

J Transp Secur (2013) 6:43–57  
DOI 10.1007/s12198-012-0102-y

---

## Crew goal setting for security control

Olive Emil Wetter · Franziska Hofer · Klaus Jonas

Received: 13 November 2012 / Accepted: 26 November 2012 / Published online: 11 December 2012  
© The Author(s) 2012. This article is published with open access at Springerlink.com

**Abstract** This study investigated the effectiveness, efficiency, and robustness of simple goal setting in airport security control. As outcome, crew performance in terms of productivity ([Experiment 1, field setting](#)) was studied. Furthermore, the moderating role of negative and positive priming due to a previous task on the impact of goals ([Experiment 2, laboratory setting](#)) was analyzed. This research builds a bridge from goal setting theory to practice and prepares the grounds for its application in security or emergency organizations. In Experiment 1, supervisors of Security Officers at a large European airport communicated goals to their subordinates without any prior intervention or training. Goals were applied to a short “peak” time span (40 min). Dependent variables were objective team-level measures of productivity, namely passenger density and throughput. Experiment 2 featured two different tasks that primed speed (negative priming: puzzle; positive priming: car race). The Frankfurter Aufmerksamkeitsinventar served as main task for obtaining speed and accuracy measures. The results show that pre-intervention goal setting can be used easily and effectively by supervisors to increase subordinates’ team performance during short interventions. Goal setting for short time spans is effective even without providing feedback. However, negative priming by a previous task may undermine the beneficial effects of goal setting.

**Keywords** Goal setting · Priming · Leadership · Airport security control · Crew performance

---

O. E. Wetter (✉) · F. Hofer  
Kantonspolizei Zürich (Zurich State Police), Airport Division, Research & Development,  
CH-8058 Zurich-Airport, Switzerland  
e-mail: weto@kapo.zh.ch

K. Jonas  
Social and Business Psychology, University of Zurich, CH-8050 Zurich, Switzerland

## Introduction

As air traffic has increased remarkably over the last years (e.g., >23 % increase in departing passengers at Swiss airports throughout the years 2004–2009, Swiss Federal Statistical Office 2010), so has the size of security providers at airports. The entity responsible for security control at European airports now often counts more than 1000 employees. Given this size, it becomes evident that management of the workforce has to be effective, efficient, and robust with regard to the specific conditions in the workplace. As a consequence of the increase in air traffic, security processes have become a significant bottleneck at some airports. Especially in Europe, many airports do not have enough space to physically enlarge their checkpoints. Therefore, apart from the abundance of research that had analyzed security control performance regarding threat detection only (e.g., Basner et al. 2008; Hofer and Schwaninger 2004; Schwaninger et al. 2010; Van Wert et al. 2009; Wales et al. 2009), a new additional focus on productivity or efficiency of security controls has been added in the past few years. Wetter et al. (2011) have reflected this development by stating that productivity is another important objective of security control apart from “just” providing security.

The impact of leadership in general on subordinates’ performance in teams and crews (e.g., Guzzo and Dickson 1996; Zaccaro et al. 2001) is well known. Optimizing leadership thus promises to be both an effective and cost-saving solution to increase performance of the human workforce. Earlier research has suggested that goal setting can be applied as a management tool for security control on an individual (Wetter et al. 2012) and an organizational level (Wetter et al. 2011). Goal setting theory by Locke and Latham (1990), including the high performance cycle (Latham and Locke 2007), is one of the most influential theories in work and organizational psychology. Its postulated effects have been shown in many laboratory and field studies (Locke and Latham 2002). However, several questions for its implementation on *crew level* in security organizations still remain to be investigated. Such organizations often operate under strong constraints regarding time and money. Security has to be provided reliably and within a certain limited time frame. In particular, the time available for interaction of supervisors with their staff, as in briefings, is often confined to a few minutes per shift. It is furthermore unclear how performing subtasks of different nature in short sequences, as it is common in security control, moderates the effects of goal setting.

In organizations, “management by objectives” (e.g., Drucker 1974, 2007; Kopelman 1986; Odiorne 1978) has become popular. Usually, goals are set for a predefined specific time period (6–12 months). During this period, goals are in place, employees work with them, and data are collected. After this time period, goal progress is evaluated together with the supervisor, and, subsequently, a new goal is set for the next time period. With respect to the specific work characteristics in security control, we studied crew level goals that are valid for a short time only (for an intervention that lasts 40 min). Some of the most important newer developments in goal setting research include indeed group goal setting (e.g., Haslam et al. 2009), as well as priming with nonconscious goals (e.g., Eitam et al. 2008). Latham and Locke (2007) stated that the study of the effects of nonconscious goals would clearly warrant additional studies. Indeed, the given example of security control unites both of these

aspects: goal setting in security control is often group goal setting because the service at the checkpoint is actually provided by a crew. Moreover, as security control often includes job rotation in intervals of 20–30 min, it can be postulated that previous subtasks can act as primes for the next subtask to be performed.

Given the above knowledge, the present study is based on the three following hypotheses:

*Hypothesis 1: Goal setting increases crew performance in short interventions.* This hypothesis aims at the *effectiveness* of goals on crew performance.

