rijksuniversiteit
groningen

Personality traits in the Dutch Special Forces: Zooming in on *how* recruits fill in questionnaires to detect discipline and dedication

Laura Haasper

Master Thesis – Talent Development & Creativity

s3146553
2022/01/23
Department of Psychology
University of Groningen
Supervisors:
Rik Huijzer & dr. Ruud den Hartigh
Second evaluator:
dr. Barbara Huijgen

Abstract

Special forces operators are constantly working in intense and unpredictable environments all over the world. In order to be successful, their tasks need to be carried out with discipline and dedication. Therefore, this study will investigate how disciplined and dedicated the recruits of the Korps Commandotroepen of the Dutch military are by assessing how they fill in a long personality questionnaire at the start of the selection period. Specifically, we compared successful recruits ($n = 35$) and the ones who drop out of the program ($n = 213$) on *how* they fill in questionnaires in terms of differences in the response times and stability in their answer patterns. We expected successful recruits to display quicker and more stable response times between subsequent items, as they are likely to be disciplined and dedicated. However, we did not find any statistical differences between the groups in the emerging clusters of response time and stability. We further expected successful recruits to display more consistent answer patterns on the different facets of the Big Five personality traits. Categorical recurrence quantification analyses revealed no difference between the graduates and non-graduates. Nevertheless, moving from exclusively analyzing outcomes of global measures to assessing actual behavior of the recruits may be an important avenue for better selection procedures in the future.

Keywords: answer patterns, Big Five, Korps Commandotroepen, military, response times

Personality traits in the Dutch Special Forces: Zooming in on *how* recruits fill in questionnaires to detect discipline and dedication

Military personnel operates in highly demanding situations and environments. Special forces operators are not only trained at land, but also in the sea and in the air, requiring a variety of knowledge and skills in different habitats. These trainings are highly demanding and aimed at preparing the recruits for every imaginable situation. For example, commandos have to work underwater, shoot, and climb in the mountains, while being exposed to different climates and smells. Special forces operators typically operate in a small team (i.e., 16 commandos; Obringer & Guzman, 2006), behaving strategically under every circumstance while being constantly under enormous pressure and physical suffering (e.g., sleep deprivation). Therefore, only few individuals with very specific characteristics are fit for these tasks. Military personnel across the globe are characterized as having certain personality traits in order to work in such demanding environments and under those special circumstances. It is common that many potential recruits drop out during the intense training because they cannot withstand the high demands (Vaara et al., 2020). The Korps Commandotroepen are the Special Forces of the Dutch military. The operators of the Korps Commandotroepen are a prestigious unit with a broad spectrum of tasks, including providing military assistance to foreign military units, special reconnaissance (i.e. target and threat assessment), direct action, all applied to counter-insurgency, counter-terrorism, hostage release operations and faction liaison operations. The tasks of the Korps Commandotroepen change rapidly, alongside current developments around the globe (Griffith, 2009). Therefore, commandos of the Korps Commandotroepen have to be prepared for anything as the missions they face are extremely uncertain (i.e., where, when and under what circumstances they will be deployed in the future). Hence, the teams are constantly on standby and always ready to be dispatched.

In order to become a member of this elite unit, recruits are facing fifteen weeks of intensive training, after they have successfully completed the first selection procedure.

Previous military training is a prerequisite to apply in order to ensure a basic fitness and skill level (e.g., shooting abilities). Before the recruits are able to start the fifteen weeks of training, they are supposed to spend multiple hours filling in various extensive questionnaires regarding personality traits, intelligence, self-efficacy, commitment, coping, optimism/pessimism, mental toughness, and resilience. The aim of the current study is to assess *how* military recruits fill in the long personality questionnaires in terms of consistency and pace. This information could provide insights into the differences in discipline and dedication between successful recruits and dropouts. In order to do this, we examined the response times and the answer patterns to the individual facets of the Big Five personality questionnaire (Costa & McCrae, 1992).

