming less relevant due to technological development and access to better equipment. Still, due to the fixed angle of the camera and microphones, it will not always be able to clearly register everything. Arguably though, this problem is similar for real-time observation if the observer does not want to disturb the process (e.g., by moving around).

Furthermore, observation is a very time consuming process. Both in the creation of the observational tool (i.e. the *coding scheme*) as well as the analysis of the data with it. For example, Weingart, Hyder and Prietula (1996) describe how a coding scheme containing 'only' 14 categories required each coder 60 hours of training to achieve a significantly high enough inter-rater reliability when analysing their data. Training is necessary, both on a theoretical as

well as on an hypotheses-level because coders naïve to these aspects of the coding scheme will not hold the same understanding as the researcher developing it. For the actual creation of the coding scheme, Baker and Salas (1992) state that it is essential that the coding scheme is mainly based on theory, especially when measuring teamwork behaviours. Weingart (1997) suggests a more hybrid approach because it is necessary to be able to decide which behaviours should be measured, and how this should be done. They suggest to first of all check whether behaviour is actually observable because the operating circumstances in which the behaviour takes place needs to be clear before a measurement tool based on theory can be designed. This is especially relevant in a situation in which a lot of different situational-evolving specific behaviours can occur while there are only a few key behaviours identified from literature. Because of the expected situational-evolving aspects of scenario-based training while interacting with an actor, this approach will also be followed for the creation of the coding scheme in this master thesis.

3.2.2 Available measures

There are several available measurement tools designed specifically for the healthcare setting, that can be investigated for inspiration. Most of them measure non-technical skills and focus on cooperation, coordination, communication, awareness and leadership. This is in line with the Big Five's model of teamwork (Paragraph 2.4). However, these readily available tools are not suitable for the context of this master thesis because they either specify a too broad of a range of constructs, or, use too job-specific underlying behaviours for clarification. For example, the Observational Teamwork Assessment for Surgery (OTAS) separates behavioural constructs (e.g. *leadership*) clarified by exemplar behaviours (e.g. *'Advises Anaesthetist if unfamiliar with operative techniques (e.g. tube insertion) to call for senior help'*) over multiple phases (pre-, intra-, and post-operational). The constructs then have to be rated on a Likert scale given a certain anchor definition (Hull *et al.,* 2010). Similar drawbacks apply to other healthcare related measurement tools, for example, the Operating Room Management Attitudes Questionnaire (ORMAQ; Sexton, Helmreich, Glenn, Wilhelm &, Merritt 2000) and Anaesthetist' Non-Technical Skills (ANTS; Fletcher, Flin, McGeorge, Glavin, Maran & Patey, 2003). A more generally applicable measurement tool is the Mayo High Performance Teamwork Scale (MHPTS; Malec, Torsher, Dunn, Wiegmann, Arnold, Brown & Phatak, 2007). However, it has only been validated on *teamwork* as its only dimension. This is a disadvantage since the goal of the coding scheme in this master thesis is to measure different teamwork behaviours. As such, when investigating teamwork behaviours in a new training setting, Schraagen, Schouten, Smit, Haas, van der Beek, van de Ven, & Barach (2010) suggest to use the Big Five model as a general applicable theoretical foundation, extended by training specific behavioural markers.

3.2.3 Theoretical foundation

The teamwork behaviours that are used in the research model, will be incorporated in the observational tool (i.e. *situational awareness, closed-loop communication, adaptive leadership, backup behaviour* and *mutual performance monitoring*). The definitions used for the coding scheme are given in Table 3.1 to 3.5. Besides their definition, *decision rules* (i.e. describing how to map out an observation to the appropriate action) and *behavioural markers* (i.e. observable behaviours structured into a set of categories) are added. The definitions, decision rules, and

behavioural markers are based upon research by Salas, Sims and Burke (2005), Salas *et al.* (1995), and Malec *et al.* (2007).

Table 3.1: Situational awareness

Situational awareness	
Definition	*The perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future.*
Decision rule	*Situational awareness behaviour is present whenever team members have a dynamic theory of the situation, which is tested and endorsed/changed by collecting and sharing information.*
Behavioural markers	▪ Team members seek and communicate information, from both the internal and external environment.

Table 3.2: Closed-loop communication

Closed-loop communication	
Definition	*The exchange of information between a sender and a receiver, irrespective of the medium.*
Decision rule	*Closed-loop communication behaviour is present when necessary information is not only distributed to other team members, but there is a continuous (verbal or non-verbal) involvement between team members whether they receive, understand, and/or agree with the information.*
Behavioural markers	▪ Verbalizing task-related activities aloud; ▪ Following up with team members to ensure a message was received; ▪ Acknowledging, either verbally or non-verbally, that a message is received; ▪ Clarifying with the sender of the message that the message received is the same as the intended message; ▪ Repeating back or paraphrasing instructions and clarifications to indicate information was heard correctly.

Table 3.3: Adaptive leadership

Adaptive leadership	
Definition	*Ability to adjust, direct and coordinate strategies based on information gathered from the environment. Altering the course of action (internal or external) in response to changing conditions.*
Decision rule	*Adaptive leadership behaviour is present when a team member interacts dynamically, interdependently, and adaptively toward a common and valued goal whenever inputs from the environment change. Each member can have a specific role to fulfil within the team, directed internally or externally towards the team's tasks or social domains.*
Behavioural markers	▪ Facilitating and synchronizing team member contributions usable for problem solving ▪ Clarifying team member roles; ▪ Seeking and evaluating information that affects team functioning; ▪ Identifying cues that a change has occurred, assigning meaning to that change, and developing a new plan to deal with the changes; ▪ Remaining vigilant to changes in the internal and external environment of the team.

Table 3.4: Backup behaviour

Backup behaviour	
Definition	*Ability to anticipate other team members' needs through accurate knowledge about their responsibilities. This includes the ability to shift workload among members to achieve balance during high periods of workload or pressure*
Decision rule	*Backup behaviour is present when a team member is unable to sufficiently perform duties and another member addresses the necessary action via verbal or non-verbal input to ensure the duties can be completed.*
Behavioural markers	▪ Shifting of work responsibilities to underutilized team members; ▪ When appropriate, roles are shifted to address urgent or emergent events; ▪ Completion of the task or parts of tasks by other team members; ▪ Team members ask each other for assistance prior to or during periods of task overload.

Table 3.5: Mutual performance monitoring

Mutual performance monitoring	
Definition	*The ability to develop a common understanding of the team environment and apply appropriate task strategies to accurately monitor teammate performance.*
Decision rule	*Mutual performance monitoring behaviour is present whenever team members actively (verbal or non-verbal) maintain awareness of other team member's behaviour while carrying out their own actions. Whenever necessary, they identify and correct emerging lapses to ensure processes run as expected.*
Behavioural markers	▪ Identifying mistakes and lapses in other team member's actions; ▪ Team members call attention to actions that they feel could cause errors or complications; ▪ Providing feedback regarding other team member's actions to facilitate self-correction; ▪ Team members acknowledge –in a positive manner- statements directed at avoiding or containing errors or seeking clarification.

3.2.4 Coding scheme development

The teamwork behaviours as defined in Paragraph 3.2.3 are defined at a macro-level. The definitions are extensive. It can therefore be difficult to observe and code them. For example, backup behaviour as a single type of behaviour can be present in more than one way. It is therefore required that the coding scheme codes behaviours on a *micro-level of measurement*. As such, the actual coding, i.e. its *units of measurement*, can be regarded as single "*utterances or actions that fit into one category"* (Kolbe, Künzle, Zala-Mezö, Wacker & Grote., p. 210). The developed coding scheme (Figure 3.2) is based on research by Kolbe *et al.* (2009).

Kolbe *et al.* (2009) emphasize the importance of adaptive coordination, defined as the "*structured patterning of within-group activities by which groups strive to achieve their goals"* (Arrow *et al.*, 2000, p. 104). Adaptive coordination is necessary to overcome failures in healthcare, given that the main source of failure is the lack of quality of teamwork due to changing circumstances (Weinger & Slagle, 2002). It is especially important whenever the context of the task is characterised by non-routine events (*NRE*), described as events being *out-of-the-ordinary* or simply *unusual* (Weinger & Slagle, 2002). Furthermore, adaptive coordination incorporates the principle of adaptability as described by Salas, Sims and Burke (2005). Kolbe *et al.* (2009, p.205) describe adaptability as a "*coordination process in and of itself rather than simply a prerequisite of coordination"*, and that different coordination mechanisms are appropriate to adapt to different events.

Based on the description of adaptive coordination by Kolbe *et al.* (2009), three types of coordination mechanisms are identified for the current research context: (1) *coordination of information*; (2) *coordination of action*; and (3) *coordination of interrelation*. For the first and

second coordination mechanisms, behaviours are subdivided into either being explicit (i.e. intentionally used verbal or written communicative behaviour), or implicit (i.e. describing shared cognitions and anticipations of actions and needs). Attitudes and behaviours of the team and its environmental operational state, and, how to act in alignment with respective requirements to achieve certain goals are described by the third coordination mechanism. It therefore goes beyond merely being implicit and explicit subdivision (Kolbe *et al.*, 2009).

The taxonomy of Kolbe *et al.* (2009) has over 40 behaviours. Even though developed specifically for anaesthesia teams, most of the behaviours are usable for the paramedics teams in the context of this master thesis. However, some of the behaviours are left out, for various reasons: (1) not being applicable (e.g., *talking to patient* or *technical alarm*); (2) not realistic (e.g. *chatting* or *teaching others*); (3) not relevant to code (e.g., *incomprehensible communication*); (4) redundancy (e.g., *approval* incorporated within *acknowledgment*); and (5) renaming and/or small changes (e.g., *non-verbal acknowledgment* rather than *listening*).

The behaviours incorporated in Figure 3.2 are defined in Table 3.6 and are based upon the original specifications by Kolbe *et al.* (2009). However, they are adjusted for the specific situation of paramedic teams. Each behaviour is clarified with an example.

Table 3.6: Definitions and examples of coding scheme behaviours

Behaviour	Definition	Example
Request for information	A direct question. Addressed to team members, patient, or the bystander.	'Does he/she use any medication that you know of?'
Providing information upon request	The answering of a direct question on request.	'The oxygen cylinder is on your left'
Verifying information	Repeating information, or, giving a verbal confirmation regarding fulfilled actions.	'His/her blood sugar level is 4.0'
Questioning information	Expressing doubt about the accuracy of information.	'Are you sure about his/her blood sugar level?'
Verbal acknowledgment (of receipt, or, understanding)	Short, verbal confirmation that the information was heard, understood, or, that team members are in agreement.	'Okay'

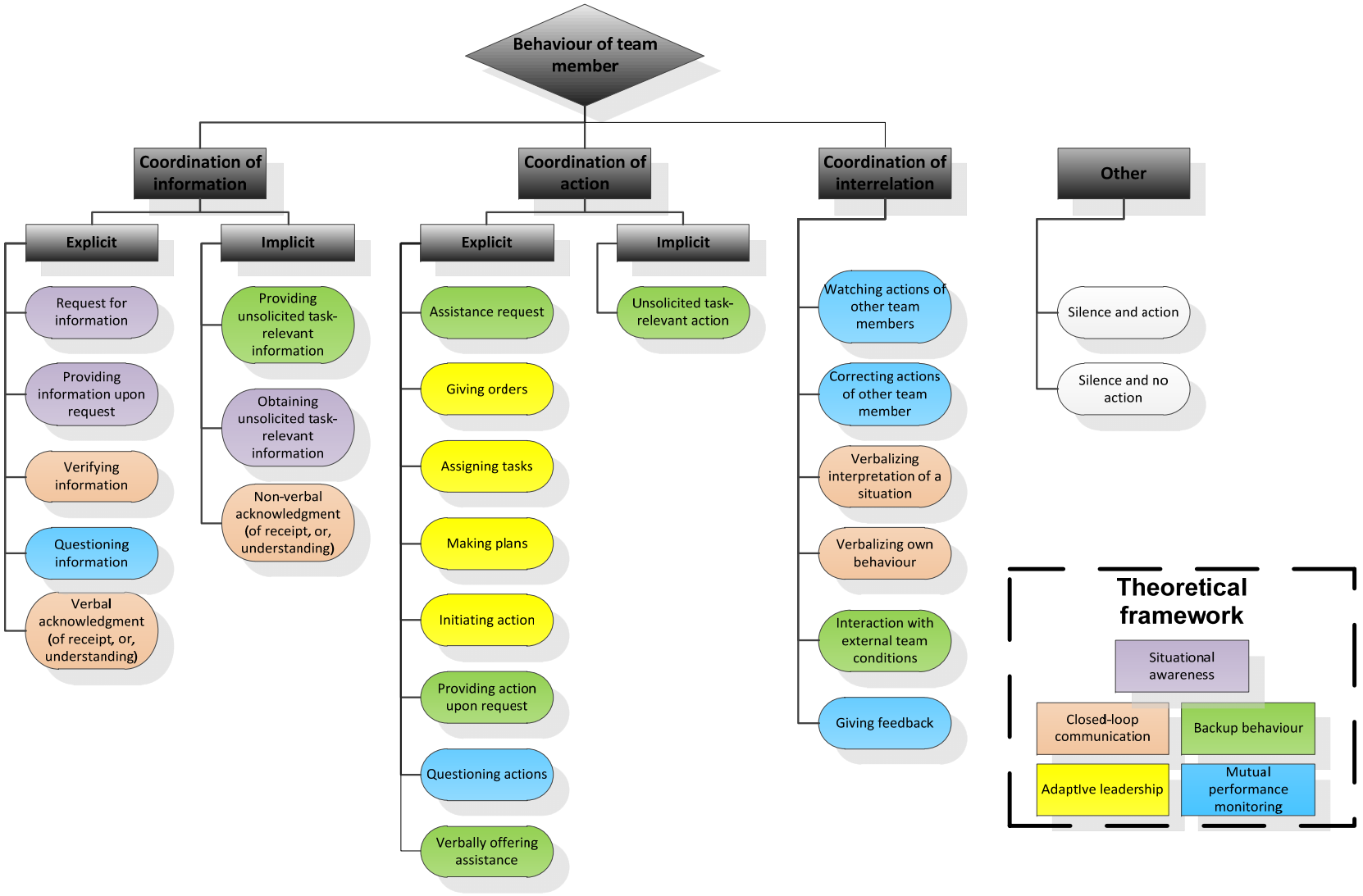

Figure 3.2: Coding scheme for teamwork behaviour

Table 3.6: Continued

Category	Definition	Example
Providing unsolicited task-relevant information	Providing information without being asked to do so.	'Breathing is normal'
Obtaining unsolicited task-relevant information	Actively gathering information without being asked to do so.	*Reading baseline on a monitor*
Non-verbal acknowledgment (of receipt, or, understanding)	Non-verbal confirmation indicating confirmation that the information was heard, understood, or, that team members are in agreement.	*Nodding*
Assistance request	Explicitly asking for help.	'Can you hand me the tube for intubation?'
Giving orders	Directives, commands, or instructions.	'You should listen very carefully now!'
Assigning tasks	Describing the division of subtasks among team members.	'I'll intubate the patient, can you watch his/her heart rate?'
Making plans	Verbalization of non-immediate considerations.	'When we are done intubating, the ambulance will take him'
Initiating action	Statements which initiate actions.	'We should give him 10cc glucose'
Providing action on request	Behaviour that is performed because it is asked to do so.	*Starting intubation whenever asked to do so*
Questioning actions	Expressing doubt about the accuracy of decisions, orders, or actions.	'Are you sure he/she needs intubation?'
Unsolicited task-relevant action	Task-relevant actions, without being asked to do so.	*Routinely intubating the patient*
Verbally offering assistance	Verbally offering help.	'Do you need any help handling that?'
Watching actions of other team member	A team member is observing actions of another team member.	*Team member is monitoring other team member's actions and his/her close environment*

Table 3.6: Continued

Category	Definition	Example
Correcting behaviour of other team members	Actions or statements aimed at correcting behaviour of another team member.	'No, the electrode should be plugged in on the other side'
Verbalizing interpretation of a situation	Verbal assessment of a team member's current situation.	'We are paramedics and obliged to help someone in need'
Verbalizing own behaviour	Verbally communicating personal task-related actions.	'I am requesting a second car'
Interaction with external team conditions	Considerations of conditions outside the team, and their possible consequences,	*Interactive behaviour (e.g. talking, walking, or continuous eye contact) with the bystander*
Giving feedback	Friendly reassurances or, less-than-sincere comments.	'You are helping us nicely by doing that!'
Silence and action	Team members work silently, and, independently.	-
Silence and no actions	Team member is doing nothing at all.	-

3.2.5 Analysis

To take the dynamic aspect of teamwork behaviour into account, each five minute scenario is divided into time slots. This has two advantages. First of all, it takes process development into account (i.e. the scenario having a dynamic nature rather than being static). This is relevant because of the expected maturation process both in the interaction with the bystander as well as within the team because the teams are not necessarily in a standard configuration. Secondly, it enables aggregation over time. Thereby acknowledging temporal effects and making it possible to include low-frequency events without them biasing the broader macro-level whenever they occur. The choice for the length of the time slots is an arbitrary choice because there is no reference point as to what is a justified choice. Nonetheless, time slots of 20 seconds are chosen as they should prevent the over-coding of behaviour (compared to a shorter time slot) and still takes the effect of prolonging behaviour into account over multiple timeslots (compared to a longer time slot) (Weingart, 1997). As such, each five minute scenario is divided into 15 time slots of 20 seconds. Both scenarios each team takes part in will be coded. The first scenario each team takes part in will be used as the pre-test score of each variable in the analysis, the second scenario as the post-test score of each variable.