*Hypothesis 2: Significant goal setting effects can be achieved with minimal investments of briefing time.* This hypothesis aims at the *efficiency* of crew goal setting.

*Hypothesis 3a: Negative priming due to a previous task decreases the impact of a goal. Hypothesis 3b: Positive priming due to a previous task enhances the impact of a goal.* This hypothesis aims at the *robustness* of goal setting with regard to specific conditions in the workplace.

This study provides new knowledge in three main aspects: First, it examines the effects of goal setting on crew performance in security control by explicitly setting a goal that is valid for a defined short period only (about 40 min). The design allows for an evaluation of the effectiveness of only the first part of the common goal setting process, i.e., goal setting and first performance by the employee, without feedback and subsequent adjustment of behavior by the employee or even multiple feedback loops. Crew performance was measured using two true crew-level measures, namely, passenger density and throughput in security control lanes. Second, the effects of simple goal setting by supervisors, i.e., without any instruction, training, or other interventions that cost time and money, are analyzed. Third, the moderating role of negative and positive priming due to different subtasks in job rotation is investigated.

## Experiment 1 (field setting)

### Method

#### *Setting and participants*

This experiment was conducted at the security control of a large European airport that is an airline hub. Data were collected from early February to mid March 2010. To ensure highest possible standardization, data collection took place always on the same weekdays during the morning peak hour. A total of six security control screening lanes, i.e., all lanes for “economy” passengers in the most important sector at the airport, were studied. The lanes all featured the same technical equipment.

The security control process basically requires the passenger to put the belongings on the roller table, pass the walk-through metal detector (WTMD), and collect the belongings again after X-ray inspection. The security control is carried out by Security Officers (SOs) between 20 and 65 years of age, that are employed by the police. Although SOs are not police officers, the work environment is embedded in police structures. All SOs and supervisors are trained and certified according to the

actual standards by the country's regulatory authority. Crew composition is ad hoc, changes on a daily basis, and is random with the exception that some criteria are followed (e.g., both sexes have to be represented in a crew). Allocation of crews to the studied sector was fully random. As a consequence, crews were randomly assigned to the three different experimental conditions (see paragraph on [variables](#) below).

Each crew operates one screening lane and consists of four SOs and one crew leader. In the studied sector of the airport, two supervisors are responsible for controlling and problem solving and have higher decision making authority. One of them is responsible for holding pre-shift briefings. This experiment was conducted during standard daily operation in order to ensure maximal ecological validity.

### *Variables*

The experiment featured one independent variable, namely, goal setting, with three different conditions: one baseline and two experimental conditions, which we labeled “do your best” and “specific difficult goal.” In the baseline condition, no intervention was made and no goal communicated to the SOs. This condition reflects standard daily operation with the only exception that two observers were present behind the six screening lanes (the same in all three conditions). The “do your best” condition used a “do your best” goal as the intervention. Prior to the shift, staff from the organization's research & development (R&D) unit contacted the supervisor in charge of the pre-shift briefing for the crews that were studied subsequently. The supervisor was given a memo stating the following: “This morning it is our goal that everybody does their best in order to achieve excellent job performance” (translated from the German). The supervisor was asked to communicate this information to the crews at the end of the pre-shift briefing.

The “specific difficult goal” condition used a specific, difficult goal as the intervention. The supervisor was given a memo stating the following: “In today's morning peak, in addition to the compliance with security regulations as usual, it is our goal to ensure that there are always 8 passengers simultaneously in a screening lane. We, as members of the present shift, can occasionally monitor that ourselves” (translated from the German). In other words, crews were told that they can monitor their own goal attainment to a certain extent, if they like. Again, the supervisor was asked to communicate this instruction to the crews at the end of the pre-shift briefing. This goal was chosen because passenger density is important (see paragraph on passenger density below), because it is specific (it contains a number), and because it is known to be difficult but realistic based on a preliminary study by Wetter et al. (2010). This preliminary study demonstrated that, given this particular airport's infrastructure, this number is usually lower (around 6), but can reach 8 if crews are motivated and work together well.

Crew performance was assessed in terms of productivity as dependent variable. Two productivity measures were recorded, namely passenger density and throughput (for a definition, see Wetter et al. 2010). Passenger density is a measure of the degree of infrastructure capacity utilization. It receives more and more attention since it denotes how well the space of a given security control infrastructure is actually used. As every  $m^2$  in an airport is costly, security control would ideally be limited to as many  $m^2$  as necessary, with as little “empty” or unused area inside as possible. In a given infrastructure, higher passenger density indicates that less space is occupied per passenger, i.e., that the available space is better utilized. Lower than optimal

passenger density would be considered as “wasting space” and economically inefficient, whereas higher than optimal passenger density could lead to disorderly, chaotic situations, which in turn could lower the security level, increase employees’ work strain, and negatively affect passenger satisfaction.

Passenger density can be defined as the number of passengers present simultaneously in one screening lane at one moment. The number of passengers present simultaneously is calculated as follows: The first passenger to be counted is the one who is at the very beginning of the security control process, i.e., the one who does the first action (usually touching the baggage tray). The last one to be counted is the one who is at the very end of the security control process (usually the one stepping away from the control lane). All passengers within this process, for example passengers loading their hand baggage into the tray, passing the WTMD, undergoing a pat-down search, or collecting their belongings, are counted.