Personality Traits in the Military

Personality is a widely used construct to understand how people behave in different situations (Buss, 1989). Specifically, the Big Five personality traits (i.e., neuroticism, extraversion, openness, agreeableness and conscientiousness) are important assets when it comes to assessing military personnel (Costa & McCrae, 1992). Each of the five personality traits consists of six facets which in turn consist of eight questions assessing that specific facet. Neuroticism is defined as generalized anxiety, meaning that people who score high in this domain are prone to feel anxious. Notably, neuroticism and emotional stability describe the same construct and are used interchangeably (Hills & Argyle, 2001). High scores on extraversion indicate that an individual is sociable and enjoys being around others. If an individual scores high on the domain of openness (to experience) it means that this individual is typically curious (e.g., willing to try out new things). Individuals who score high on agreeableness are perceived as kind, sympathetic, cooperative, warm and considerate (Rothmann & Coetzer, 2003). Conscientiousness is defined as being virtuous and rule-abiding (Hill & Roberts, 2011).

Previous research indicates that successful operators in the special forces show similar patterns of certain personality traits, namely low scores of neuroticism and agreeableness as well as high scores of extraversion and conscientiousness (Braun et al., 1994). This is partly in line with previous findings on successful leadership in the military. Successful leaders in the military score lower on neuroticism and higher on extraversion, openness, agreeableness, and conscientiousness compared to unsuccessful leaders (Johnson & Hill, 2009, McDonald et al., 1990, Campbell et al., 2010). Furthermore, high scores of emotional stability (i.e., low neuroticism) and low scores on extraversion increase the probability for success in the military (Hartmann et al., 2003). Interestingly, open-mindedness has also been found to be an important trait for Norwegian military operators (Boe et al., 2017). However, it should be noted that the personality traits may not remain stable for special forces operators. In fact, special forces operators with much deployment experience show an increase in emotional stability over time (Skoglund et al., 2020). A shift in the opposite direction occurs in agreeableness. People without deployment experience show higher scores on agreeableness compared to operators with deployment experience (Skoglund et al., 2020). This decline in agreeableness seems to intensify with every deployment as it has been found that operators with five or more deployments show a drastic decrease in agreeableness (Skoglund et al., 2020). Overall, it seems that personality traits like neuroticism serve an important purpose when it comes to operating in a small group rather than alone (Halfhill et al., 2005). Thus, psychological characteristics are generally found to be associated with training success in the military.

Besides the psychological characteristics, the traits discipline and dedication are known to be important for successful military personnel, as they are displaying the so-called military mindset (Akrami et al., 2007; Hall, 2011). Discipline is described by how willing someone is to obey to orders and to follow clear rules (Covaleskie, 1992) and is highly related to conscientiousness (Jung et al., 2017). In a military context, this means that the commandos

have to follow orders precisely even in light of extreme exhaustion or other adverse

circumstances. Dedication is described by the level of commitment to one's tasks (Mart,

2013). Considering the tasks that special forces operators are facing during their careers (e.g.,

demanding environments, dangerous situations), dedication to that profession is key (Beebe-

Mocilac, 2007). Generally, discipline and dedication may be best derived from the actual

behavior that people display, rather than self-reports on questionnaires. As proposed by

Baumeister et al. (2007), assessing actual behavior, which is rarely done in psychology

anymore, provides more insight and predictive power than the total scores of questionnaires.

A first step towards zooming in on behavioral measures from questionnaire studies is

examining the temporal structure (i.e., response times) and answer patterns of the process of

filling in the questionnaires. Therefore, it may be interesting to not only consider the outcome

of the questionnaire but rather the process of *how* the recruits fill in the questionnaire.

Response Times

Not only do the answers to the facets yield information about the recruits, but also the

time in which they complete a questionnaire. Response times are commonly used to assess

how fast people respond to different stimuli (e.g., implicit association test; Greenwald et al.,

1998). Generally, response times can reveal information on how much attention an individual

is paying but also on how cognitively loaded an item is (Zheng & Cook, 2012). For example,

emotional responses (i.e. low cognitive load) are supposed to be faster than responses where

cognitive reasoning is required (Yan & Tourangeau, 2008). Furthermore, response times are

not only influenced by external loads (i.e., number of questions asked, number and type of

answer categories) but also by age, education, and experience of the individual (Toepoel et

al., 2009).