The analyses are divided into exploratory analyses and confirmatory analyses. First of all, the exploratory analyses are used to gain initial insight into the analytical abilities of the

observational tool. For example, a graphical analysis at a time-slot level is used to investigate whether the choice for the 20 second time-slots seems justified. Furthermore, it is interesting to analyse if there are changes (percentage-wise) at the micro-level of behaviour, between groups and between scenarios. Only looking at changing totals of observed macro-level behaviour ignores information of possible changes on a micro-level (i.e. more behaviour of a certain macro-level behaviour does not necessarily mean an expected effective type of micro-level behaviour occurs more in a certain situation).

Secondly, results of the confirmatory analyses are used to assess whether there is support for an important sub-part of the research model (hypothesis 1 to 5, Paragraph 2.6); the expected effect team training has on the amount of observed teamwork behaviours.

3.3 Survey

In this paragraph, a description of the variables measured by the survey will be given, and of their methods of analysis. Participants were asked questions about variables of: (1) *teamwork behaviour*, including closed-loop communication, adaptive leadership, backup behaviour and mutual performance monitoring; and (2) *team performance*. In Paragraph 3.3.1 a description is given of the scale measures used for each of the variables. In Paragraph 3.3.2. a description is given of their construct validity, internal consistency, and, interrater reliability and agreement.

3.3.1 Scale measures

All variables were measured at an individual level. The total list of items for each variable is given in Appendix B.

Teamwork behaviour variables

Closed-loop communication. Closed-loop communication was measured using a 5-point Likert scale (1 = *totally does not apply* to 5 = *does apply strongly*), with a total of three items An example of a closed-loop communication item is: "*in this team, we verified communicated information"*. Items used are based on the scale used by Gevers *et al.* (2010).

Backup behaviour. Backup behaviour was measured using a 5-point Likert scale (1 = *totally does not apply* to 5 = *does apply strongly*), with a total of three items. An example of a backup behaviour item is: "*in this team, we anticipated on things the other person(s) needed"*. Items used were based on the scale used by Gevers *et al.* (2010).

Mutual performance monitoring. Mutual performance monitoring was measured using a 5-point Likert scale (1 = *totally does not apply* to 5 = *does apply strongly*), with a total of three items. An example of a mutual performance monitoring item is: "*in this team, we attended to impending mistakes and complications"*. Items used are based on the scale used by Gevers *et al.* (2010).

Adaptive leadership. Adaptive leadership was measured using a 5-point Likert scale (1 = *totally does not apply* to 5 = *does apply strongly*), with a total of six items. An example of an adaptive leadership item is: "*in this team, we adapted whenever it was required, given the situation*". Items used are based on the scale used by Gevers *et al.* (2010).

Team performance variable

Team performance. Team performance was measured using a 7-point Likert scale (1 = *really bad* to 7 = *really good*), with a total of five items. An example of a team performance item is: "*in this team, the quality of care given was …"*. Items used are based on the scale used by Gevers *et al.* (2010).

3.3.2 Internal consistency, construct validity, and, interrater reliability and agreement results

Construct validity

To test the construct validity of the scale measures used (i.e. whether items of a certain scale in fact measure the same construct as intended), Principle Component Analyses (PCA) with Direct Oblimin rotation was used. The results of the PCA are given in Appendix C.

For the different teamwork behaviours, no clear factor pattern exists for both extraction based on a pre-determined number of factors as well as based on the Kaiser criterion (Kaiser, 1960). Based on the Scree-plot's point of inflexion, it appears that teamwork behaviour consists of only one factor. However, since the PCA is only a First Order Factor Analysis for exploratory purposes and shows no consistent clear pattern, based on the principle of their *theoretical meaningfulness*, no items are removed. Instead, all the different types of teamwork behaviour variables are combined as one additional *teamwork behaviour variable* that is used in the analyses. For team performance, the PCA does not raise any concerns to use it as a single variable.

Internal consistency

Because the variables of teamwork behaviour and team performance consisted of multiple items measuring the same variable, Cronbach's α's were computed at each moment in time participants filled in the survey (i.e. T1, T2, and T3) as an indication for the internal consistency of the scales. Cronbach's α is a lower-bound of the estimated reliability. Even though some researchers (e.g., Hair, Black, Babin & Anderson, 2009) suggests that $\alpha \geq .8$ is good and $.60 \leq \alpha \leq .80$ is reasonable, for the social sciences $\alpha \geq .70$ is considered good. Results are displayed in Table 3.7.

Even though the Cronbach's α of backup behaviour at T1 is $\leq .6$, all of the averages are considered reasonable (two out of six) or good (four out of six) values. As such, there is no reason for concern of their internal consistency.

Table 3.7: Internal consistency (Cronbach's α)

Variable	*Items*	α (*T1)*	α (*T2)*	α (*T3)*	α (*Average)*
Closed-loop communication	3	.78	.82	.88	.83
Adaptive leadership	6	.87	.90	.87	.88
Backup behaviour	3	.55	.62	.77	.64
Mutual performance monitoring	3	.65	.66	.73	.68
Teamwork behaviour	15	.91	.92	.93	.92
Team performance	5	.92	.94	.95	.94

Interrater reliability and agreement

The variables of teamwork behaviour and team performance will be analysed at a team level, but they were measured at an individual level. It should be investigated whether there is sufficient conformity between individuals within the team to justify their aggregation on a team-level. Several measures for comparison were computed[7], as there is sufficient debate in literature which one (and how many) should be used (Bartko, 1976): (1) *interrater reliability* (intra-class correlation coefficients ICC1 and ICC2); and (2) *interrater agreement* (within-group interrater agreement (r_{wg})).

The intra-class correlation coefficients are a description of relative consistency by assessing correlations. ICC1 is used to evaluate whether there is sufficient within-team homogeneity to justify aggregating the individual scores at a team-level. Values of ICC1 range from -1 to 1, and as the value of ICC1 becomes larger, it is acceptable to assume that the rating of an individual will provide an reliable estimate of the team's mean. Bliese (2000) states that values of ICC1 $>.12$ are acceptable to make this assumption. Secondly, ICC2 as an absolute value is an estimation of the reliability of the team means. ICC2 is related to ICC1 as a function of team size (Bliese, 1998). When considering ICC2, teams with many people will provide a more stable estimate than teams with fewer people (Klein & Kozlowski, 2000). In general, Bliese (2000) states that values of ICC2 $>.50$ are acceptable. Taking into account the team sizes of two of the paramedic teams means that a sufficient ICC1 of .12 will only provide an ICC2 of .19[8]. Or the other way around, to obtain an ICC2 of $>.50$, an ICC1 value of approximately .33 is required. The results of the ICC2 should therefore be interpreted with caution

Lastly, r_{wg} is the degree to which ratings from individuals are interchangeable (Kozlowski & Hattrup, 1992) thereby being a measure of absolute consensus. James *et al.* (1984) suggest using the uniform distribution for r_{wg} and that values $>.70$ are acceptable. Results for ICC1, ICC2 and r_{wg} are given in Table 3.8.

[7] Using the excel-tool to accompany: Biemann, T., Cole, M.S., & Voelpel, S. (2012). Within-group agreement: on the use (and misuse) of rwg and rwg(j) in leadership research and some best practice guidelines. *The Leadership Quarterly*, 23, 66-80.

[8] $\text{ICC(2)} = \frac{k(\text{ICC(1)})}{1 + (k-1)\text{ICC(1)}}$, in which k is the team size (Bliese, 2000, p.357).

Table 3.8: Interrater reliability (ICC1 and ICC2) and interrater agreement (r_{wg} uniform (Standard Deviation))

Variable	ICC1	ICC2	r_{wg} uniform (SD)
Closed-loop communication T2	.13	.23	.75 (.34)
Closed-loop communication T3	.28	.44	.85 (.21)
Backup behaviour T2	.16	.27	.78 (.26)
Backup behaviour T3	.36	.52	.89 (.17)
Adaptive leadership T2	.04	.08	.81 (.25)
Adaptive leadership T3	.20	.34	.87 (.21)
Mutual performance monitoring T2	.34	.50	.83 (.21)
Mutual performance monitoring T3	.23	.38	.84 (.22)
Teamwork behaviour T2	.19	.32	.84 (.21)
Teamwork behaviour T3	.31	.47	.90 (.17)
Team performance T2	.29	.44	.86 (.20)
Team performance T3	.19	.32	.86 (.21)

ICC2 is $\leq .50$ for most variables. This indicates a low reliability of the between-team means. As mentioned earlier, the ICC2 values should be interpreted with caution because of the small team sizes of the paramedic teams. Therefore, because all r_{wg} values are $> .70$, and all but one ICC1 values are $> .12$, aggregation is justified.

4 Results

In this chapter, the results of the findings based on the data of the experimental design data are presented. In Paragraph 4.1 the results based on the data acquired by the observational tool are given. In Paragraph 4.2 the results based on the data acquired by the individual variables from the survey are given.

4.1 Results of the observational tool

For the observational tool, a distinction was made between exploratory analyses and confirmatory analyses.

Of the 61 teams that took part in the experimental design, the video data of 53 teams will be analysed of which 23 teams were part of the control group and 30 teams of the experimental group. The teams that will not be used in the analyses were part of the session that included the three teams with a team size of three (instead of two). These teams were used to test the tool in the early stages of analysis. Being a learning method, this was used in an attempt to minimize miscoding of the other sessions. Both scenarios each team took part in were analysed. The first scenario each team takes part in will be used as the pre-test score of each variable in the analysis, the second scenario as the post-test score of each variable.

Of the 1590 time slots (i.e. 53 (*teams*) x 15 (*number of slots per scenario*) x 2 (*number or scenarios per group*)), 2.6% is missing (i.e. 42 time slots, distributed over 10 teams). The missing data is not missing at random as only the last time slots (starting from the last one of the scenario) are missing. The missing data can have two reasons, either the scenario (1) *escalated*, i.e. the scenario ends abruptly due to escalating circumstances; or (2) had a *measurement error*, i.e. the scenario is stopped too soon by the scenario leader. From the ten teams missing data, three of them were due to scenario escalation (accounting for 43% of the missing time slots, with a minimum of four and a maximum of nine missing time slots per escalated scenario). However, the escalated scenarios are divided among conditions. Both reasons for missing data are not a reason for concern whether or not the groups should be considered an outlier. Escalation was always initialized by a single remark not specifically related to behaviour in the time slots before, by either one of the team members or the bystander (e.g. the bystander stating that a female team member "*acted like an arrogant slut"*). Whereas measurement error was fairly consistent (i.e. scenarios consistently stopped after four minutes, not related to the actual state of the scenario). Visual inspection does not suggest otherwise. Because the teams with missing data would be 'missing' behaviours whenever the behaviours per time slot are aggregated into a total number of behaviours per team per scenario, the mean value of each behaviour each time slot (taking into account the scenario-specifics) is added via data imputation.

4.1.1 Exploratory analyses

In Paragraph 3.2.5 a description was given of two types of analyses: (1) whether the choice for the arbitrarily chosen 20 second time slots seems justified; and (2) if there are percentage-wise

differences between micro-level behaviours (taking into consideration differences between group, and scenario).

20 second time slots

To analyse whether the 20 second time slots appears to be a justified choice, it can be analysed visually (Figure 4.1) and analytically (Table 4.1). Even though the analytical results (i.e. Skewness, Kurtosis and Kolmogorov-Smirnov) of Table 4.1 are all significant, suggesting a deviation from normality, the significance of the deviation seems to be an artefact of the large sample size (n = 1548). Visual inspection of the frequency distribution (Figure 4.1) is an indication that the 20 second time interval is an appropriate time slot for observing the relevant teamwork behaviours.

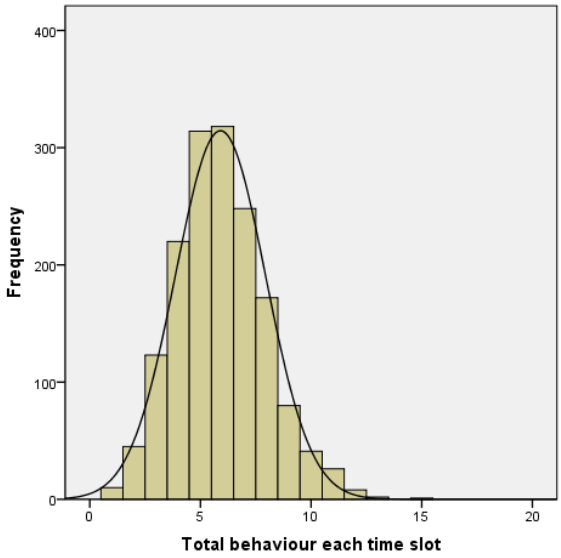

Figure 4.1: Frequency distribution of the 'total number of behaviours per time slot' (n = 1548)

Table 4.1: Descriptive statistics for the 'total number of behaviours per time slot' (n = 1548, Mean (M), standard deviation (SD), Kurtosis (K), Skewness (S), and Standard Error (SE)).

Statistic		Z-score
M	5,9	
SD	2.04	
S	0.384	6.295**
SE S	0.061	
K	0.285	2.336*
SE K	0.122	
K-S		
0.122**		

*Significant at p<.05, ** Significant at p<.001, two-tailed.