To ensure a certain minimum distance between measurements, a maximal sampling frequency of 1/min was defined. Sampling was done visually by two observers who were present behind the security control screening lanes. Each of those observers sampled data from three security control lanes sequentially, i.e., started at lane one and sampled, then moved on to lane two and sampled, then moved on to lane three and sampled, then began again with lane one. Like this, a certain number of samplings of passenger density per control lane resulted after the measurement period. Subsequently, those were averaged into one value of passenger density per control lane per measurement period.

The second productivity measure was passenger throughput. Throughput is very important for different stakeholders, because it is used for capacity calculations and for infrastructure and personnel planning. It is also important for ensuring certain minimum connecting times for transferring passengers and for the calculation of queuing times for security control. If throughput is neglected, “airlines or airports whose security screening procedures are particularly time consuming can expect to lose business to their competitors” (Gilliam 1979, p. 117). Throughput is a widely recognized measure in the applied field and is used for benchmarking between European airports. We propose the following definition and operationalization of passenger throughput: Passenger throughput denotes the maximum capacity of a security control lane in passengers/h.

For both passenger density and throughput, it must be ensured that during measurement there are more passengers available than the number that the security control lane can actually handle. If the number of passengers is lower than the capacity of the security control lane, the measure only denotes the number of passengers and is not an indicator of maximal capacity. Therefore, measurements must only take place if there is a queue of passengers, i.e., at least one passenger queuing up and waiting in front of the lane. Throughput was measured automatically using the relevant function of the WTMD’s standard commercial software.

Goal setting is expected to affect passenger density because density can be controlled by the SOs themselves via the way they work and guide passengers. However, it is unclear if it will also affect throughput. Apart from employee behavior, throughput significantly depends on external factors as well, such as “number of manual baggage inspections to be done” and “day temperature” (Wetter et al. 2010). Throughput was chosen as second measure for productivity because it is known to

correlate with passenger density (Wetter et al. 2010) and because it is widely used. Both are true crew-level measures—that is, one value per crew is generated. The two dependent variables were collected for all crews present (population level) in the respective sector during the respective time interval.

In addition to these measured variables, employees were asked to recall the goal after the peak. They were asked the question, “Evaluation of the briefing: What goal were you assigned by the supervisor at today’s briefing?” Possible answers were “There was no goal assigned” (with a check box) or “The following goal was assigned:” (with an empty line for writing in the goal). In this way, it was assessed whether SOs thought a goal had been assigned, and if yes, what they thought the goal had been.

### *Procedure*

Prior to the work shift, the supervisor in charge of the pre-shift briefing was told that staff persons from R&D would be present during the morning peak hour. For the “do your best” and the “specific difficult goal” condition, he was given a memo with the goal (see section “[Variables](#)”). He then communicated this information to the SOs at the briefing. During the briefing, staff persons from R&D were present as passive listeners. After the briefing, SOs opened the assigned security control lanes to passengers. As soon as there were enough passengers waiting, that is, when there were queues and no more gaps between passengers, measurements started. In order to measure passenger density, two people from R&D circulated behind the security control lanes, each of them observing three lanes. The measurement period for both, passenger density and throughput, stopped as soon as there were gaps between passengers approaching the security control lanes instead of continuous queues. Measurements lasted 39 min on average ( $M=39.34$  min,  $SD=9.42$  min). A few minutes afterwards, the goal recall questionnaire was administered to the SOs.

The baseline condition was devised in order to guarantee the same amount of observer presence and for recording and measuring passenger density in exactly the same way as in the conditions with goal setting. Thus, the procedure of the baseline condition was identical to the one described above with the only difference that no memo was given to the supervisor, and, as a consequence, no goal setting took place.

### Results for Experiment 1

#### *Manipulation check*

The observers always noted during the supervisor’s briefing if the instruction from the memo was in fact communicated to the crews, which served as manipulation check. This was always the case, thus, the success rate of the manipulation check being 100 %.

#### *Interrater reliability*

Interrater reliability for passenger density was established as follows: After measurements were finished, both observers (the first author being one of them) plus a third person who had not been involved in data collection measured the same security control lanes in an additional session. Thirty measurements were recorded

independently by these three observers, and they were subsequently compared. Ratings of the two “trained” observers who collected data for the experiment correlated with  $r=.94$ . The correlation between the ratings of observer 1 and the “new” observer was  $r=.92$ ; the correlation between observer 2 and the “new” observer was  $r=.91$ .