Generally, discipline and dedication can be addressed by two different expressions,

namely temporal (i.e., response times) and structural (i.e., answer patterns). Hence, discipline

is related to response times with a medium speed and a high degree of stability, whereas

answer patterns are related to high consistency within the facets. We will investigate the response times per item for the personality questionnaire that recruits fill in during their first week of training and determine whether the response pattern can discriminate between successful applicants who ultimately complete the training and unsuccessful ones who do not make the final selection. Specifically, focused work may reflect medium response times indicating that questions are read carefully with some fluctuations between questions. However, to the best of our knowledge, no research has been conducted on assessing response times in personality questionnaires on military recruits. Relatedly, extracting information about the stability in response times to distill the discipline and dedication of the recruits is an interesting new avenue. Therefore, we aim to take a next step by examining how recruits fill in the personality questionnaire, rather than focusing exclusively on the outcomes of the personality questionnaires.

Previous research has already unraveled the importance of temporal structure in human behavior. For example, Van Orden and colleagues (2003) demonstrated that high-level cognitive performance shows specific patterns in the underlying temporal structures. Specifically, the timeseries of the sequential task elements revealed a pattern that is defined by nested signals that demonstrate both long-term stability and short-term flexibility (see also Wijnants et al., 2009, 2012). This specific pattern has even been demonstrated in critical biological processes, such as healthy cardiovascular functioning (Goldberger et al., 2002). Both these different lines of research point to the fact that the aggregated outcome of the underlying temporal patterns measured at a single moment in time may be less important than understanding the temporal structure that ultimately produces these outcomes. In line with the arguments by Baumeister and colleagues (2007), one single overall score on a personality trait is not enough to resemble the behavior of an individual. Therefore, we will evaluate *how* recruits fill in the personality questionnaires. Specifically, these structures may be informative

of the discipline and dedication that are essential to succeed in the special forces training (Akrami et al., 2007).

Answer Patterns

When evaluating personality traits, the outcome (i.e., the score) of the questionnaire is typically assessed. Baumeister and colleagues (2007) point out that psychologists seem to neglect real behavior and that human behavior is extracted from descriptions of (possible) behavior that is not necessarily accurate (e.g., what people say they did, what they want to do and what they would do). Interestingly, real behavior can be obtained from *how* the questionnaires are filled in (i.e., assessing temporal answer patterns). These underlying answer patterns can be informative because they give an indication on how consistent a recruit scores on one trait. This is also in line with the findings of van Orden et al. (2003), who point out that it is not only the outcome, but the exact composition of the pattern that is informative of people's behavior (i.e., the answer to a previous question influences the answer to the following question). Therefore, looking into the individual responses to each construct and its facets will reveal additional information with regards to how consistent a recruit scores on each trait. Recruits with a rather 'stable' composition of personality traits and recruits who seriously fill in the questionnaire are more likely to show consistent answer patterns. As military personnel is supposed to display a special 'military culture', reflected by dedication and discipline, they are expected to show rather consistent patterns in their responses to different items belonging to the same facet (Akrami et al., 2007; Hall, 2011). However, two recruits can obtain similar overall scores on one personality trait (or facet) with different underlying answers (i.e., more or less consistency). Furthermore, previous research found that individuals who are scoring high or low on the constructs of the Big Five personality traits show faster response times compared to those who score average (Hedlund, 2010).