Micro-level of observed behaviour

Inspecting single utterances of observed behaviour (i.e. on a micro-level) of the post-test (i.e. the second scenario each team took part in), has the benefit that it specifically targets changes in a certain utterance of behaviour (e.g., *giving orders*) rather than the macro-level category to which it belongs (i.e. *adaptive leadership*). This could identify if there are changes between observed micro-level behaviours on which specific focus was put during the training. But also if a certain situation (e.g., the differences between scenarios) can cause people to behave differently on a micro-level to perform better, something which the total amount of observed macro-level behaviour does not investigate.

Results of the percentage-wise changes of the behaviours at a micro-level between *experimental group* and *control group*, and, between the *heroin* scenario and *diabetes* scenario, are calculated. The cases with changes of $\geq 10\%$ are shown in Figure 4.2 and will be described below. All other between-group and between-scenario differences can be found in Appendix D.

Experimental group vs. control group

For *adaptive leadership*, the experimental group shows percentage-wise more *giving orders* behaviour for both the *diabetes* scenario ($\Delta +12.8\%$) and the *heroin* scenario ($\Delta +15.4\%$), and less *initiating action* behaviour for the *diabetes* scenario ($\Delta -10.7\%$).

For *backup behaviour*, the experimental group shows percentage-wise less *unsolicited task-relevant action* for both the *diabetes* scenario ($\Delta -13.3\%$) and the *heroin* scenario ($\Delta -10.4\%$).

Diabetes scenario vs. heroin scenario

For *adaptive leadership*, the *diabetes* scenario shows more *initiating action* for both the experimental group ($\Delta +13.2\%$) and the control group ($\Delta +15.9\%$), and, more *making plans* for the experimental group ($\Delta +10.3\%$). The *diabetes* scenario shows less *giving orders* for both the experimental group ($\Delta -25.1\%$) and the control group ($\Delta -22.5\%$).

For *backup behaviour*, the *diabetes* scenario shows less *unsolicited task-relevant action* for the experimental group ($\Delta -11.4\%$).

For *mutual performance monitoring*, the *diabetes* scenario shows more *giving feedback* for both the experimental group *(*$\Delta +21.3\%$) and the control group ($\Delta +20.5\%$), and less *watching the actions of other team members* for both the experimental group ($\Delta -24.7\%$) and the control group ($\Delta -18.8\%$).

Figure 4.2: Post-test pie-chart percentage distribution for micro-level behaviours (>10% changes). Comparison between experimental group vs. control group, and, diabetes scenario vs. heroin scenario.

4.1.2 Confirmatory analyses

The correlations between macro-level teamwork behaviours of the coding scheme are given for the experimental group (Table 4.2, above the diagonal, n = 30) and for the control group (Table 4.2, below the diagonal, n = 23). Descriptive statistics are given in Table 4.3, describing sample size, means and standard deviations. After that, the hypotheses are discussed separately.

Table 4.2: Correlations between post-test macro-level behaviours, aggregated at a team-level (n = 53). Experimental group above the diagonal, control group below.

Teamwork behaviour	1	2	3	4	5
1. Situational awareness	-	.31	-.14	-.04	.18
2. Closed-loop communication	.18	-	-.42*	-.05	.59**
3. Adaptive leadership	-.22	.54**	-	.04	-.58**
4. Backup behaviour	-.25	-.52*	-.23	-	-.14
5. Mutual performance monitoring	.36	.37	-.03	-.50*	-

*= Correlation is significant at p<.05, **= Correlation is significant at p<.01, two-tailed

For the experimental group, there are two negative correlations that are significant: between *adaptive leadership* and *closed-loop communication* ($r = -.42, p < .05$), and, between *adaptive leadership* and *mutual performance monitoring* ($r = -.58, p < .01$). This indicates that if more adaptive leadership behaviour was observed, less closed-loop communication and mutual performance monitoring behaviour was observed, and vice versa. There is also one significant positive correlation, between *closed-loop communication* and *mutual performance monitoring* ($r = .59, p < .01$). This indicates that if more closed-loop communication behaviour was observed, also more mutual performance monitoring behaviour was observed.

For the control group, there are two negative correlations that are significant: between *backup behaviour* and *closed-loop communication* ($r = -.52, p < .05$), and, between *backup behaviour* and *mutual performance monitoring* ($r = -.50, p < .05$). This indicates that if more backup behaviour was observed, less closed-loop communication and mutual performance monitoring behaviour was observed, and vice versa. There is also one significant positive correlation, between *adaptive leadership* and *closed-loop communication* ($r = .54, p < .01$). This indicates that if more adaptive leadership behaviour was observed, also more closed-loop communication behaviour was observed, and vice versa.

Method of analysis

Hypothesis 1 to 5 are tested by performing a two-way (*group*, *scenario*) ANCOVA for each of the hypotheses. Each ANCOVA has two levels for group (experimental, control) and two levels for scenario (heroin, diabetes), and includes the post-test score of the hypothesized teamwork behaviour variable as the dependent variable and the pre-test score of it as a covariate (as suggested by as suggested by Todman and Dugard (2007)).

Situational awareness

Hypothesis 1 stated a positive relation between *training* and the observed amount of *situational awareness*. No significant relation between *training* and *situational awareness* was found

`

Table 4.3: Descriptive statistics of observed post-test macro-level teamwork behaviours, at a team-level (sample size (n = 53), mean (M), standard deviation (SD))

Group	n	Teamwork behaviour*	M	SD	Scenario	n	Teamwork behaviour	M	SD
Control	23	Total	76.26	7.90	**Diabetes**	10	Total	80.00	6.99
		SA	13.04	4.86			SA	15.40	2.55
		CLC	16.43	3.94			CLC	18.50	3.50
		AL	10.13	3.60			AL	9.70	3.62
		BB	30.39	4.34			BB	28.10	4.38
		MPM	6.29	2.60			MPM	8.30	1.95
					Heroin	13	Total	73.38	7.57
							SA	11.23	5.49
							CLC	14.85	3.60
							AL	10.46	3.69
							BB	32.15	3.51
							MPM	4.69	1.84
Experimental	30	Total	98.59	10.24	**Diabetes**	16	Total	101.87	10.29
		SA	12.07	4.46			SA	13.68	4.54
		CLC	22.74	5.47			CLC	26.44	3.85
		AL	17.68	6.39			AL	13.69	5.51
		BB	40.08	4.16			BB	38.94	4.61
		MPM	6.02	4.79			MPM	9.13	4.29
					Heroin	14	Total	94.85	9.13
							SA	10.23	3.70
							CLC	18.52	3.68
							AL	22.24	3.75
							BB	41.39	3.26
							MPM	2.46	2.13

*SA = Situational awareness, CLC = Closed-loop communication, AL = Adaptive leadership, BB = Backup behaviour, MPM = Mutual performance monitoring

($M_{control} = 13.04, M_{experimental} = 12.07$; F(1,48) = 1.040, p=.313, η^2 = .021). As such, hypothesis 1 is not supported. No interaction effect was found between *group* and *scenario* (F(1,48) = .080, p=.779). Similarly, no significant effect was found for the pre-test score of *situational awareness* as a covariate (F(1,48) = .895, p=.349).

Closed-loop communication

Hypothesis 2 stated a positive relation between *training* and the observed amount of *closed-loop communication*. A positive relation between *training* and *closed-loop communication* was found ($M_{control} = 16.43, M_{experimental} = 22.74$; F(1,48) = 19.00, p < .01, η^2 = .284), indicating a large effect[9]. As such, hypothesis 2 is supported. No interaction effect was found between *group* and *scenario* (F(1,48) = 3.136, p=.083). Similarly, no significant effect was found for the pre-test score of *closed-loop communication* as a covariate (F(1,48) = 2.353, p=.132).

Adaptive leadership

Hypothesis 3 stated a positive relation between *training* and the observed amount of *adaptive leadership.* A positive relation between *training* and *adaptive leadership* was found ($M_{control} = 10.13$, $M_{experimental} = 17.68$; F(1,48) = 39.94, p < .01, η^2 = .454), indicating a large effect. As such, hypothesis 3 is supported. Furthermore, the interaction between *group* and *scenario* was found to be significant as well (F (1,48) = 10.59, p < .05). This interaction is depicted in Figure 4.3.

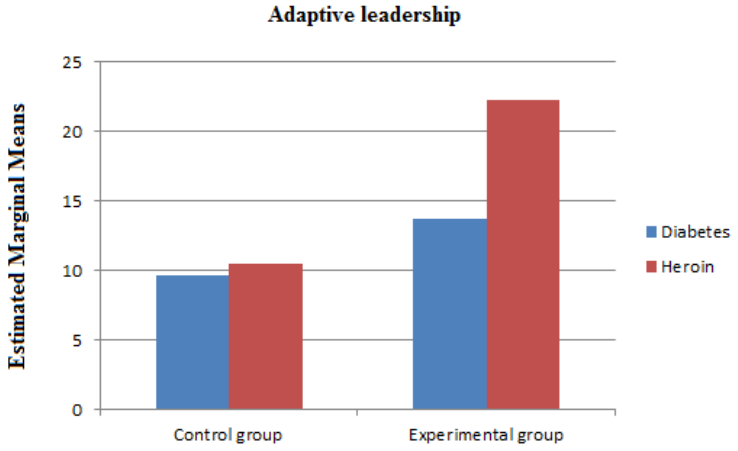

Figure 4.3: Interaction effect between group and scenario for Adaptive leadership

The simple slopes of the interaction effect were significant for both types of *scenario* ($t = 2.28, p < .05$ for the *diabetes* scenario and $t = 4.32, p < .001$ for the *heroin* scenario)[10]. Especially the effect of *training* on *adaptive leadership* was stronger for the heroin scenario

[9] Even though some authors argue that η^2 can be biased and its interpretation can be difficult (e.g., Field (2005)), Cohen (1988, p.283) gives an index indicating a small effect for η^2 > .0099, a medium effect for η^2 >.0588 and a large effect for η^2 >.1379,. Cohen's index will be used for the interpretation of the effect sizes.

[10] Simple intercepts, simple slopes, and regions of significance in MLR 2-Way interactions - http://quantpsy.org/interact/mlr2.htm , Preacher, Curran and Bauer (2012)

compared to the diabetes scenario ($\Delta_{control} = +.76$ whereas $\Delta_{experimental} = +8.55$). No significant effect was found for the pre-test score of *adaptive leadership* as a covariate (F(1,48) = .606, p=.440).

Backup behaviour
Hypothesis 4 stated a positive relation between *training* and the observed amount of *backup behaviour.* A positive relation between *training* and *backup behaviour* was found ($M_{control} = 30.39$, $M_{experimental} = 40.08$; F(1,48) = 22.47, p < .001, η^2 = .319), indicating a large effect. As such, hypothesis 4 is supported. No interaction effect was found between *group* and *scenario* (F(1,48) = 2.422, p=.126). The pre-test score for *backup behaviour* is a significant covariate in the model (F(1,48) = 4.125, p < .05). Thus, even after controlling for initial differences between the *experimental group* and *control group*, differences remain.

Mutual performance monitoring
Hypothesis 5 stated a positive relation between *training* and the observed amount of *mutual performance monitoring.* No significant relation between *training* and *mutual performance monitoring* was found ($M_{control} = 6.29$, $M_{experimental} = 6.02$; F(1,48) = .782, p =.381, η^2 = .016). As such, hypothesis 5 is not supported. No interaction effect was found between *group* and *scenario* (F(1,48) = 3.643, p=.062). Similarly, no significant effect was found for the pre-test score of *mutual performance monitoring* as a covariate (F(1,48) = .180, p=.674).

Total amount of observed behaviour
An alternative test for hypotheses 1 to 5 is testing whether there is a positive relation between *training* and the *total amount of observed behaviour* (i.e. the five macro-level teamwork behaviours combined, analysed in the same way as suggested by the *method of analysis*). A positive relation between *training* and *total amount of observed behaviour* was found ($M_{control} = 76.26$, $M_{experimental} = 98.59$; F(1,48) = 29.09, p < .001, η^2 = .377), indicating a large effect. No interaction effect was found between *group* and *scenario* ((F(1,48) = .031, p=.861). The pre-test score for *total amount of observed behaviour* is a significant covariate in the model (F(1,48) = 4.154, p<.05). Thus, even after controlling for initial differences between the *experimental group* and *control group*, differences remain.

4.2 Results of the survey data

This paragraph will start with assessing the raw data acquired by the survey. Data on the post-test variables is missing for one of the 119 participants. Moreover, data from one participant is removed due to unrealistic values (e.g., *experienced number of bystander conflicts* ≥ 30, whereas the average is 1.03). This leaves a total of 118 participants, together forming 61 teams. Of these 61 teams, 37 teams were in the experimental group and 24 teams were in the control group. For two of the 61 teams, data was available for one participant only.

Descriptive statistics for team performance are given in Table 4.4, describing sample size, mean and standard deviation. The correlations between teamwork behaviour variables is given for the experimental group (Table 4.5, above the diagonal) and for the control group (Table 4.5, below

the diagonal). Descriptive statistics for each of the teamwork behaviour variables is given in Table 4.6, describing sample size, means and standard deviations.

Table 4.4: Descriptive statistics of post-test team performance, aggregated at a team-level (n = 61)

Group	n	Variable	M	SD	Scenario	n	M	SD
Control	24	Team performance	5.19	.48	**Heroin**	13	5.04	.45
					Diabetes	11	5.39	.46
Experimental	37	Team performance	5.34	.73	**Heroin**	14	4.76	.68
					Diabetes	23	5.70	.51

Table 4.5: Correlations between post-test teamwork behaviours, aggregated at a team-level (n = 61). Experimental group above the diagonal, control group below

Teamwork behaviour	1	2	3	4	5
Closed-loop communication	-	.68**	.57**	.75**	.70**
Adaptive leadership	.42*	-	.72**	.77**	.52**
Backup behaviour	.53**	.78**	-	.79**	.53**
Mutual performance monitoring	.69**	.75**	.68**	-	.68**
Team performance	.61**	.26	.16	.45	-

*= Correlation is significant at p<.05, **= Correlation is significant at p<.01, two-tailed

All but two correlations are significant and all are (highly) positively correlated. This indicates that if the score for a teamwork behaviour was higher, this is similar for the other teamwork behaviours.

Method of analysis

To test whether the teamwork behaviours act as mediating variables between *training* and *team performance*, as described by the research model and hypothesized by hypotheses 6 to 11, the classical procedure by Baron and Kenny (1986) is used as a reference point. Baron and Kenny (1986) suggest there is mediation whenever three conditions are met (*causal step approach*): (1) the suggested main effect of the independent variable on the outcome variable is not equal to zero (i.e. the effect of *training* on *team performance*). Otherwise there is no effect to explain; (2) the effect of the independent variable on the mediating variable is not equal to zero (i.e. the effect of *training* on any of the suggested *teamwork behaviours*). Otherwise there is no mediating variable; and (3) the effect of the mediating variable on the outcome variable is not equal to zero (i.e. the effect of the *teamwork behaviours* on *team performance*). Otherwise the mediator variable does not function as a sufficient explanation in the model. However, Zhao, Lynch & Chen (2009) state that adhering to the causal step approach is too strict: there can still be an effect left if the first step is not met. So, whatever the result of the first step is, it often remains interesting to analyse at least the second step as well.