### Productivity

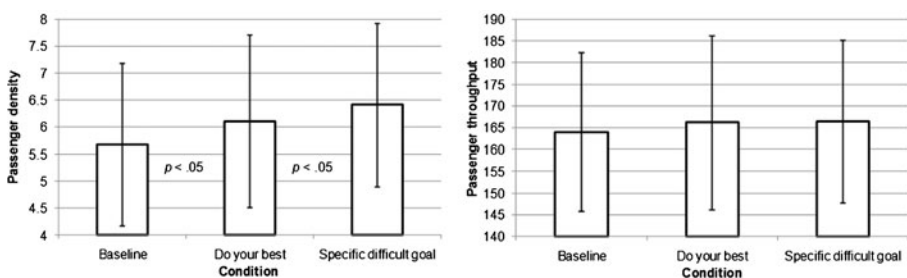
Because it makes sense to integrate both, passenger density and throughput, into a construct “productivity”, and because the preliminary study (Wetter et al. 2010) had demonstrated that they are correlated with  $r=.27$ , a MANOVA was calculated. The factor (between) was “Goal” (baseline, “do your best,” “specific difficult goal”). The dependent variable was “productivity” as a construct. The MANOVA revealed a significant effect of “Goal” with  $F(4, 196)=5.006$ ,  $p<.001$  and an effect size of  $\eta_p^2=.093$  on productivity. In order to find out which of the two dependent variables, that constitute the construct “productivity”, contributed to this result, univariate analyses were calculated.

Univariate tests revealed that “Goal” had a significant effect on passenger density with  $F(2, 98)=10.971$ ,  $p<.001$  and an effect size of  $\eta_p^2=.183$ . Simple contrasts revealed significant differences between all possible comparisons: baseline vs. “do your best” with  $p<.05$ , “do your best” vs. “specific difficult goal” with  $p<.05$ , baseline vs. “specific difficult goal” with  $p<.001$  (Fig. 1, left). On the other hand, there was no significant effect of “Goal” on throughput with  $F(2, 98)=0.177$ ,  $p=.838$  (Fig. 1, right).

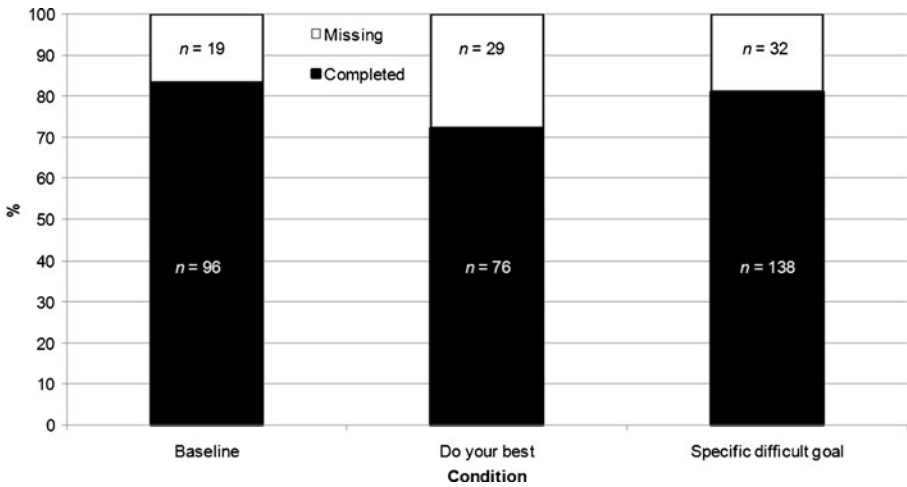
### Goal recall

The questionnaire containing the goal recall question was filled out individually by more than 79 % of the SOs overall. Figure 2 shows the exact numbers for the three conditions. In the baseline condition, the question (“What goal was assigned by the supervisor at today’s briefing?”) revealed that SOs had no goals in mind: Out of a total of  $N=96$  individual responses, 56 % cited a new regulation that had nothing to do with a goal, 24 % cited other information that had been communicated in the briefing, 19 % stated explicitly that no goal had been communicated, and 1 % reported “do not know.”

Since the dependent variables are crew-level measures, answers to the goal recall question for the “do your best” and the “specific difficult goal” condition were



**Fig. 1** Experiment 1: Means and standard deviations for passenger density (left) and throughput (right) for the three conditions

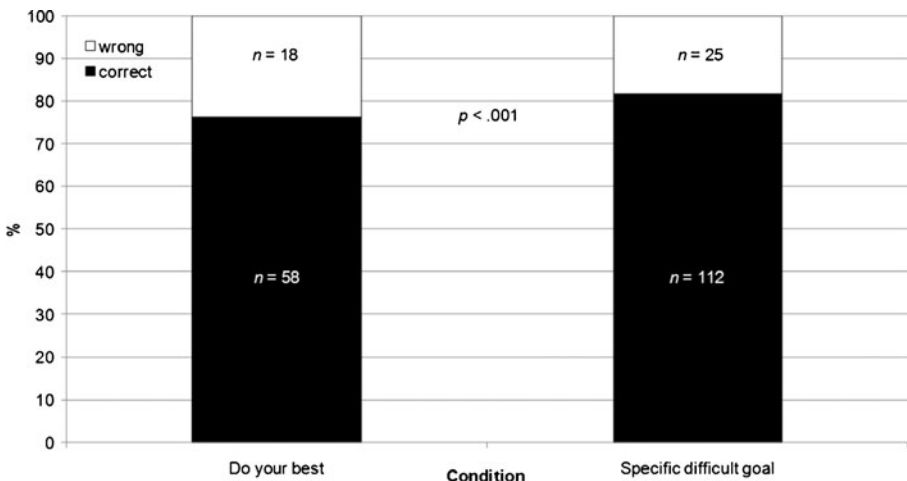


**Fig. 2** Experiment 1: Questionnaire completed and missing for the three conditions

categorized as either “correct” or “wrong” for the whole crew. They were regarded as “correct” if at least three out of five crew members stated the correct goal. Otherwise, they were regarded as “wrong” for the whole crew. On the individual level, correctness of the reported goal was judged conservatively. Figure 3 shows the goal recall performance in the “do your best” and the “specific difficult goal” conditions. Free recall of the goal succeeded significantly more often in the “specific difficult goal” condition than in the “do your best” condition, with  $\chi^2(1, N=137)=26.12, p<.001$ .