The Current Study

The current study provides a first insight into *how* recruits of the Dutch special forces fill in personality questionnaires. Notably, instead of only using general outcomes of the questionnaires we are zooming in on the underlying answer patterns and response times. Specifically, we will analyze the answer patterns and response times of 249 recruits of the Dutch special forces who filled in the questionnaire during the first week of the intense training. The fifteen weeks of intense training are part of the tough selection procedure. There are two aims in this study. First, we will analyze whether response times are indicators of who is going to pass the training and who does not. We expect that people who are focused while filling in the questionnaire (i.e., displaying stable response times) demonstrate traits known to be important in military personnel (i.e. discipline, dedication). Therefore, we will compare the pace and stability of the response times between graduates and non-graduates. Hence, we hypothesize that specific clusters of response times distinguish successful recruits from non-successful recruits. The second aim is to test whether graduates show different underlying answer patterns to the Big Five personality traits compared to non-graduates. Subsequently, we will not only assess the overall score on each of the five personality traits but the underlying patterns (i.e., different scoring patterns per construct). Generally, successful military personnel is expected to show more discipline and dedication and therefore, we also expect them to score more consistently (i.e., stable) within those constructs than non-graduates (Akrami et al., 2007). Therefore, we hypothesize that graduates respond more consistently between different facets within one trait compared to non-graduates.

Method

Participants

The total sample consisted of 262 males who applied for the selection procedure at the KCT. The applications of the recruits were initially screened before they were admitted to the next selection round during which the current data was collected. Given the sensitivity of the data, no information on age and demographic background of the recruits could be revealed.

Two participants had to be removed because they did not fill in all questions (i.e., missing to fill in all the facets of the questionnaire).

Materials & Procedure

The contents and procedure of this study were reviewed and approved by the ethical committee of Psychology at the University of Groningen. After the first broad selection procedure, the recruits entered the training phase at the Korps Commandotroepen. In the first week of their training they had to fill in multiple extensive questionnaires. The questionnaires were completed digitally (i.e., by using tablets) at the military base. Before the recruits started the questionnaires, they were informed about the procedure and that the results of the questionnaires would not be used for selection purposes. Furthermore, there was no time limit and anonymity was secured by anonymous usernames. The majority of the recruits finished the questionnaire after approximately 30 minutes. Note that the recruits were able to drop out at any time and that only the successful recruits filled in a second questionnaire assessing the personality traits after successful completion of the training. The questionnaire used to assess the Big Five personality traits was the Dutch version of the NEO-PI-3 (Hoekstra & de Fruyt, 2014). The questionnaire consists of 240 questions assessing the following constructs: 1) neuroticism 2) extraversion, 3) openness, 4) agreeableness and 5) conscientiousness. Each of these individual traits is further subdivided into 6 different facets with 8 questions per facet (see Table 1).

Table 1

Overview of the Facets for Each Construct. Note that the initial assessment was done in Dutch. Values in square brackets indicate the 95%

confidence interval for each correlation.

Construct	Facet	Example item	Correlation
Neuroticism	Anxiety	I am easily scared of something	.57 [.51, .68]
	Irritation	Even small annoyances can frustrate me	.48 [.28, .57]
	Depression	Sometimes I feel completely worthless	.70 [.39, .70]
	Shame	In company I worry I'm making a fool of myself	.60 [.30, .60]
	Impulsivity	It takes effort for me to resist my desires	.46 [.22, .47]
	Vulnerability	I often find it difficult to come to a decision	.49 [.46, .60]
Extraversion	Cordiality	I am known as a warm and friendly person	.45 [.26, .56]
	Sociability	I like partying with lots of people	.62 [.32, .62]
	Dominance	I am dominant, powerful and self-confident	.53 [.40, .67]
	Energy	I am a very active person	.51 [.19, .51]
	Adventurism	I like the excitement of the roller coaster	.45 [.20, .46]
	Cheerfulness	I am a cheerful and lively person	.63 [.19, .63]
Openness	Fantasy	I have a very vivid imagination	.50 [.32, .60]
	Aesthetics	Some kinds of music fascinate me immensely	.37 [.31, .65]
	Feelings	My feelings about things are important to me	.40 [.22, .52]
	Change	I find it interesting to start new hobbies	.23 [.23, .40]
	Ideas	I like to solve problems or puzzles	.49 [.41, .57]
	Values	I think people should honor traditional values instead of questioning them	.20 [.20, .39]
Agreeableness	Trust	My first reaction is to trust people	.61 [.37, .63]
	Sincerity	I can't cheat on anyone, even if I wanted to	.32 [.22, .62]
	Caring	In general, I try to be attentive and caring	.48 [.25, .51]
	Complacency	When I am insulted, I just try to forgive and forget	.31 [.22, .45]
	Modesty	I don't like to talk about myself and my achievements	.42 [.37, .53]
	Sympathy	I feel sympathy for people who have suffered a worse fate than me	.44 [.27, .49]
Conscientiousness	Efficiency	I have most things in my life pretty well under control	.50 [.44, .56]
	Orderliness	I keep my things neat and clean	.56 [.21, .56]
	Reliability	I try to perform all tasks assigned to me conscientiously	.41 [.15, .49]
	Ambition	I work hard to achieve my goals	.61 [.25, .62]
	Self-discipline	I am a productive person who always gets a job done	.43 [.35, .67]
	Thoughtfulness	I rarely make hasty decisions	.54 [.35, .67]