As such, the main effect of *training* on *team performance* is first of all analysed via a two-way (*condition*, *scenario*) ANCOVA, having two levels for group (experimental, control) and two levels for scenario (heroin, diabetes), including the post-test score of the team performance as the

Table 4.6: Descriptive statistics of the post-test teamwork behaviour variables (sample size (n)=61, mean (M), standard deviation (SD)

Group	n	Teamwork behaviour	M	SD	Scenario	n	Teamwork behaviour	M	SD
Control	24	Closed-loop communication	3.77	.41	Diabetes	11	Closed-loop communication	3.90	.29
		Adaptive leadership	3.92	.43			Adaptive leadership	3.96	.31
		Backup behaviour	3.94	.35			Backup behaviour	3.84	.39
		Mutual performance monitoring	3.98	.39			Mutual performance monitoring	3.85	.29
					Heroin	13	Closed-loop communication	3.65	.46
							Adaptive leadership	3.89	.52
							Backup behaviour	4.01	.30
							Mutual performance monitoring	4.09	.43
Experimental	37	Closed-loop communication	3.72	.63	Diabetes	23	Closed-loop communication	3.92	.47
		Adaptive leadership	3.99	.42			Adaptive leadership	4.09	.36
		Backup behaviour	3.82	.52			Backup behaviour	3.90	.48
		Mutual performance monitoring	3.85	.55			Mutual performance monitoring	3.94	.50
					Heroin	14	Closed-loop communication	3.34	.72
							Adaptive leadership	3.88	.52
							Backup behaviour	3.71	.58
							Mutual performance monitoring	3.73	.62

dependent variable and the pre-test score of it as a covariate. If the first condition of Baron and Kenny's (1986) procedure is met (i.e. hypothesis 6 is supported), the second and third condition (related to hypothesis 7 to 11) are analysed via the SPSS-script for *(multiple) mediator models* described by Preacher and Hayes (2008). If the first condition is not met, the direct effect of *training* on each of the *teamwork behaviours* is still analysed by performing a two-way (*condition, scenario*) ANCOVA for each teamwork behaviour. Each ANCOVA has two levels for group (experimental, control) and two levels for scenario (heroin, diabetes), and includes the post-test score of the hypothesized teamwork behaviour variable as the dependent variable and the pre-test score of it as a covariate.

Team performance

Hypothesis 6 stated a positive relation between *training* and *team performance*. No significant relation between *training* and *team performance* was found ($M_{\text{control}} = 5.19$, $M_{\text{experimental}} = 5.34$; F(1,56) = .014, p =.328). However, the interaction between *group* and *scenario* is significant (F(1,56) = 19.72, p < .05), and depicted in Figure 4.4. Further inspection indicates that the simple slope of the interaction effect is significant for the *experimental group* ($t = 5.13, p < .05$) but not for the control group ($t = 1.60, p = 0.11$). This indicates that there is a significant difference of *team performance* between scenarios ($M_{heroin} = 4.76, M_{diabetes} = 5.70$) in the *experimental group*, but not in the *control group* ($M_{heroin} = 5.04, M_{diabetes} = 5.39$). No significant effect was found for the pre-test score of *team performance* as a covariate (F(1,56) = .972, p=.328).

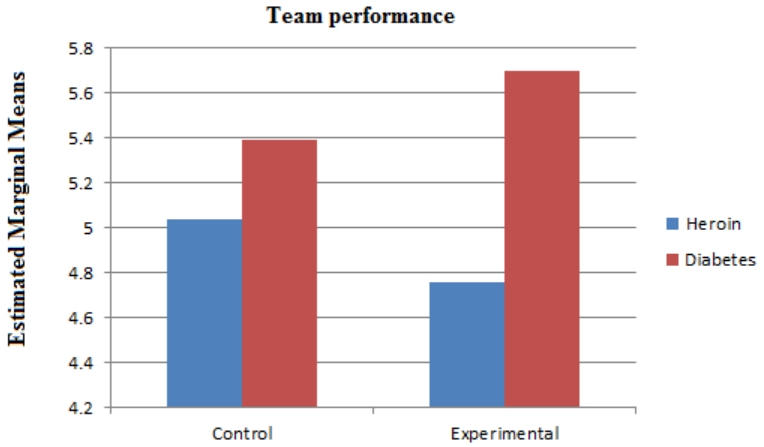

Figure 4.4: Interaction effect between group and scenario for team performance

Situational awareness

Hypothesis 7 stated that *situational awareness* mediates the effect of *training* on *team performance*. The data for *situational awareness* was only acquired by the use of the observational tool and not by the survey. There was no significant effect found for *training* on *situational awareness* (Paragraph 4.1.2, p.38,40), as was suggested by hypothesis 1 . Similarly, no significant effect for *training* on *team performance* was found, as was suggested by hypothesis 6. Because both hypothesis 1 and hypothesis 6 are not supported, neither is hypothesis 7.

Closed-loop communication

Hypothesis 8 stated that *closed-loop communication* mediates the effect of *training* on *team performance*. Even though the main effect of *training* on *team performance* is not significant, the effect of *training* on *closed-loop communication* was still tested. No significant relation between *training* and *closed-loop communication* was found ($M_{\text{control}} = 3.77$, $M_{\text{experimental}} = 3.72$; F(1,56) = .594, p =.444). As such, hypothesis 8 is not supported. No interaction effect was found between *group* and *scenario* (F(1,56) = 2.214, p=.142). The pre-test score for *closed-loop communication* is a significant covariate in the model (F(1,56) = 5.967, p< .05, η^2 = .096). Thus, there were initial differences between the *experimental group* and *control group*.

Adaptive leadership

Hypothesis 9 stated that *adaptive leadership* partially mediates the effect of *training* on *team performance*. Even though the main effect of *training* on *team performance* is not significant, the effect of *training* on *adaptive leadership* was still tested. No significant relation between *training* and *adaptive leadership* was found ($M_{\text{control}} = 3.92$, $M_{\text{experimental}} = 3.99$; F(1,56) = .013, p =.908). As such, hypothesis 9 was not supported. No interaction effect is found between *group* and *scenario* (F(1,56) = 2.940, p=.092). The pre-test score of *adaptive leadership* is a significant covariate in the model (F(1,56) = 14.115, p < .05, η^2 = .201). Thus, there were initial differences between the *experimental group* and *control group.*

Backup behaviour

Hypothesis 10 stated that backup behaviour partially mediates the effect of training on team performance. Even though the main effect of *training* on *team performance* is not significant, the effect of *training* on *backup behaviour* was still tested. No significant relation was found between *training* and *backup behaviour* ($M_{\text{control}} = 3.94$, $M_{\text{experimental}} = 3.82$; F(1,56) = .188, p =.667). As such, hypothesis 10 is not supported. No interaction effect was found between *group* and *scenario* (F(1,56) = 3.531, p=.065). Similarly, no significant effect was found for the pre-test score of *backup behaviour* as a covariate is not significant (F(1,56) = 2.659, p=.109).

Mutual performance monitoring

Hypothesis 11 stated that mutual performance monitoring mediates the effect of training on team performance. Even though the main effect of *training* on *team performance* is not significant, the effect of *training* on *mutual performance monitoring* was still tested. No significant relation found between *training* and *mutual performance monitoring* was found ($M_{\text{control}} = 3.98$, $M_{\text{experimental}} = 3.85$; F(1,56) = .919, p =.342). As such, hypothesis 11 is not supported. No interaction effect was found between *group* and *scenario* (F(1,56) = 3.929, p=.052). The pre-test score of *mutual performance monitoring* is a significant covariate in the model (F(1,56) = 5.148, p<.05 , η^2 =.084). Thus, there were initial differences between the *experimental group* and *control group*.

Teamwork behaviour

Closed-loop communication, adaptive leadership, backup behaviour and mutual performance monitoring were combined as one weighted average *teamwork behaviour* variable, based on the results of the Principle Component Analysis (Paragraph 3.3.2). Descriptive statistics for the teamwork behaviour variable are given in Table 4.7. No significant relation between *training* and *teamwork behaviour* was found ($M_{\text{control}} = 3.90$, $M_{\text{experimental}} = 3.87$; F(1,56) = .202, p =.655). The interaction effect between *group* and *scenario* was not significant (F(1,56) = 4.00, p=.050). Even though the interaction effect was not significant, this was only by the smallest of margins. It is therefore depicted in Figure 4.5 for visualization. The pre-test score for teamwork behaviour is a significant covariate in the model (F(1,56) = 8.796, p<.05, η^2=.136). Thus, there were initial differences between the *experimental group* and *control group*.

Table 4.7: Descriptive statistics of teamwork behaviour (sample size (n = 61), mean (M), and standard deviation (SD))

Group	N	Variable	M	SD	Scenario	N	M	SD
Control	24	Teamwork behaviour	3.90	.35	**Heroin**	13	3.90	.41
					Diabetes	11	3.91	.28
Experimental	37	Teamwork behaviour	3.87	.46	**Heroin**	14	3.69	.50
					Diabetes	23	3.98	.40

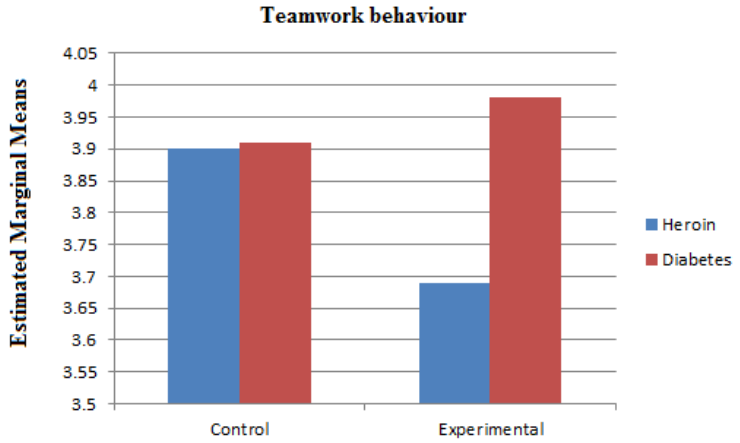

Figure 4.5: Interaction effect between group and scenario for teamwork behaviour

5. Conclusion

In this chapter I will summarize the main conclusions of this master thesis. In Paragraph 5.1 a recapitulation of the research objective and research methods is given. In Paragraph 5.2 conclusions for the results of the observational tool are presented. In Paragraph 5.3 conclusions for the results of the survey are presented. In Paragraph 5.4 the implications on a theoretical and practical level are given. In Paragraph 5.5 a description is given of the research's limitations. Lastly, in Paragraph 5.6 a suggestion is made for two future research directions.

5.1 Research objective and research methods

The objective of this master thesis was to s*uggest and analyse a research model that incorporates the expected effects training focussed on teamwork behaviour has on team performance, for paramedic teams experiencing undesirable behaviour*. The research model is depicted in Figure 5.1.

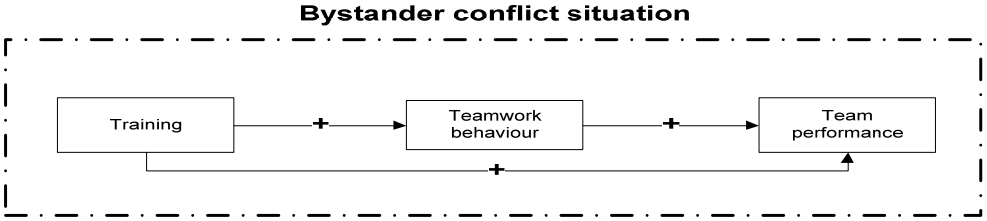

Figure 5.1: Research model

The theoretical framework (Chapter 2) on which the research model is based, describes the *bystander conflict* to be part of undesirable behaviour, and is defined as "*the perception of an individual (within the primary process) that his/her task(s) or goals are frustrated by one or more persons (outside the primary process) through irrelevant behaviour that interrupts focused concentration on the primary task being carried out*". It is suggested that certain teamwork behaviours can buffer the expected negative effect the bystander conflict has on team performance, for paramedic teams experiencing it. This suggestion is based on the demand-resource perspective, as described by the Job Demand-Resources (JD-R; Bakker and Demerouti, 2007); describing that job resources can buffer the negative effects high job demands have on job outcomes. If a paramedic experiences a bystander conflict, it is experienced as an additional demand which is expected to negatively affect time performance as an outcome.

It was argued that paramedic teams have a complex and intensive work flow pattern and the critical lever to maintain or improve performance is whether teamwork behaviour is sufficient (Salas *et al.*, 1997). Teamwork behaviour can be characterised as a non-technical skill, and training is expected to be the most efficient way to improve it. As such, it is often used in the healthcare setting because of their high within-team interdependence (e.g., McGaghie *et al.*, 2006; Rosen *et al.*, 2008; Bowers *et al.*, 1998; Salas *et al.*, 2001; Shapiro *et al.*, 2004; Baker *et al.*, 2003). Training teamwork behaviour is expected to make them part of the team's resources. They can then be used to buffer the negative effect the bystander conflict has as an additional demand, as well as increase team performance in general.

Data acquired during an experiment was used to analyse the research model. The experiment was held at the RAV (*Regionale Ambulance Voorziening*) Brabant Zuid-Oost, responsible for the ambulance services. 119 participants divided over 61 teams took part in the experiment, consisting of two videotaped five minute scenarios (*heroin scenario* and *diabetes scenario*) per team. The 61 teams consisted of 37 teams in the experimental group and 24 teams in the control group. The two scenarios were each characterised by a different bystander conflict situation. Each participant filled in a survey before the first scenario, and after each of the two scenarios. The experimental group received a 1.5 hour training between the two scenarios. The control group only received unstructured feedback between the two scenarios.

There were three sub-goals formulated within the master thesis:

1. *Create an observational tool usable for coding different relevant teamwork behaviours during the experimental design the participants take part in;*
2. *Assess teamwork behaviour observed in the context of the experimental design with the observational data, to evaluate the research model;*
3. *Assess teamwork behaviour and team performance in the context of the experimental design with the survey data, to evaluate the research model.*

Using more than one method of acquiring and analysing data is essential because using only one method can lead to *deficient measurement* which may *beget deficient decisions* about the training's effectiveness (Rosen *et al.*, 2008). To make use of observations, it is essential to have a *robust tool* for analysis, being *necessary to benchmark good teamwork skills* (Hull *et al.*, 2011). Because most of the *observational tools* within literature were not suited for the master thesis' context, one was created (first sub-goal). The second and third sub-goal answer the question whether the research model is a sufficiently good representation of the relationships between training, teamwork behaviour and team performance in a bystander conflict setting.

5.2 Observational tool

The observational tool has five macro-level teamwork behaviours (i.e. *situational awareness, closed-loop communication, adaptive leadership, backup behaviour,* and *mutual performance monitoring*), based on the 'Big Five' model suggested by Salas *et al.* (2005) and other research (e.g., Salas *et al.*, 1995). The five teamwork behaviours contain 23 different behaviours (with two 'control' options of '*no teamwork behaviour*'). These behaviours are organised in a coding scheme by whether the behaviour is implicit or explicit, and, whether it coordinated information, action, or interrelations between persons. The behaviours are based upon research by Kolbe *et al.* (2009). Their *coding units,* i.e. their units of measurement, are single *utterances or actions that fit into one category*. As such, they function on a micro-level. These micro-level behaviours were aggregated within 20 second time slots, which proved to be an adequate choice for observing behaviour without over-coding it (compared to shorter timeslots) or under-coding it (whenever behaviour is prolonging) (Weingart, 1997).