Discussion of Experiment 1

The results demonstrate that using this specific form of goal setting, i.e., goal setting for a short time span without prior instruction, training, or other interventions for supervisors, as well as without feedback for employees, increased productivity in an airport’s



**Fig. 3** Experiment 1: Goal recall performance for the “do your best” and “specific difficult goal” condition



security control. The specific difficult goal was most effective, followed by the “do your best” goal. Working without goal setting was least effective. The specific, difficult goal could more easily be recalled than the “do your best goal.” An open question is how effects of goal setting interact with other effects that can occur in the field. In order to investigate the interaction of goal setting with one such effect (priming due to job rotation), the following laboratory experiment was designed.

## Experiment 2 (laboratory setting)

### Method

#### *Setting and participants*

Experiment 2 was conducted under standardized conditions in a computer laboratory between April and July 2011. Participants were students of psychology, who were offered a 1-hour credit point, as well as volunteers who were offered CHF 15 for participation. After exclusion of 13 participants due to a negative result for the manipulation check, the data of 65 participants between 18 and 36 years of age ( $M=23.28$  years,  $SD=4.08$  years), 69 % of which were female, remained for analysis. Most participants were students (85 %) and a minority represented the working population (12 %). The rest of the participants followed other main activities or did not indicate their main activity. Assignment of participants to conditions was random. Table 1 provides an overview of the number of participants in each condition (one control and two experimental conditions). The three conditions are described in the next section.

#### *Procedure and variables*

During the first 6 min of the experiment, all participants completed the paper-pencil version of the Frankfurter Aufmerksamkeitsinventar (FAIR; Moosbrugger and Oehlschlägel 1996) as a baseline test. Test form B was used because its retest reliability is higher ( $r=.91$ ) than the retest reliability of form A ( $r=.85$ ). As dependent variables, speed and accuracy were measured. For measuring speed, the number of marked signs minus the number of line errors ( $G-F_L$ ) was calculated. For measuring accuracy, the number of missing errors plus the number of false alarm errors, multiplied by 2, were calculated  $((F_V + F_A) * 2)$ . After this first baseline measurement, the experimenter assigned the goal for the two experimental conditions (negative and positive priming) as follows: “In 20 min, I will ask you to solve the same

**Table 1** Experiment 2: Distribution of participants by conditions

Condition ( <i>n</i> )	Female ( <i>n</i> ; %)	Male ( <i>n</i> ; %)
Control (33)	25; 75.76	8; 24.24
Experimental “puzzle” (17)	10; 58.82	7; 41.18
Experimental “car race” (15)	10; 66.67	5; 33.33
Total (65)	45; 69.23	20; 30.77

test again. Your goal will then be to solve six lines more than you just did” (translated from the German). Thus, participants received a specific, difficult speed goal for the second testing. Participants in the control condition were only informed that they would solve the test again in 20 min (no goal).

During the next 20 min, an interim task had to be solved. For this task, a duration of 20 min was chosen because this is a typical duration for job rotation at a security control, where crew members change position in just that interval. During this time period, participants in the experimental condition “puzzle” played an online puzzle which served as a prime for accuracy; progress in a puzzle can only be made if one works accurately. The puzzle had 192 pieces, was easy to operate, and showed a picture of a sea landscape that was conceptually irrelevant to both speed and accuracy. The size of the puzzle was chosen so that no participant could finish in 20 min. Participants in the experimental condition “car race” played an online car race, which served as prime for speed. In order to unlock new tracks, participants had to rank first to third in the race. The car race was easy to operate. Participants in the control condition answered a questionnaire (adapted from Rolfa 2011) which featured questions about experiences and preferences which primed neither speed nor accuracy. Sample questions are “Have you ever visited a pyramid/been swimming at midnight/touched an iceberg?”, which had to be answered with yes or no and an additional rating (3-point scale) of how desirable participants found the respective experience. There were enough questions to keep participants busy. After 20 min, participants were interrupted in their activity and asked to either switch off the computer screens or to put aside the questionnaire.

The FAIR was then solved a second time. The same measures were collected as in the baseline measurement. Subsequently, the participants completed a questionnaire with seven domains. This questionnaire assessed goal recall (manipulation check), goal salience, goal commitment (three items), previous familiarity with the interim task, assumed underlying concept of the interim task (speed, accuracy, neither of both), fatigue, as well as some demographic variables.

## Results for Experiment 2

### *Manipulation and randomization checks*

A total of 13 participants had to be excluded based on the following reasons: Three participants did not solve the FAIR according to the instructions and had to be excluded because no values for speed and accuracy could be calculated. Five participants did not recall the goal correctly in the questionnaire and two participants did not answer this question at all. These participants were excluded because it is not certain if they had heard and understood the goal correctly. Furthermore, three participants in the experimental condition “car race” indicated that this interim task had accuracy (instead of speed) as an underlying concept. If, in these people, the car race should have primed accuracy rather than speed, the manipulation did not work as intended, leading to their exclusion from the sample.