The constructs are assessed by various numbers of statements with a five-point Likert scale ranging from strongly disagree to strongly agree. The individual facets are not assessed successively one by one, but the corresponding items of a given facet recur in regular intervals. The resulting order of the questions was the same for each participant. The response times per question were measured in milliseconds.

Data Analysis

The data was analyzed using Matlab 2018a and SPSS Statistics 27. We tested whether response time patterns in the items of the personality questionnaire can distinguish graduates from non-graduates. Data of the non-graduates ($n = 214$) was compared to the recruits who successfully finished the program ($n = 35$). The response times of 249 recruits were assessed in milliseconds on a total of 240 questions. Specifically, focused and steady work on the questionnaire may reflect desirable behaviors for the military (i.e., hypothesis 1). This means that the questionnaire should not be filled in carelessly reflecting very brief, but very steady response times.[1] Also, unfocused work or increased mental exhaustion reflected by either very long response times and highly erratic patterns may be undesirable. In order to test the two facets of (un)desirable behavior, we clustered by overall pace and temporal pattern (i.e. stability).

In order to analyze the temporal patterns of the response times, we assessed the response times per item for each individual. This means that we calculated indicators for temporal stability for each person. Specifically, we determined lag-1 autocorrelations (Von Neumann, 1941). Lag-1 autocorrelations have already been applied to detect critical changes temporal patterns in human behavior (cf. Hill et al., 2020). Lag-1 autocorrelation reflects the stability, or inertia, of a signal. Here, a time series is linearly related to a lagged (i.e. previous) version of itself (Riley & Greenhall, 2004). The lag-1autocorrelation can range between -1

---

[1] Note that this response pattern is highly unlikely in this context as the recruits have a strong desire to "perform well" on the personality test in order to increase their chances of attaining the desired end goal (i.e., being selected for the special forces).

and 1. A value around 0 there indicates that there is no consistent linear relationship between two subsequent measurement points. In our case, this would mean that there is no stability in the response patterns for subsequent items. If the lag-1 autocorrelation is 1 it means that the times are the same or continuously increase across items. Contrary, if the lag-1 autocorrelation is -1 it shows a negative relationship meaning that a small response time is followed by a higher response time and vice versa. Thus, lag-1 autocorrelations provide an insight into the specific pattern with which subsequent questions are answered. While the lag-1 autocorrelation provides a standardized indicator of the temporal stability, it does not yield much information on the fluctuations within a signal. Note that the lag-1 autocorrelation is conducted for each participant's response times separately before the scores are aggregated for further analyses.

In the next step we cluster the data on pace (i.e., response time) and stability (i.e., lag-1 autocorrelation). This means that up to four distinct clusters may emerge, namely [fast/erratic], [slow/erratic], [slow/steady] and [fast/steady]. We did this by starting to plot the data into one cluster in order to compute the distance of the single cluster centroid. Afterwards, we added cluster two, three and four, respectively. In a final step, we plotted the four clusters for visual inspection (i.e., cluster formation). Besides the visualization of whether adding more clusters can distinguish between graduates and non-graduates, we conducted the according chi-square tests for clusters two, three and four (i.e., with graduates and non-graduates by cluster group). Note that no chi-square test for cluster one can be conducted as every recruit is part of this cluster. Thereby, we assess whether the emerging clusters indeed converge to the four distinct patterns (i.e., [fast/erratic], [slow/erratic], [slow/steady], [fast/steady]). Our hypothesis that response times allow us to make predictions about whether a candidate will pass the selection will be supported if the four clusters can distinguish between successful and unsuccessful recruits.