It was hypothesized that training increases the absolute number of observed macro-level behaviour, which would indicate a positive effect of training on teamwork behaviour. Furthermore, the micro-level behaviours can indicate whether there are percentage-wise

improvements on a lower level of behaviour. This takes into account whether instructions in the training to use a specific micro-level behaviour in a certain situation was effective.

First of all, for the macro-level behaviours it was found that there was significantly more *closed-loop communication*, *adaptive leadership*, *backup behaviour*, and *total amount of teamwork behaviour* observed for paramedic teams receiving training. However, this was not the case for *situational awareness* and *mutual performance monitoring*. This partially supports the sub-part of the research model that training increases teamwork behaviour for paramedic teams while experiencing a bystander conflict. For the two macro-level behaviours not increased, this can possibly be explained by the relative technical task simplicity of the scenarios[11]. Based on their research of teams in anaesthesiology, Gaba *et al.* (2001) argue that they all performed with acceptable levels of technical performance because they should all have had sufficient training on that part. I suggest this is similar for the paramedic teams in this master thesis, as training technical skills is a mandatory part of their job. Regarding situational awareness, it is assumed that all teams have some basic protocol for assessing a new situation (e.g., via obtaining information). For a simple technical task, the maximum of necessary required information to be aware of what is happening in that situation can be relatively quickly reached. Regarding mutual performance monitoring, the simplicity of the technical task can cause a lack of the need to monitor what another team member is doing because it is perceived to be less likely something can actually go wrong. This is an interesting finding as it suggests that when there is an additional demand in a simple medical situation the coordination of action (e.g., backup behaviour and adaptive leadership) appears to be more important behaviour for team members. It is suggested that whenever the technical complexity of the medical situation increases, this preference is expected to shift more to coordinating information and interrelations (Byström & Järvelin, 1994). Furthermore, for *adaptive leadership*, there was a significant interaction effect as well when taking the differences between scenarios the paramedic teams took part in into account. The effect of training on adaptive leadership was greater for the *heroin* scenario than for the *diabetes* scenario. This is an interesting finding as it suggests that if the situation becomes more difficult (i.e. the heroin scenario), the dynamic delegation of leadership behaviour becomes more salient to handle the situation (Klein *et al.*, 2006; Heifetz *et al.*, 2009).

Secondly, when looking at the micro-level of behaviour, there was more *'giving orders'* for the paramedic teams receiving training, but less '*unsolicited task-relevant actions'*. This suggests that certain reactive actions can become more salient than anticipative actions. This is an interesting finding, as it was argued that anticipation of the unknown rather than reacting to it when something already has occurred, was an important aspect in a complex and dynamic work setting characterised by non-routine events (Paragraph 3.2.4). However, it is suggested this is based on the nature of the bystander conflict when compared to other non-routine events a paramedic team encounters. Unsolicited task-relevant action is mainly characterised by routine behaviour with not much interaction. It is suggested that team members realise this can cause an escalation of the situation if the bystander feels ignored by this kind of behaviour. In that case, the routine behaviour is replaced by more steering and reacting behaviour such as '*giving orders'* to demarcate the situation.

[11] The scenarios consisted of basic technical skills, such as intubation or insulin injection.

When looking at the differences between the two scenarios, in the heroin scenario, there was less '*giving feedback'* and '*initiating action'*. But, there was more '*giving orders'* and '*watching actions of other team members'*. This is an interesting finding as it suggests that in a more difficult situation (i.e. the heroin scenario) there is more demarcating behaviour imposing restrictions and behaviour monitoring those set restrictions. Rather than team members attempting relational strategizing (Giebels, 2002). This also suggests that a training focus on specific micro-level behaviour in a certain situation is effective to some extent.

Lastly, results of the correlations between the macro-level behaviours are in line with the Attentional Focus Model (AFM; Karau & Kelly, 1992). The AFM (Paragraph 2.3) represents the link between time pressure and group effectiveness, suggesting that time pressure interacts with several factors and can decrease attentional focus, thereby possibly narrowing the use of teamwork behaviour necessary to overcome a situation. The high negative correlations suggest that paramedic teams experience time pressure and as such focus their behaviour onto certain types of behaviour while decreasing others. This is an interesting finding as the results suggests that behaviours more associated with the coordination of information and interrelation become less salient than behaviours associated with coordination of action (and vice versa), whenever demands increase. Suggesting that even *if* training would increase the absolute number of observed behaviour for all macro-level teamwork behaviours in the scenario, this is not always true for a specific moment *within* the scenario (i.e. not all types of teamwork behaviour occur simultaneously in a certain time frame).

5.3 Survey

Participants' perception of their teamwork behaviour and team performance was measured by using multiple surveys. The survey data was used to test the research model, stating that teamwork behaviours (i.e. *closed-loop communication*, *adaptive leadership*, *backup behaviour*, *mutual performance monitoring* and *situational awareness*) mediate the expected positive effect of training on team performance. However, the premises of the research model were not supported.

First of all, results indicate there is no effect found for the paramedic teams receiving training and their team performance. This is an interesting result because most scientific literature suggests training has a positive effect on team performance (e.g., Salas *et al.,* 2008). It is suggested that because of the technical simplicity of the scenarios, the paramedic teams perceive they have performed well independent of the possible side-effects the bystander conflict might have (e.g., on rational thinking). If that would be the case, training to handle the conflict does not have a significant effect on the team's perception of team performance. It is furthermore suggested that this is in line with the basic structure of the survey items measuring team performance: four out of five items describe outcome-related performance (e.g., *quality of care*) rather than process-related performance (e.g. *efficiency of work*). However, there is an interaction effect for the paramedic teams receiving training and the different scenarios. The interaction effect suggests that the effect of training on team performance is positive for the diabetes scenario but negative for the heroin scenario. This is an interesting result as it suggests that paramedic teams starting with the *easy* scenario (i.e. the diabetes scenario) might be less motivated to learn during the training because they perceive that everything went rather well and

this causes their intrinsic motivation to be lower (Lepper & Henderlong, 2000). Then, when confronted with a more difficult scenario (i.e. the heroin scenario) after the training, they may perceive their performance has decreased if their behaviour would be similar. Because the more difficult aspects of what could go wrong are highlighted by the training, but are also more actively remembered in general (Kensinger, 2007). It similarly suggests that paramedic teams starting with the *difficult* scenario are motivated more during the training, as long as it is not experienced being *too* difficult (Lepper & Henderlong, 2000). Then, when confronted with an easier scenario after the training the teams may perceive their performance has increased. It is suggested that this occurs because of the combination that teams have learned more during the training because they are intrinsically motivated more, and, the scenario being inherently easier.

Secondly, no effect was found for the paramedic teams receiving training and all of the suggested teamwork behaviours. As such, there is no support for the teamwork behaviours acting as mediating variable in the research model. It should be noted that the pre-test scores for closed-loop communications, adaptive leadership, mutual performance monitoring and the weighted average of all four teamwork behaviours combined as one variable were significant. This suggests that there were initial differences between the control group and the experimental group. Even though the interaction effect between paramedic teams receiving training and the weighted average of the combined teamwork behaviours was not significant, this was only by the smallest of margins. It is therefore suggested that this result can be used as an indication there might be similar reason like there was for team performance as why there were no effects found. When using a survey to acquire data, team members use a *retrospective relative score* (i.e. a comparison between '*what I could have done*' and '*what did I actually do*'). It is suggested that paramedic teams starting with the easier scenario may perceive there is a gap between *what they could have done* and w*hat they actually did*, when taking part in the more difficult scenario after the training. If this gap is relatively big, they may perceive that their teamwork behaviours have decreased. Similarly, paramedic teams starting with the more difficult scenario may perceive the gap to be smaller because they perceive the training to be more effective, when taking part in the easier scenario afterwards. As such, they might perceive their teamwork behaviours have increased.

5.4 Theoretical and practical implications

Theoretical implications

This master thesis contributes theoretically to the scientific literature in two ways.

First of all, a research model proposing a way to improve a social issue is investigated. A social issue which is getting increasingly more media and political attention, yet having barely been the topic of scientific research. Mixed support was found for this research model, describing how teamwork behaviour can buffer the effect the bystander conflict is expected to have on team performance, and, how training can improve this situation. There are some notable differences between the conclusions for the data acquired by the observational tool (Paragraph 5.1) and conclusions for the data acquired by the survey (Paragraph 5.2). It is suggested that the participant's perception of their teamwork via retrospective impressions is different than the rater's assessment of them. Such differences in results when using different methods of obtaining

data is something future research has to take into account. Zimmerman, Roussea, Duffy, Devers, Gillies, Wagner, Draper, Shortell and Knaus (1994) demonstrated something similar during an organizational case study at an Intensive Care Unit (ICU). Using a combination of their implicit judgment (i.e. by directly observing and interviewing clinical and organizational researchers), organizational questionnaires and self-evaluations of staff members, they failed to distinguish between levels of ICU performance (i.e. the high-low risk-adjusted survival ration). They proposed that you cannot solely rely on one method of measurement.

Secondly, a methodology based on existing literature to create an observational tool to observe different basic teamwork behaviours is suggested, while being adaptable to context specifics. Such an observational tool is relevant for two separate reasons. First of all, current measurement instruments for teamwork behaviour are mostly limited to self-reports via surveys, and self-reports are known for being susceptible to bias. Several forms of bias can influence the validity of the results, for example, inaccurate recall (measurement invalidity) and subjective assessment (inter-rater invalidity). Besides bias, using more than one methodology of obtaining and analysing data is critical to evaluate training effectiveness (Rosen *et al.*, 2008). Secondly, most of the observational measurement tools available are fairly new (i.e. observational research of teamwork behaviour is still in its infancy) and therefore most of them are too context specific. This causes them to not be usable in a more general context. As such, an observational tool with a more general applicable theoretical foundation, adaptable to context specifics by adjusting decision rules and behavioural markers can circumvent these issues. And such a tool is thereby a welcome addition to the infant state the field is in.

Practical implications

The progress of the development of training programmes for the healthcare sector has significantly increased in the last decade and shows some promising results. A new training programme that combines aspects of existing training programmes with new aspects, and applies them to a specific problem-context that is recognized as becoming increasingly more important, is an important condition to continue this progress. Given the problem-context of the bystander conflict, the training and experimental design described in this master thesis was the first of its kind. Focussing on, and analysing other training contents in the same setting (e.g., conflict management, individual resources, other scenarios) and/or applying them in different sectors (e.g., other sectors of public safety) should be the next steps in the training's development process. All in all, continuing this approach will be necessary to improve coping with the *bystander conflicts* that personnel with a public services task increasingly experience, being one of the current social issues getting the most attention in our society. Besides the specific bystander conflict setting, training non-technical skills such as teamwork behaviour is also in general relevant for paramedic teams. It is therefore advised for such a programme to be included as an optional choice for the RAV-personnel to satisfy the needs of their training portfolio. It is furthermore advised that the current programme (i.e. *with* a bystander conflict focus) should be a one-time mandatory training for all RAV-personnel that has not taken part in the training's initial form. Because of the likelihood that they will experience a bystander conflict sooner or later, it is suggested this will decrease the negative consequences whenever they do, *if* they have experienced it in a training for the first time. If they have already experienced a bystander conflict, the training can increase the probability that a future encounter will be dealt with

accordingly after its occurrence, something which is currently often not the case (Abraham *et al.*, 2011).

Furthermore, the issue how other people should be trained to use the observational tool of this master thesis for future analytical purposes has to be addressed. A training guide consisting of an extended theoretical background based on the foundation of Paragraph 3.2.3 and Paragraph 3.2.4 and the literature referenced to in those paragraphs should be created. Previously analysed footage of training sessions can then be used for comparison (e.g. Cohen, 1960) whenever others categorize utterances of observed behaviour to the behaviours of the coding scheme, via *concept learning*. That is, "*the search for and listing of attributes that can be used to distinguish exemplars from non exemplars of various categories"* (Bruner, Goodnow & Austin, 1967, p.233). Obviously also new observational data of multiple raters can be compared with each other.

5.5 Research limitations

There are of course some limitations that should be considered when interpreting the results of this master thesis.

ICC2 values

Not all indices-criteria for aggregating individual data on a team-level were met. The criterion that was not met gives an indication that the differences between groups were small. This suggests that it might be difficult to detect emerging relationships when the means of those groups are taken into account.

Validation of the observational tool

Since the observational tool in this master thesis was only used by one person to analyse data, nothing can be said about its inter-rater agreement. In the future it should be investigated if the same results are found if the tool is: (1) tested by multiple persons independent from each other using the same data; and (2) applied in a different context.

No objective measurement of team performance

Team performance was only measured by the retrospective impression of participants, by means of a survey. However, an objective measurement of team performance is preferred, especially for a medical scenario with technical simplicity. For example, even if the set goal of patient-care is reached, it might still be interesting to see how long it took the team members to start the intubation process, or stabilize the patient altogether.

5.6 Future research direction

In this paragraph, I suggest two important future research directions.

Measuring team performance in healthcare

Jeffcott and Mackenzie (2008, p.188) argue that "*effective team performance is important to measure in order to determine how clinicians should be trained for safe and effective patient care*", yet that it "*is challenging to measure*". They suggest several methodologies, including clinical surveys, direct observation, and real-life video-based analysis, each with their own

deficiencies. Most of these deficiencies were already mentioned in this master thesis. For example, self-reported behavior by clinical surveys does not always match behavior as it is watched by others and it "*cannot be relied upon in isolation to capture the complexity of teamwork in a healthcare setting"* (Jeffcott & Mackenzie, 2008, p.192). Furthermore, video-analysis is time consuming, equipment is still not fail proof with regard to capturing all occurring events and it requires a specific coding scheme for the context. As such, Jeffcott and Mackenzie (2008, p.189) suggest that a "*consensus on a theoretical framework or model of team performance would provide researchers with the ability to describe and test hypotheses concerning the interrelationships of various performance measures, as well as the interdependencies of the predictors and the outcome criteria, across healthcare settings"*. Even though this 'take home message' is currently already four years old, not much progress has been made since, to the best of my knowledge. As such, because it is a critical issue for the scientific field of training in healthcare, it should receive more attention.

Sequential coding

Sequential coding can identify whether there are optimal *cycles of behaviour*. Some researchers assume that for a given group task there is a systematic behavioural optimal way to proceed that can explain variance "*over and above more traditional indices…such as the total amount*" (Tschan, 2002, p.616). However, others argue it is mainly task, situation, or group dependent (Hirokawa, 1983; Predmore, 1995).

What most approaches advocating this concept have in common, is a cycle of action regulation that consist of (1) a goal to achieve; (2) orientation and preparation of action; (3) execution of action; and (4) monitoring action and evaluation. The last step can either lead to termination if the goal is achieved, or correcting behaviour. Interruptions during this cycle may change the process, impeding performance (Tschan, 1995).

Given the research setting in this master thesis, the bystander is likely to interrupt the cycle of action regulation of the paramedic team, thereby impeding behaviour. Using the classification of the macro-level teamwork behaviours and the action regulation cycle, *situational awareness* and *closed-loop communication* could function as the orientation and preparation phase. Similarly, *adaptive leadership* and *backup behaviour* as the execution of action phase, and, *mutual performance motoring* as the monitoring and evaluation phase. If an objective measure of *team performance* is used (i.e. the time it takes to start a procedure), it can be investigated if teams that perform better have certain cycles. And, if a training mentioning this concept can improve them. A similar proposition can be tested for the micro-level behaviours, and if there are recurrent cycles for teams that perform better (e.g., '*giving orders'* always follows '*verbalising interpretation of a situation'*). As such, the micro-level behaviours can be investigated in a more meaningful way than the current percentage-wise distributions.