In order to test whether the randomization of participants to conditions was successful, one-way ANOVAs with the factor “condition” (experimental “puzzle”, experimental “car race”, control) and the dependent variables baseline performance in

speed and accuracy, goal commitment, age, educational level, and level of exhaustion due to interim task were conducted. This analysis revealed no significant difference between participants in the different conditions regarding any of the dependent variables (all  $p$ s > .206 except educational level with  $p$  > .094). Randomization can thus be regarded as successful.

### Hypothesis testing

The descriptives are given in Table 2. For quantifying the speed improvement between the first and the second testing,  $\Delta$ s was introduced (right column).  $\Delta$ s denominates the difference between the speed measures from the second and the first testing. It was decided to include goal commitment, age, and educational level as covariates in the subsequent statistical analyses. Goal commitment has been shown in many studies to be an important moderator for goals to be effective (e.g., Hollenbeck and Klein 1987; Locke and Latham 1984). The goal commitment scale was adapted from the scale used by Wetter et al. (2012) and showed a reliability of Cronbach's  $\alpha$  = .61 (three items). Age was chosen as a second covariate since basic cognitive skills decline over time (Li et al. 2001) and because the realization of goals has been shown to depend upon cognitive mechanisms (Wetter et al. 2012). "Educational level" was included as a third covariate because it can be assumed that people with a higher educational level also perform better on cognitive tasks.

In order to test Hypothesis 3a, a mixed ANCOVA with the between-participants factor "condition" (experimental "puzzle", control) and the within-participants factor "measurement" (1st and 2nd testing) was calculated. The dependent variable was the FAIR's speed measure. As covariates, goal commitment, age, and educational level were included. The result shows a main effect of "measurement" ( $F(1, 45) = 48.62, p < .001, \eta_p^2 = .52$ ), i.e., participants were significantly faster in the second testing than in the first one. An interaction of "condition" and "measurement" shows that the speed improvement between the two testings was significantly lower for participants in the experimental condition "puzzle" than for participants in the control condition ( $F(1, 45) = 2.92, p < .05$ , one-tailed,  $\eta_p^2 = .06$ ). Thus, Hypothesis 3a is supported by the data. Further significant results of this analysis were an interaction of "measurement" and age ( $F(1, 45) = 4.96, p < .05, \eta_p^2 = .10$ ), as well as an interaction of "measurement" and goal commitment ( $F(1, 45) = 12.70, p < .001, \eta_p^2 = .22$ ).

In order to test Hypothesis 3b, a mixed ANCOVA with the between-participants factor "condition" (experimental "car race", control) and the within-participants factor "measurement" (1st and 2nd testing) was calculated. The dependent variable was the FAIR's speed measure. As covariates, goal commitment, age, and educational

**Table 2** Experiment 2: Speed measure in 1st and 2nd testing, as well as speed improvement ( $\Delta$ s)

Condition	1st testing ( <i>M</i> ; <i>SD</i> )	2nd testing ( <i>M</i> ; <i>SD</i> )	$\Delta$ s ( <i>M</i> ; <i>SD</i> )
Control	391.76; 98.91	520.06; 101.76	128.30; 50.23
Experimental "puzzle"	377.06; 77.32	491.35; 87.09	114.29; 36.96
Experimental "car race"	427.67; 79.94	560.00; 73.20	132.33; 47.89

level were included. The result shows a main effect of “measurement” ( $F(1, 43)=33.09$ ,  $p<.001$ ,  $\eta_p^2=.44$ ), i.e., participants were significantly faster in the second testing than in the first one. However, the higher improvement in speed between the two testings for participants in the experimental condition “car race” compared to participants in the control condition was not significant (“condition” X “measurement”,  $F(1, 43)=0.20$ ,  $p=.328$ , one-tailed). The effect size of this non-significant result would be very small ( $\eta_p^2=.005$ ) but pointing into the predicted direction. A power analysis shows that a sample size of  $N=1689$  participants would be necessary for being able to decide this question with sufficient power in the given design. Thus, in this experiment, we found no evidence for Hypothesis 3b. Furthermore, there was a significant interaction of “measurement” and goal commitment ( $F(1, 43)=4.40$ ,  $p<.05$ ,  $\eta_p^2=.09$ ).

## General discussion

Simple goal setting without feedback for a time span of about 40 min can significantly increase crew performance of SOs in terms of productivity at work. Experiment 1, conducted as field experiment at an airport’s security control, has shown that the tenet of goal setting theory, i.e., that specific difficult goals have an advantage over “do your best” goals and no goals, is also valid for teamwork settings in security. As shown by Experiment 2, this effect of goal setting might actually interact with other effects such as priming due to job rotation.