References

Abraham, M., Flight, S., & Roorda, W. (2011). *Agressie en geweld tegen werknemers met een publieke taak – Onderzoek voor Veilige Publieke Taak 2007-2009-2011* (DSP – groep).

Arrow, H., McGrath, J.E., & Berdahl, J.L. (2000). *Small groups as complex systems: formation, coordination, development, and adaption.* Thousand Oaks: Sage.

Baker, D.P., Day, R., & Salas, E. (2006). Teamwork as an essential component of high-reliability organizations. *Health Service Research,* 41(4), 1576-1598.

Baker, D.P., Gustafson, S., Beaubien, J.M., Salas, E., & Barach, P. . Medical team training programs in health care. In: Henriksen, K. Battles, J.B. Marks, E.S., & Lewin, D.I. (Eds.). *Advances in patient safety: from research to implementation.* AHRQ Publication, Rockville MD, No. 05, Feb 2005.

Baker, D.P., & Salas, E. (1992). Principles for measuring teamwork skills. *Human Factors,* 34(4), 469-475.

Bakker, A.B., & Demerouti, E. (2007). The Job Demands-Resources model: state of the art. *Journal of Managerial Psychology,* 22(3), 309-328.

Bakker, A.B., Demerouti, E., De Boer, E. & Schaufeli, W.B. (2003). Job demands and job resources as predictors of absence duration and frequency. *Journal of Vocational Behavior,* 62, 341-356.

Bakker A.B., Demerouti, E., Taris, T., Schaufeli, W.B. & Schreurs, P. (2003). A Multi-group analysis of the Job-Demands-Resources model in four home care organizations. *International Journal of Stress Management,* 10, 16-38.

Bakker, A.B., Demerouti, E. & Verbeke, W. (2004). Using the Job Demands-Resources model to predict burnout and performance. *Human Resource Management,* 43, 83-104.

Bandow, D. (2001). Time to create sound teamwork. *The Journal of Quality and Participation,* 24, 41-47.

Baron, R.M., & Kenny, D.A. (1986). Moderator-mediator variables distinction in social psychological research: conceptual, strategic, and stochastical considerations. *Journal of Personality and Social Psychology,* 51(6), 1173-1182.

Bartko, J.J. (1976). On various intraclass correlation reliability coefficients. *Psychological Bulletin,* 83(5), 762-765.

Bliese, P.D. (1998). Group size, ICC values and group-level correlations: A simulation. *Organizational Research Methods,* 1, 355-373.

Bliese, P.D. (2000). Within-group agreement, non-independence, and reliability: Implications for data aggregation and analysis. In Klein, K.J., & Kozlowski, S.W.J. (Eds.), *Multilevel theory, research, and methods in organizations: foundations, extensions, and new directions.* p.349-381. San Fransisco: Jossey-Bass.

Bowers, C.A., Salas, E. & Rhodenizer, L. (1998). It is not how much you have but how you use it: toward a rational use of simulation to support aviation training. *International Journal of Aviation Psychology*, 8, 197-208.

Bruner, J., Goodnow, J.J., & Austin, G.A. (1967). *A study of thinking*. New York: Science Editions.

Burke, C.S., Stagl, K.C., Klein, C., Goodwin, G.F., Salas, E., Halpin, S.M. (1996). What type of leadership behaviors are functional in teams? A meta-analysis. *The Leadership Quarterly*, 17, 288-307.

Byström, K., & Järvelin (1994). Task complexity affects information seeking and use. *Information Processing & Management*, 31(2), 191-213.

Campion, M.A., Papper, E.M., & Medsker, G.J. (1996). Relations between work team characteristics and effectiveness: a replication and extension. *Personnel Psychology*, 49, 429-452.

Cannon-Bowers, J.A., Tannenbaum, S.I., Salas, E., & Volpe, C.E. (1995). Defining team competencies and establishing team training requirements. In R. Guzzo, E. Salas, & associates (Eds.), *team effectiveness and decision making in organizations* (pp. 333-380). San Francisco: Jossey-Bass.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement,* 20(1), 37-46.

De Dreu, C.K.W. (2008). The virtue and vice of workplace conflict: food for (pessimistic) thought. *Journal of Organizational Behavior,* 29, 5-18.

De Dreu, C.K.W., & Weingart, L.R. (2003). Task versus relationship conflict, team performance, and team member satisfaction: a meta-analysis. *Journal of Applied Psychology,* 88(4), 741-749.

Demerouti, E., Bakker, A.B., De Jonge, J., Janssen, P.P.M & Schaufeli, W.B. (2001). Burnout and engagement at work as a function of demands and control. *Scandinavian Journal of Work, Environment and Health,* 27, 279-286.

Demerouti, E., Bakker, A.B., Nachreiner, F., & Schaufeli, W.B. (2001). The job demands-resources model of burnout. Journal of Applied Psychology, 86, 499-512.

Diener, E. & Fujita, F. (1995). Resources, personal strivings, and subjective well-being: a nomothetic and ideographic approach. *Journal of Personality and Social Psychology*, 68, 926-935.

Doten, G.W., Cockrell, J. T.,& Sadacca, R. (1966). *The use of teams in image interpretation: Information exchange, confidence, and resolving disagreements* (Tech. Rep. 1151). Washington, DC: U.S. Army Personnel Research Office.

Driskell, J.E. & Salas, E. (1991). Group decision making under stress. *Journal of Applied Psychology,* 76(3), 473-478.

Driskell, J.E. & Salas, E. (1992). Collective behavior and team performance. *Human Factors,* 34(3), 277-288.

Endsley, M.R. (1988). Design and evaluation for situation awareness enhancement. *Proceedings of the Human Factors Society*, 97-101.

Endsley, M.R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37(1), 32-64.

Erp, K.J.P.M. van, Demerouti, E., Rispens, S., & Gevers, J.M.P. (201). *Don't stand so close to me: Bystander conflict – defining an important phenomenon.* 24th Annual IACM Conference, Istanbul, July 3-6, 2011.

Fleishman, E. A.,Mumford, M. D., Zaccaro, S. J., Levin, K. Y., Korotkin, A. L., & Hein, M. B. (1991). Taxonomic efforts in the description of leader behavior: A synthesis and functional interpretation. *Leadership Quarterly*, *2*, 245-287.

Fletcher, G., Flin, R., McGeorge, P., Glavin, R., Maran, N., & Patey, R. (2003). Anaesthetist' Non-Technical Skills (ANTS): evaluation of a behavioral marker system. *British Journal of Anaesthesia*, 90(5), 580-588.

Gevers, J., van Erven, P., de Jonge, J., Maas., M., & de Jong, J. (2010). Effect of acute and chronic job demands on effective individual teamwork behaviour in medical emergencies. *Journal of Advanced Nursing*, 66(7), 1573-1583.

Giebels, E. (2002). Beinvloeding in gijzelingsonderhandelingen: de tafel van tien, *Nederlands Tijdschrift voor de Psychologie*, 57, 145-154.

Goodman, P.S., Ravlin, E., & Schminke, M. (1987). Understanding groups in organizations. In L.L. Cummings & B.M. Staw (Eds.), *Research in organizational behavior,* 121-173. Greenwich, CT: JAI Press.

Hackman, J.R., & Morris, C.G. (1975). Group tasks, group interaction process, and group performance effectiveness: A review and proposed integration. In L. Berkowitz (Ed.), *Advances in experimental social psychology*, 45-99. New York: Academic Press.

Hair, J.F., Black, W.C., Babin, B.J., & Anderson, R.E. (2009). *Multivariate Data Analysis: a global perspective (7th edition).* Pearson Education, Inc: New Jersey.

Hartel, C.E.J., Smith, K., & Prince, C. (1991). *Defining aircrew coordination: Searching mishaps for meaning*. Paper presented at the Sixth International Symposium on Aviation Psychology, Columbus, OH.

Heifetz, R., Grashow, A., & Linksy, M. (2009). Leadership in a (permanent) crisis. *Harvard Business Publishing*.

Helmreich, R.L., Merritt, A.C., & Wilhelm, J.A. (1999). The evolution of crew resource management in commercial aviation. *International Journal of Aviation Psychology*, 9(1), 19-32.

Hershcovic, M.S., & Barling, J. (2010). Towards a multi-foci approach to workplace aggression: A meta-analytical review of outcomes from different perpetrators. *Journal of Organizational Behavior,* 31(1), 24-44.

Hines, S., Luna, K., & Lofthus, J. (2008). Becoming a High Reliability Organization: Operational advice for hospital leaders. *AHRQ Publication, Rockville, MD: Agency for Healthcare Research and Quality*.

Hirokawa, R.Y. (1988). Group communication and decision making performance: a continued test of the functional perspective. *Human Communication Research,* 14, 487-515.

Hobföll, S.E. (1989). Conservation of resources: a new attempt at conceptualizing stress. *American Psychologist*, 44(3), 513-524.

Hull, L., Arora, S., Kassab, E., Kneebone, R., & Sevdalis, N. (2010). Observational teamwork assessment for surgery: content validation and tool refinement. *American College of Surgeons*, 212(2), 234-243.

Ilgen, D.R., Hollenbeck, J.R., Johnson, M., & Jundt, D. (2005). Teams in organizations: from input-process-output models to IMOI models. *Annual Review Psychology*, 56, 517-543.

James, L.R., Demaree, R.J., & Wolf, G (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology*, 69, 85-98.

Jeffcott, S.A., & Mackenzie, C.F. (2008). Measuring team performance in healthcare: Review of research and implications for patient safety. *Journal of Critical Care*, 23, 188-196.

Jett, Q. R., & George, J. M. (2003). Work interrupted: A closer look at the role of interruptions in organizational life. *Academy of Management Review*, 28, 494-507.

Jones, G., & George, J. (1998). The experience and evolution of trust: Implications for cooperation and teamwork. *Academy of Management Review*, 23, 531-546.

Jonge, J. de, Mulder, M.J.G.P., & Nijhuis, F.J.N. (1999). The incorporation of different demand concepts in the Job Demand-Control model: effects on health care professionals. *Social Science and Medicine,* 48, 1149-1160.

Kahn, R.L. & Byosserie, P. (1992). Stress in organizations. In Dunette, M.D. & Hough, L.M. (Eds.). *Handbook of Industrial and Organizational Psychology.* Vol. 3, Palo Alto, CA,: Psychologists Press.

Kaiser, H. F. (1960). *The application of electronic computers to factor analysis*. Educational and Psychological Measurement, 20, 141-151

Karasek, R.A. (1979). Job demands, job decision latitude, and mental strain: implications for job design. *Administrative Science Quarterly,* 24, 285-308.

Karasek, R.A. (1998). Demand/Control Model: a social, emotional, and physiological approach to stress risk and active behavior development. In: Stellman, J.M. (Ed.), *Encyclopaedia of occupational health and safety*. Geneva: ILO.

Karau, S.J., & Kelly, J.R. (1992). The effects of time scarcity and time abundance on Group performance quality and interaction process. *Journal of Experimental Social Psychology,* 28, 542-571.

Karau, S.J., & Kelly, J.R. (2003). Time pressure and team performance: an attentional focus integration. *Research on Managing Groups and Teams*, 6, 185-212.

Kensinger, E.A. (2007). Negative emotion enhances memory accuracy: Behavioural and neuroimaging evidence. *Current Directions in Psychological Science,* 16(4), 213-218.

Kerckhaert, A., & De Ruig, L. (2010). *Aanpak van agressie en geweld in internationaal – Quickscan naar aard, omvang en beleid in vier landen* (Research voor Beleid).

Klein, K.J., & Kozlowski, S.W.J. (2000). From micro to meso: critical steps in conceptualizing and conducting multilevel research. *Organizational Research Methods*, 3(3). 211-236.

Klein, K.J., Ziegert J.C., Knight, A.P., & Xiao, Y. (2006). Dynamic delegation: shared, hierarchical and deinvidiualized leadership in extreme action teams. *Administrative Science Quarterly,* 51, 590-621.

Kolbe, M., Künzle, B., Zala-Mezö, E., Wacker, J., & Grote, G. (2009). Measuring coordination behavior in anaesthesia teams during induction of general anaesthetics. In: Flin, R. & Mitchell, L. (Eds.). *Safer surgery: analyzing behavior in the operating theatre,* 203-221. Ashgate, Aldershot, UK.

Kohn, L., Corrigan, J., & Donaldson, M. (1999). *To err is human: building a safer health system*. Washington DC: National Academy Press.

Kleinman, D.L., & Serfaty, D. (1989). Team performance assessment in distributed decision making. In R. Gilson, J.P. Kincaid, & B. Goldez (Eds.), *Proceedings of the Interservice Networked Simulation for Training Conference* (pp. 22-27). Orlando: University of Central Florida.

Kozlowski, S.W.J. & Hattrup, K. (1992). A disagreement about within-group agreement: Disentangling issues of consistency versus concensus. *Journal of Applied psychology,* 77, 161-167.

Kozlowski, S.W.J., & Ilgen, D.R. (2006). Enhancing the effectiveness of work groups and teams. *Association for Psychological Science,* 7(3), 77-124.

Leary, T. (1957). *Interpersonal diagnosis of personality; a functional theory and methodology for personality evaluation.* Oxford, England: Ronald Press.

Leiter, M.P. (1993). Burnout as a developmental process: considerations and models. In Schaufeli, W.B., Maslach, C. & Marek, T. (Eds.). *Professional burnout: recent developments in theory and research.* Washington DC: Taylor and Francis.

Lepper, M.R., & Henderlong, J. (2000). Turning "play" into "work" and "work" into "play": 25 years of reseach on intrinsic versus extrinsic motivation. In Sansone, C. & Harackiewicz, J.M. (Eds.). *Iintrinsic and extrinsic motivation: The search for optimal motivation and performance.* San Diego, CA: Academic Press.

LePine, J.A., Podsakoff, N.P., & LePine, M.A. (2005). A meta-analytic test of the challenge stressor-hindrance stressor framework: an explanation for inconsistent relationships among stressors and performance. *Academy of Management Journal,* 48(5), 764-775.

Malec, J., Torsher, L.C., Dunn, W.F., Wiegmann, D.A., Arnold, J.J., Brown, D.A., & Phatak, V. The Mayo High Performance Teamwork Scale: reliability and validity for evaluating key crew resource management skills. *Simulation in Healthcare,* 2(1), 2007.

Marks, M.A., Mathieu, J.E., & Zaccaro, S.J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review,* 26(3), 356-376.

McGaghie, W.C., Issenberg, S.B., Petrusa, E.R. & Scalese, R.J. (2006). Effects of practice on standardized learning outcomes in simulation-based medical education. *Medical education,* 40(8), 792-797.

McGrath, J.E. (1964). *Social psychology: A brief introduction.* New York: Holt, Rinehart, & Winston.

McIntyre, R. M., & Salas, E. (1995). Measuring and managing for team performance: Emerging principles from complex environments. In R. A. Guzzo & E. Salas (Eds.), *Team effectiveness and decision making in organizations* (pp. 9-45). San Francisco: Jossey-Bass.

Pearce, C.L., & Conger, J.A. (2003). *Shared leadership: reframing the hows and whys of leadership.* Thousand Oaks, CA: Sage.

Pearce, C.L. & Sims, H.P. (2002). The relative influence of vertical vs. shared leadership on the longitudinal effectiveness of change management teams. *Group Dynamics: Theory, Research, and Practice*, 6, 172-197.

Porter, C. O. L. H., Hollenbeck, J. R., Ilgen, D. R., Ellis, A. P. J., West, B. J., & Moon, H. (2003). Backup behaviors in teams: The role of personality and legitimacy of need. *Journal of Applied Psychology, 88*, 391-403.

Preacher, K.J., & Hayes, A.F. (2008). SPSS and SAS procedures for estimating indirect effects in simple mediation models. *Behavior Research Methods, Instruments & Computers*, 36(4), 717-731.

Pruitt, D., & J. Rubin. (1986). *Social conflict: Escalation, stalemate and settlement.* New York: Random House.

Rispens, S., & Jehn, K.A. (2010). *Conflict in workgroups: constructive, destructive, and conflict asymmetry*. In D. De Cremer, K. Murnighan, & R. van Dick (Eds.). Social Psycholohy and Organizations.

Roberts, K.H. (1990). Managing high reliability organizations. *California Management Review*, 32(4), 101-113.

Roberts, K.H., & Rousseau, D.M. (1989). Research in nearly failure-free, high reliability organizations: having the bubble. *Transactions on engineering management*, 36(2), 132-139.

Rosen, M.A., Salas, E., Wilson, K.A., King, H.B., Salisbury, M., Augenstein, J.S., Robinson, D.W. & Birnbach, D.J. (2008). Measuring team performance in simulation-based training: Adopting best practices for healthcare. *Simulation in Healthcare*, 3(1), 33-41.

Rouse, W.B., & Morris, N.M. (1986). On looking into the black box: prospects and limits in the search for mental models. *Psychological Bullitin*, 100(3), 349-363.

Salas, E., Burke, S.C., Bowers, C.A. & Wilson, K.A. (2001). Team training in the skies: does Crew Resource Management (CRM) training work? *Human Factors,* 43- 641-674.

Salas, E., & Cannon-Bowers, J.A. (2000). The anatomy of team training. In S. Tobias & J.D. Fletcher (Eds.), *Training & retaining: a handbook for business, industry, government and the military* (pp. 312-335). New York: Macmillan.

Salas, E., & Cannon-Bowers, J.A. (2001). The science of training: A decade of progress. *Annual Review of Psychology*, 52, 471-499.

Salas, E., Diaz-Granados, D., Klein, C., Shawn Burke, C., Stagl, K.C., Goodwin, G.F., & Halpin, S.M. (2008). Does team training improve team performance? A meta-analysis. *Human Factors and Ergonomics Society*, 50, 903-933.

Salas, E., Prince, C., Baker, D.P., & Shrestha, L. (1995). Situation awareness in team performance: Implications for measurement and training. *Human Factors*, 37(1), 123-136.

Salas, E., Rosen, M.A., Burke, C.S., Nicholson, D., & Howse, W.R. (2007). Markers for enhancing team cognition in complex environments: the power of team performance diagnosis. *Aviation, Space, and Environmental Medicine*, 78(5), 77-85.

Salas, E., Rosen, M.A., & King, H. (2007). Managing teams managing crises: principles of teamwork to improve patient safety in the emergency room and beyond. *Theoretical issues in ergonomics science*, 8(5), 381-394.

Salas, E., Sims, D.E., Burke, C.S. (2005). Is there a "Big Five" in teamwork? *Small Group Research*, 36, 555-599.

Schnall, P.L., Landsbergis, P.A. & Baker, D. (1994). Job strain and cardiovascular disease. *Annual Review of Public Health*, 15, 381-411.

Schraagen, J.M., Schouten, T., Smit, M., Haas, F., van der Beek, D., van de Ven, J., & Barach, P. (2010). Assessing and improving teamwork in cardiac surgery. *Quality & Safety in Health Care*, 19, 1-6.

Sexton, J.B., Helmreich, R.I., Glenn, D., Wilhelm, J.A. & Merritt, A.C. (2000). Operating room management attitudes questionnaire (ORMAQ). Technical report 00-2 of the University of Texas.

Shapiro, M.J., Morey, J.C., Small, S.D., Langford, V., Kaylor, C.J., Jagminas, L., Suner, S., Salisbury, M.L., Simon, R., & Jay, G.D. (2004). Simulation based teamwork training for emergency department staff: does it improve clinical team performance when added to an existing didactic teamwork curriculum? *Quality & Safety Health Care*, 13, 417-421.

Siegel, A.I., & Federman, P.J. (1973). Communications content training as an ingredient of effective team performance. *Ergonomics, 16, 403-416.*

Siegrist, J. (1996). Adverse health effects of high effort-low reward conditions. *Journal of Occupational Health Psychology*, 1, 27-41.

Staw, B.M. (1975). Attribution of the “causes” of performance: A general alternative interpretation of cross-sectional research on organizations. *Organizational Behavior and Human Performance*, 13, 414-432.

Stout, R.J., Salas, E., & Fowlkes, J.E. (1997). Enhancing teamwork in complex environments through team training. *Group Dynamics: Theory, Research, and Practice*, 1, 169-182.

Temkin-Greener, H., Gross, D., Kunitz, S.J., & Mukamel, D. (2004). Measuring interdisciplinary team performance in a long term care setting. *Medical care*, 42(5), 472-481.

Tesluk, P., Mathieau, J.E., Zaccaro, S.J., & Marks, M. (1997). Task and aggregation issues in the analysis and assessment of team performance. In: Brannick, M.T., Salas, E., & Prince, C. (Eds.). *Team performance assessment and measurement: Theory, methods, and applications.* Mahwah, NJ: Lawrence Erlbaum.

Todman, J., & Dugard, P. (2007). *Approaching multivariate analysis: an introduction for psychology*. Psychology Press; 1st edition.

Tziner, A., & Kopelman, R.E. (2002). Is there a preferred performance rating format? A non-psychometric approach. *Applied Psychology: An International Review*, 51, 479-503.

Vegchel, N. van, Jonge, J. de, Bakker, A.B., & Schuafeli, W.B. (2002). Testing global and specific indicators of rewards in the Effort-Reward Imbalance model: does it make any difference? *European Journal of Work and Organizational Psychology*, 11, 403-421.

Wagner, J. A. (1995). Studies of individualism-collectivism: Effects on cooperation in groups. *Academy of Management Journal, 38*, 152-172.

Wagner, C., & De Bruijne, M. (2007). *Onbedoelde schade in ziekenhuizen* (EMGO & NIVEL).

Webber, S.S. (2002). Leadership and trust facilitating cross-functional team success. *Journal of Management Development*, 21, 201-214.

Weingart, L.R. (1997). How did they do that? The ways and means of studying group processes. *Research in Organizational Behavior*, 19, 189-239.

Weingart, L.R., Hyder, E.B., & Prietula, M.J. (1996). Knowledge matters: the effect of tactical descriptions on negotiation behavior and outcome. *Journal of Personality and Social Psychology*, 20(6), 1205-1217.

Weinger, M.B., & Slagle, J. (2002). Human factors research in anaesthesia patient safety: techniques to elucidate factors affecting clinical task performance and decision making. *Journal of the American Medical Association*, 9, S58-S63.

Wickens, C.D. (1984). *Engineering psychology and human performance*, Columbus, OH: C.E. Merrill Pyblishing Co.

Zaccaro, S.J., Rittman, A.L., & Marks, M.A. (2001). Team leadership. *Leadership Quarterly*, 12, 417-451.

Zhao, X., Lynch, J.G., & Chen, Q. (2010). Reconsidering Baron and Kenny: Myths and truths about mediation analysis. *Journal of Consumer Research,* 37(2), 197-206.

Zimmerman, J.E., Rousseau, D.M., Duffy, J., Devers, K., Gillies, R.R., Wagner, D.P., Draper, E.A., Shortell, S.M., & Knaus, W.A. (1994). Intensive care at two teaching hospitals: an organizational case study. *American Journal of Critical Care*, 3(2), 129-138.

Appendix A: Training specifics

Training-day schedule

Table A1 shows the basic structure of a one day, including two sessions.

Table A1: Training day schedule

Time	Procedure	
8:30 – 8:45	Introduction to **all** the participants, including: -**General introduction** (trainer and research team) -**Topic introduction** (Goal**:** awareness and training skills) -Schedule of the workshop -**Basic assumptions** (being qualified and motivated) -**Confidentiality** (not a test, scientific research, and anonymity)	
8:45 – 9:00	**Fill in questionnaire T1** (group 1 and 2)	
9:00 – 9:15	Group 1 – **scenario 1**	Group 2 – break
9:15 – 9:30	Group 1 – **questionnaire T2**	Group 2 – **scenario 1**
9:30 – 9:45	Group 1 – break	Group 2 – **questionnaire T2**
9:45 – 11:10	Group 1 and 2 - **Feedback or training**	
11:10 – 11:25	Break – processing training material	
11:25 – 11:40	Group 1 – **scenario 2**	Group 2 – break
11:40 – 11:55	Group 1 – **questionnaire T3**	Group 2 – **scenario 2**
11:55 – 12:10	Group 1 – break	Group 2 – **questionnaire T3**
12:10 – 12:30	**Feedback and wrap-up morning session**	
12:30 – 13:00	**Break**	
13:00 – 13:15	**Fill in questionnaire T1** (group 3 and 4)	
13:15 – 13:30	Group 3 – **scenario 1**	Group 4 – break
13:30 – 13:45	Group 3 – **questionnaire T2**	Group 4 – **scenario 1**
13:45 – 14:00	Group 3 – break	Group 4 – **questionnaire T2**
14:00 – 15:25	Group 3 and 4 - **Feedback or training**	
15:25 – 15:40	Break – processing training material	
15:40 – 15:55	Group 3 – **scenario 2**	Group 4 – break
15:55 – 16:10	Group 3 – **questionnaire T2**	Group 4 – **scenario 2**
16:10 – 16:25	Group 3 - break	Group 4 – **questionnaire T2**
16:25 – 16:45	**Feedback and wrap-up afternoon session**	

Scenario specifics

Heroin scenario (**anti-social bystander**)[12]

The *bystander* is a person living in the nearby neighbourhood and could hear the sirens approach. Because it appears to be close, the bystander decides to go and have a look. It appears to be what he already expected: another drug addict from the clinic nearby. Only recently one of them

[12] Based upon the classification of van Erp *et al.* (2011)

actually threatened the bystander, and he is fed up with all the troubles the drug addicts are responsible for in his neighbourhood! The bystander's **attitude**: *Why don't you help people who actually deserve to be helped: he will just be lying here again next week anyway. Just let him rot and don't waste money that can be used in better ways. I don't pay taxes to help people like him*. **Starting point:** oh, you don't have to hurry, I know these kinds of people, they don't deserve to be helped (!).

Diabetes scenario (**emotional bystander**)
The bystander is living in an apartment complex and finds his neighbour lying on the ground. The bystander is the one calling 112, and is waiting for the paramedics to arrive. However, it is taking a long time and the bystander has experience with situations going wrong because paramedics arrived too late: the bystander's father got paralyzed this way. The bystander's **attitude**: the bystander knows the neighbour is diabetic and has homecare, but, thinks the paramedics should focus on the broken arm and does not hesitate to make that as clear as possible. **Starting point:** why didn't you get here any earlier (!?)

Training content:

Topic	Content
Bystander conflict	**Question**: who had experience with bystander conflicts in their normal work setting (i.e. do participants experience it to be actually relevant?) **Explanation**: bystander (conflict) definition. *In this introduction of the training session the actual topic at hand will be introduced. How did the participants experience the first scenario, and, is this something (perhaps to a lesser or greater extent) they have experienced in real life?*
Methods of handling	**Question:** how was the scenario handled? *This follow-up question assesses how the participants have experienced the scenario more specifically, i.e. via which behaviours they tried to handle the situation: what was effective, and, what was not?* **Methods of handling:** influencing techniques of the '*Table of Ten'* (Giebels, 2002)– what people do and what they *should* do, in different bystander-situations is visualized by the 'Rose of Leary'- framework (Leary, 1957). *Which influencing techniques did participants use, and, can they can think of other methods of handling conflicts they did not use? How do people react to certain techniques, and, why do some work while others do not. Can the participants identify why their own behavior worked, or not, and why? Use the visualizing framework at a base-level, for example, by only drawing the two axis. Where do the participants think their used influencing strategies fit?* **Extension**: describing different types of bystanders (emotional, anti-social, idealistic)

Can the participants identify where all the bystanders are placed on the Rose of Leary framework? Furthermore, which type of bystander did they think they have encountered in the scenario. Do they (now) recognize how to handle the situation, given the framework just described?

Goal: *handling* different types of bystanders with an array of different influencing techniques while maintaining or improving performance

Teamwork

Before-situation:

-Awareness: what is teamwork
– Why does it matter?

Let participants identify and explain what teamwork behaviors they use in practice and why those are important: in general but also in other situations, as would be the case in the scenario. Explain that being aware of the different behaviours and their use is the first step in using them effectively.

During-situation (use of):

Teamwork:
-Communicating: signaling, supporting
-Leadership via adaptability

After-situation:

-Emotional support
-Evaluation

Focus on the fact that besides knowing (before) and applying (during) the concepts of teamwork to handle a situation, the situation itself can still have an effect some time after it has occurred. To cope with this after-effect and gain additional insights, feedback and evaluation are important parts of the process.

Appendix B: Scale measures

The text of each measure is taken from the surveys that were filled in after the first and second scenario. The text in the survey that was filled in before the first scenario is similar, but obviously did not mention the line about '*the scenario you just took part in*'. Furthermore, for the teamwork behaviours, the last column stating to which teamwork behaviour variable the item belongs was only added in this Appendix for clarification.

Team performance

Hieronder volgen enkele vragen over de prestatie van het team waarmee u zojuist het scenario heeft geoefend

In dit team was…

Table B1: Scale measures team performance

	Uitermate slecht	Erg slecht	Slecht	Niet slecht en niet goed	Goed	Erg goed	Uitermate goed
1. De kwaliteit van de hulpverlening…	1	2	3	4	5	6	7
2. De zorg voor patientveiligheid…	1	2	3	4	5	6	7
3. De efficientie waarmee het werk werd gedaan…	1	2	3	4	5	6	7
4. De mate waarin doelstellingen verwezenlijkt zijn…	1	2	3	4	5	6	7
5. De teamprestaties als geheel…	1	2	3	4	5	6	7

Teamwork behaviour

In het scenario werkte u samen met iemand anders. U vormde samen een team. De volgende vragen gaan over dat team.

In dit team…

Table B2: Scale measures teamwork behaviour

	In zeer geringe mate	In geringe mate	Een beetje	In sterkte mate	In zeer sterke mate	
1. Voorzagen we elkaar tijdig van duidelijke informatie.	1	2	3	4	5	*Closed-loop communication*
2. Verifieerden we de informatie die werd gecommuniceerd.	1	2	3	4	5	
3. Luisterenden we naar wat elk van ons had in te brengen.	1	2	3	4	5	
4. Hielpen we elkaar bij de taakuitvoering.	1	2	3	4	5	*Backup behaviour*
5. Anticipeerden we op hetgeen de ander nodig had.	1	2	3	4	5	
6. Wisselden we van rol als de situatie hier om vroeg.	1	2	3	4	5	
7. Zagen we van elkaar wat we aan het doen waren.	1	2	3	4	5	*Mutual performance monitoring*
8. Attendeerden we elkaar op dreigende fouten of complicaties.	1	2	3	4	5	
9. Hielden we oog voor de omgeving, ook in geval van calimiteiten.	1	2	3	4	5	
10. Durfde ieder teamlied, waar nodig, de leiding te nemen.	1	2	3	4	5	*Adaptive leadership*
11. Gaven we elkaar, waar nodig, instructies over de taakuitvoering.	1	2	3	4	5	
12. Droeg ieder van ons oplossingen aan vanuit zijn eigen expertise.	1	2	3	4	5	
13. Paste het team zich aan als de situatie daar om vroeg.	1	2	3	4	5	
14. Presteerde het team goed, ook in omstandigheden die niet optimaal waren.	1	2	3	4	5	
15. Ging het team goed om met onverwachte veranderingen.	1	2	3	4	5	

Appendix C: Construct validity

All absolute scores < .4 were suppressed.