Experiment 1 has reliably shown effects of a very simple application of goal setting in a security control environment. Nevertheless, one variable of the construct “productivity”, namely throughput, was not significantly affected by goal setting. There are three possible reasons for this lack of significance. First, SOs were told or instructed to pay attention to throughput in neither the “do your best” nor in the “specific difficult goal” condition. However, they were instructed in the “specific difficult goal” condition to care for passenger density. Second, data from throughput have rather large standard deviations, which makes it more difficult to find significant effects. Third, throughput is significantly influenced by external factors, on which crew members have no influence, which is a major difference to passenger density. This limitation can be illustrated by the fact that the two external factors “number of manual baggage inspections to be done” and “day temperature” account for 25 % and 22 % respectively of the variations in throughput (Wetter et al. 2010).

The goal recall question might provide additional explanation for variance in the effectiveness of the different goals: The specific, difficult goal (including the number “8”) could more easily be recalled after the morning peak hour than the abstract “do your best goal.” Findings from cognitive psychology strongly support this idea: Walker and Hulme (1999) showed that concrete words are easier to recall than abstract words. We would thus like to suggest that a goal with a specific number can be more easily retained in memory and be subsequently recalled than an unspecific goal.

But what about social facilitation (Zajonc 1965) or the “Hawthorne Effect” (Landsberger 1958) as alternative explanations for the results? In order to prevent those from coming into effect in our experiment, it was ensured that there was exactly

the same amount of presence of supervisors and observers in all conditions including the baseline. The drawback of this is that possible overall reactivity, caused by observer presence, cannot be ruled out. However, the presence of observers was necessary in order to measure passenger density.

Locke et al. (1988, p. 23) stated, “if there is no commitment to goals, then goal setting does not work.” Similar to that, Latham and Yukl (1975, p. 824) write, “goals that are assigned to a person (e.g., by a supervisor) have an effect on behavior only to the degree that they are consciously accepted by the person.” French et al. (1966) pointed out, however, that this relation was only observed when employees had a past history of participation. In both our experiments, we used an authoritative way of setting goals without offering participation to supervisors, employees, or participants. Regarding police and security work, where authoritative goal setting is predominant, it can be argued that our setting was ecologically valid. Moreover, Locke and Latham (1990) found a correlation of  $r=.58$  between set goals and personal goals, suggesting that a significant part of authoritatively set goals will later be accepted and treated as personal goals.

Experiment 2 suggests that supraliminal negative priming caused by a previous task (“puzzle”) may significantly impair the effect of goal setting. Descriptive results of the condition with positive priming (reinforcement due to playing a car race) of the goal did not allow for an interpretation. One possible explanation for the lacking effect of positive priming could be that the car race did not sufficiently prime speed: 44.44 % of the participants did not perceive the car race as speed relevant; probably because it lasted for a fixed duration. This could have reduced the positive priming effect in a significant number of participants. It can thus be concluded that positive priming should be investigated with a different design and different stimuli in order to find out more about its potential effects.

The outcome of Experiment 2 is in line with previous findings about subconscious goals. Stajkovic et al. (2006) have shown that subconscious goals, activated through priming, interact with conscious goals on the same performance task. Latham et al. (2010) reviewed the latest findings on subconscious goals in the workplace. According to them, it is an unanswered question which methods of priming are effective. Our research suggests that previous activities or tasks, as in job rotation, may trigger priming. Even though our manipulation in Experiment 2 was not aimed at activating subconscious goals, this mechanism could explain the results. For application, though, it only matters that engaging in a certain task may act as priming and thus significantly moderate the effects of (conscious) goal setting on a later task.

The materials and tasks used in Experiment 2 are conceptually close to tasks in security and emergency services. The FAIR consists of a visual discrimination task, requiring participants to recognize relevant information and disregard irrelevant or distracting information. This is conceptually similar to the baggage screening task at a security control or a visual check regarding ticket validity at an entrance. We believe that the following suggestions for practice could thus be deducted: Similar tasks should be grouped together in order to minimize alternations between dissimilar tasks, thus minimizing possible negative priming effects. Furthermore, time gaps between goal setting and the goal-relevant task should be avoided. Nevertheless, further research about the combination of goal setting and job rotation using different designs is highly warranted.

The present experiment demonstrated that simple goal setting can be used at briefings in order to enhance work performance during short interventions. Even though we have operationalized work performance as productivity in this study in order to provide a new focus, we of course strongly advocate that goals be used sensibly by supervisors and managers. In security, this means that the most important goal, i.e., providing security, must not be neglected and must be given the appropriate weight in the process of goal setting. In practice, an often heard argument is that providing security and ensuring a certain level of productivity contradict each other. Regarding this aspect, Wetter et al. (2012) have shown that the risk for inducing speed-accuracy trade-offs can be reduced by setting goals that do not overload employees or even by setting dual goals. Goal setting is easy to use, efficient, effective, and robust, provided the tenets of the goal setting theory are followed. It can thus be recommended to install pre-intervention goal setting as a meaningful complement to programs such as “management by objectives”, which focus on longer time periods.

**Acknowledgments** We thank MSc Mirjam Baur for her contribution to Experiment 2.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.

## References

- Basner M, Rubinstein J, Fomberstein KM, Coble MC, Ecker A, Avinash D, Dinges DF (2008) Effects of night work, sleep loss, and time on task on simulated threat detection performance. *Sleep* 31:1251–1259
- Drucker P (1974) *Management: Tasks, responsibilities, practices*. Butterworth-Heinemann, Oxford
- Drucker P (2007) *The practice of management* (Classic Drucker Collection edition). Butterworth-Heinemann, Oxford
- Eitam B, Hassin RR, Schul Y (2008) Nonconscious goal pursuit in novel environments: the case of implicit learning. *Psychol Sci* 19:261–267
- French JR, Kay E, Meyer HH (1966) Participation and the appraisal system. *Hum Relat* 19:3–19
- Gilliam RR (1979) An application of queuing theory to airport passenger security screening. *Interfaces* 9:117–122
- Guzzo RA, Dickson MW (1996) Teams in organizations: recent research on performance and effectiveness. *Annu Rev Psychol* 47:307–338
- Haslam SA, Wegge J, Postmes T (2009) Are we on a learning curve or a treadmill? The benefits of participative group goal setting become apparent as tasks become increasingly challenging over time. *Eur J Soc Psychol* 39:430–446
- Hofer F, Schwaninger A (2004) Reliable and valid measures of threat detection performance in X-ray screening. In: Sanson LD (ed) *Proceedings of the 38th Carnahan Conference on Security Technology*. Institute of Electrical and Electronics Engineers, Albuquerque, pp 303–308
- Hollenbeck JR, Klein HJ (1987) Goal commitment and the goal-setting process: problems, prospects, and proposals for future research. *J Appl Psychol* 72:212–220
- Kopelman RE (1986) *Managing productivity in organizations*. McGraw-Hill, New York
- Landsberger HA (1958) *Hawthorne revisited: Management and the worker: Its critics, and developments in human relations in industry*. Cornell University, Ithaca
- Latham GP, Locke EA (2007) New developments in and directions for goal-setting research. *Europ Psychol* 12:290–300
- Latham GP, Yukl GA (1975) A review of research on the application of goal setting in organizations. *Acad Manage J* 18:824–845
- Latham GP, Stajkovic AD, Locke EA (2010) The relevance and viability of subconscious goals in the workplace. *J Manage* 36:234–255

- Li SC, Lindenberger U, Skiström S (2001) Aging cognition: from neuromodulation to representation. *Trends Cogn Sci* 5:479–486
- Locke EA, Latham GP (1984) Goal settings for individuals, groups, and organizations. Science Research Associates, Chicago
- Locke EA, Latham GP (1990) A theory of goal setting and task performance. Prentice Hall, Englewood Cliffs
- Locke EA, Latham GP (2002) Building a practically useful theory of goal setting and task motivation. *Am Psychol* 57:705–717
- Locke EA, Latham GP, Erez M (1988) The determinants of goal commitment. *Acad Manage Rev* 13:23–39
- Moosbrugger H, Oehlschlägel J (1996) FAIR. Frankfurter Aufmerksamkeitsinventar [Frankfurt Attention Inventory]. Huber, Bern
- Odiorne GS (1978) MBO: A backward glance. *Bus Horizons* 21:14–24
- Office SFS (2010) Mobilität und Verkehr 2010 [Mobility and traffic 2010]. Swiss Federal Statistical Office, Neuchâtel
- Rolfa (2011) 200 Dinge, die man im Leben gemacht haben sollte [200 things you should have done in your life]. Retrieved from <http://rolfa.ch/200dinge>
- Schwaninger A, Hardmeier D, Riegelnic J, Martin M (2010) Use it and still lose it? *GeroPsych* 23:169–175
- Stajkovic AD, Locke EA, Blair ES (2006) A first examination of the relationships between primed subconscious goals, assigned conscious goals, and task performance. *J Appl Psychol* 91:1172–1180
- Van Wert MJ, Horowitz TS, Wolfe JM (2009) Even in correctable search, some types of rare targets are frequently missed. *Atten Percept Psychophys* 71:541–553
- Wales AWJ, Anderson C, Jones KL, Schwaninger A, Horne JA (2009) Evaluating the two-component inspection model in a simplified luggage search task. *Behav Res Meth* 41:937–943
- Walker I, Hulme C (1999) Concrete words are easier to recall than abstract words: evidence for a semantic contribution to short-term serial recall. *J Exp Psychol* 25:1256–1271
- Wetter OE, Lipphardt M, Hofer F (2010) External and internal influences on the security control process at airports. In: Pritchard DA, Sanson LD (eds) Proceedings of the 44th International Carnahan Conference on Security Technology. Institute of Electrical and Electronics Engineers, San José, pp 301–309
- Wetter OE, Fuhrmann H, Lipphardt M, Hofer F (2011) Bringing adversaries together: The importance of a common management-level approach in complex work domains. In: Faundez-Zanuy MS, Espinosa-Duró V, Sanson LD (eds) Proceedings of the 45th International Carnahan Conference on Security Technology. Institute of Electrical and Electronics Engineers, Mataró, pp 20–25
- Wetter OE, Wegge J, Jonas K, Schmidt KH (2012) Dual goals for speed and accuracy on the same performance task: Can they prevent speed-accuracy trade-offs? *J Pers Psychol* 11:118–126
- Zaccaro SJ, Rittman AL, Marks MA (2001) Team leadership. *Leadership Quart* 12:451–483
- Zajonc RB (1965) Social facilitation. *Science* 149:269–274