Extraction based on theoretical framework, i.e. on a fixed number of factors (n = 4). For each item n = 119.

Table C1: Teamwork behaviour, Principle Component Analyses, Direct Oblimin rotation, Pattern Matrices

Item*	Components T1				Components T2				Components T3			
	1	2	3	4	1	2	3	4	1	2	3	4
CLC1		.90						.77		-.79		
CLC2		.85						.93		.82		
CLC3								.67		-.72		
BB1				-.52		.86			.73			
BB2		.57				.78			.69			
BB3			.72				.81		.40			.56
MPM1				-.79	.68				.73			
MPM2				-.48					.47			.45
MPM3				-.62	.52							.63
AL1	.54		.42				.79					.85
AL2				-.44	.50							.74
AL3	.70				.69							.57
AL4	.72				.76						.65	
AL5	.71				.87						.94	
AL6	.83				.87						.88	

*CLC = Closed-loop communication, BB = Backup behaviour, MPM = mutual performance monitoring, AL = Adaptive leadership

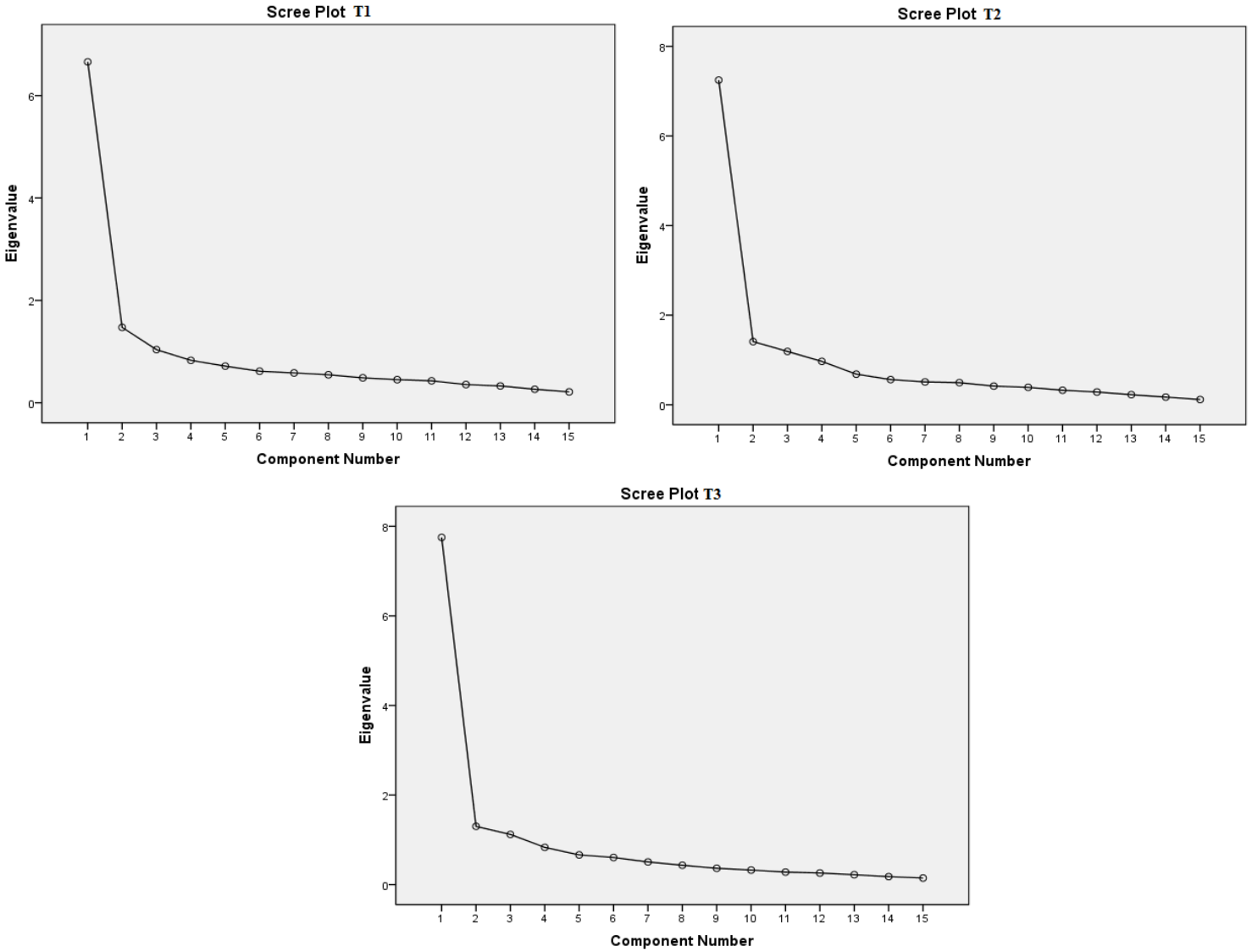

Figure C1: Teamwork behaviour, Principle Component Analyses, Direct Oblimin rotation, Scree-plots

All absolute scores < .4 were suppressed.

Extraction based on Eigenvalue > 1. For each item n = 119.

Table C2: Teamwork behaviour, Principle Component Analyses, Direct Oblimin rotation, Pattern Matrices

Item	Components T1			Components T2			Components T3		
	1	2	3	1	2	3	1	2	3
CLC1		.93			.71			-.86	
CLC2		.83			.49			-.84	
CLC3		.52			.78			-.81	
BB1		.65			.82			-.63	
BB2		.67			.77		.50	-.56	
BB3			.78			.72	.80		
MPM1		.43		.67			.49		
MPM2			.59		.47		.80		
MPM3	.48			.61			.63		
AL1			.54			.73	.77		
AL2			.55	.49			.54		
AL3	.56			.69			.45		
AL4	.64			.84					.50
AL5	.80			.90					.86
AL6	.79			.89					.80

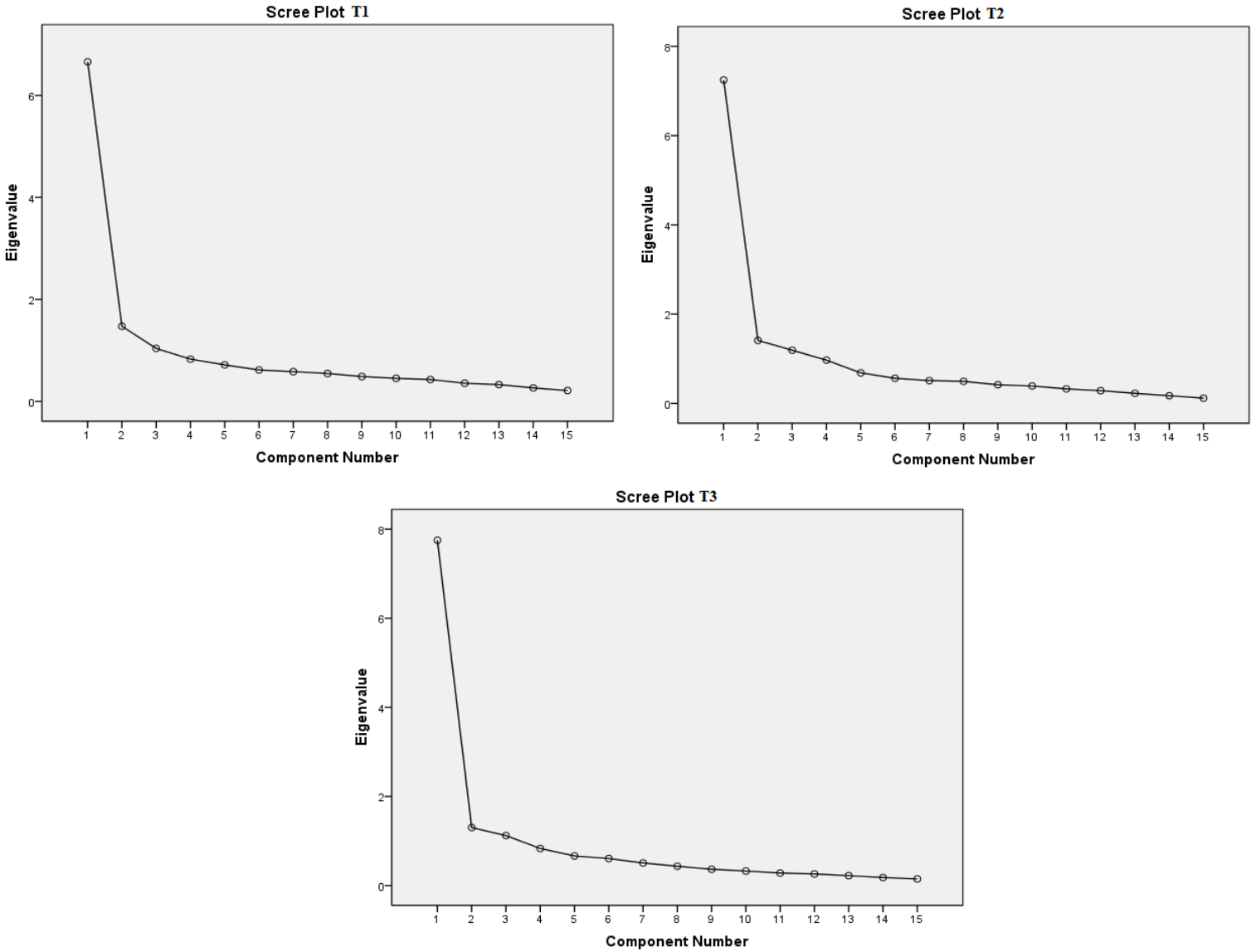

Figure C2: Teamwork behaviour, Principle Component Analyses, Direct Oblimin rotation, Scree-plots

Table C3: Team performance, Principle Component Analyses, Direct Oblimin rotation, Pattern Matrices

Item	Components T1	Components T2	Components T3
	1	1	1
TP1	.85	.92	.96
TP2	.84	.83	.91
TP3	.88	.92	.90
TP4	.86	.93	.93
TP5	.95	.92	.89

Appendix D: Micro-level percentage distribution

Table D1: 'Differences between condition'-focus

		Diabetes			Heroin		
Macro-level behaviour*	Micro-level behaviour	EG** (%)	CG (%)	Δ = EG – CG (%)	EG (%)	CG (%)	Δ = EG – CG (%)
SA	Request for information	42.0	40.9	1.1	34.1	32.2	1.9
	Providing information upon request	7.3	10.5	-3.2	12.1	10.3	1.8
	Obtaining unsolicited task-relevant information	50.7	48.6	2.1	53.8	57.5	-3.7
CLC	Verifying information	12.1	12.4	-.3	16.4	10.9	5.5
	Verbal acknowledgment	16.3	24.0	-7.7	25.2	24.9	.3
	Non-verbal acknowledgment	24.3	21.5	2.8	26.4	20.7	5.7
	Verbalizing interpretation of the situation	37.1	29.2	7.9	29.0	37.3	-8.3
	Verbalizing own behaviour	10.2	12.9	-2.7	2.9	6.2	-3.3
AL	Giving orders	46.1	33.3	12.8	71.2	55.8	15.4
	Assigning tasks	8.7	7.0	1.7	7.2	6.6	.6
	Making plans	16.4	20.2	-3.8	6.10	14.0	-7.9
	Initiating action	28.8	39.5	-10.7	15.6	23.6	-8
BB	Providing unsolicited task-relevant action	6.3	8.1	-1.8	6.1	6.2	-.1
	Assistance request	9.5	6.7	2.8	5.4	7.9	-2.5
	Providing action upon request	26.6	21.6	5	26.9	20.8	6.1
	Unsolicited task-relevant action	22.7	36.0	-13.3	34.1	44.5	-10.4
	Verbally offering assistance	.2	.3	-.1	.7	.5	.2
	Interaction with external team conditions	34.7	27.3	7.4	26.8	20.1	6.7
MPM	Questioning information	3.4	2.1	1.3	0	4.9	-4.9
	Questioning actions	0	1.1	-1.1	0	0	0
	Watching actions of other team members	61.5	71.4	-9.9	86.2	90.2	-4
	Correcting actions of other team members	0	0	0	0	0	0
	Giving feedback	35.1	25.4	9.7	13.8	4.9	8.9

*CLC = Closed-loop communication, BB = Backup behaviour, MPM = mutual performance monitoring, AL = Adaptive leadership

** EG = experimental group, CG = control group

Table D2: 'Difference between scenario'- focus

		Experimental group			Control group		
Macro-level behaviour*	Micro-level behaviour	D** (%)	H (%)	Δ = D – H (%)	D (%)	H (%)	Δ = D – H (%)
SA	Request for information	42.0	34.1	7.9	40.9	32.2	8.7
	Providing information upon request	7.3	12.1	-4.8	10.5	10.3	.2
	Obtaining unsolicited task-relevant information	50.7	53.8	-3.1	48.6	57.5	-8.9
CLC	Verifying information	12.1	16.4	-4.3	12.4	10.9	1.5
	Verbal acknowledgment	16.3	25.2	-8.9	24.0	24.9	-.9
	Non-verbal acknowledgment	24.3	26.4	-2.1	21.5	20.7	.8
	Verbalizing interpretation of the situation	37.1	29.0	8.1	29.2	37.3	-8.1
	Verbalizing own behaviour	10.2	2.9	7.3	12.9	6.2	6.7
AL	Giving orders	46.1	71.2	-25.1	33.3	55.8	-22.5
	Assigning tasks	8.7	7.2	1.5	7.0	6.6	.4
	Making plans	16.4	6.10	10.3	20.2	14.0	6.2
	Initiating action	28.8	15.6	13.2	39.5	23.6	15.9
BB	Providing unsolicited task-relevant action	6.3	6.1	.2	8.1	6.2	1.9
	Assistance request	9.5	5.4	4.1	6.7	7.9	-1.2
	Providing action upon request	26.6	26.9	-.3	21.6	20.8	.8
	Unsolicited task-relevant action	22.7	34.1	-11.4	36.0	44.5	-8.5
	Verbally offering assistance	.2	.7	-.5	.3	.5	-.2
	Interaction with external team conditions	34.7	26.8	7.9	27.3	20.1	7.2
MPM	Questioning information	3.4	0	3.4	2.1	4.9	-2.8
	Questioning actions	0	0	0	1.1	0	1.1
	Watching actions of other team members	61.5	86.2	-24.7	71.4	90.2	-18.8
	Correcting actions of other team members	0	0	0	0	0	0
	Giving feedback	35.1	13.8	21.3	25.4	4.9	20.5

*CLC = Closed-loop communication, BB = Backup behaviour, MPM = mutual performance monitoring, AL = Adaptive leadership

** D = Diabetes scenario, H = Heroin scenario