In order to test whether the consistency in the different facets of the personality traits is higher in the graduates compared to the non-graduates (Hypothesis 2), we first computed the mean score for each participant on each of the 30 facets. To assess the response patterns of the 249 recruits, we had to match the answers to the questions based on the construct they belong to. Again, the data of the non-graduates ($n = 214$) was compared to the graduates ($n = 35$). Then, we conducted ANOVAs for the graduates and non-graduates to assess whether the facets within one personality trait significantly differ from the others. To illustrate, we tested whether (at least) one of the six facets of the trait extraversion was significantly different from the others. We expected that this would never be the case for the graduates while the group of the non-graduates would show significant differences between the facets of the traits. However, similar to the problem of multiple testing and sample size difference outlined above, we also considered the effect size measure *eta squared*. Specifically, we expected small effect sizes ($\eta^2 = .01$) for the graduates and small to medium sizes ($\eta^2 = .06$) for the non-graduates. Thus, our hypothesis is supported if we find no significant differences in the facets for the graduates as well as smaller effect sizes compared to the non-graduates.

In the next step, we analyzed the responses to each of the five constructs (i.e., neuroticism, extraversion, openness, agreeableness and conscientiousness) per facet. We hypothesized that successful graduates (i.e., recruits) display more consistency in their answers per facet compared to unsuccessful recruits. To test this hypothesis, we applied a categorical recurrence quantification analysis (RQA) to the response pattern of the 240 items (recoded for reversed items) of each participant. A RQA compares a timeseries with time-delayed copies of itself. Specifically, the categorical RQA returns an indication of how many points in the timeseries belong to the same category (e.g., the same score on a Likert scale) given the specified delay (i.e., recurrence rate, Coco & Dale, 2014; Dale & Spivey, 2005). This means that we can trace whether similar answers were given to items of the same facets within one questionnaire. Moreover, choosing the right delay parameter (τ) allows us to

quantify the recurrence for the entire questionnaire by facet. For example, the NEO-PI-3

assesses five personality traits with 6 facets in regular intervals. Therefore, the items assessing

one specific facet are exactly 30 'steps' away from each other (e.g., 1, 31, 61, …, 211). In

such a configuration, setting τ to 30 ensures that for the entire questionnaire all items within

a facet are compared to each other, but not to items of other facets. Thus, the categorical RQA

would return a single value reflecting the facet-specific self-similarity of the response pattern

for the entire questionnaire. A high recurrence rate corresponds to high consistency in the

response pattern. Our hypothesis that successful recruits display more consistency in their

answers per facet compared to unsuccessful recruits is confirmed if an independent samples t-

test indicates a higher recurrence rate for the successful recruits compared to the unsuccessful

recruits.

Results

Hypothesis 1 stated that specific clusters of response times and stability can

distinguish successful recruits from non-successful recruits. Contrary to our expectations, we

did not find the proposed formation of clusters (i.e., [fast/erratic], [slow/erratic],

[slow/steady], [fast/steady]) (see Figure 1). We did not find a systematic distribution for

cluster two ($\chi^2 = 0.02$, $p = 0.96$, $V = 0.003$), cluster three ($\chi^2 = 3.57$, $p = 0.17$, $V = 0.12$) and

cluster four ($\chi^2 = 3.83$, $p = 0.28$, $V = 0.28$). Furthermore, the Elbow plot indicates that cluster

four does not add much to our analysis. This means that we did not find evidence that

response times are predictive about whether a recruit will pass the program.

Figure 1

Representation of the four emerging clusters determined by speed (in ms) and stability

(AR(1)). The clusters are marked by consecutively adding different colors (cluster 1, red;

cluster 2, blue; cluster 3, green; cluster 4, yellow). The centroids are marked by the black X.
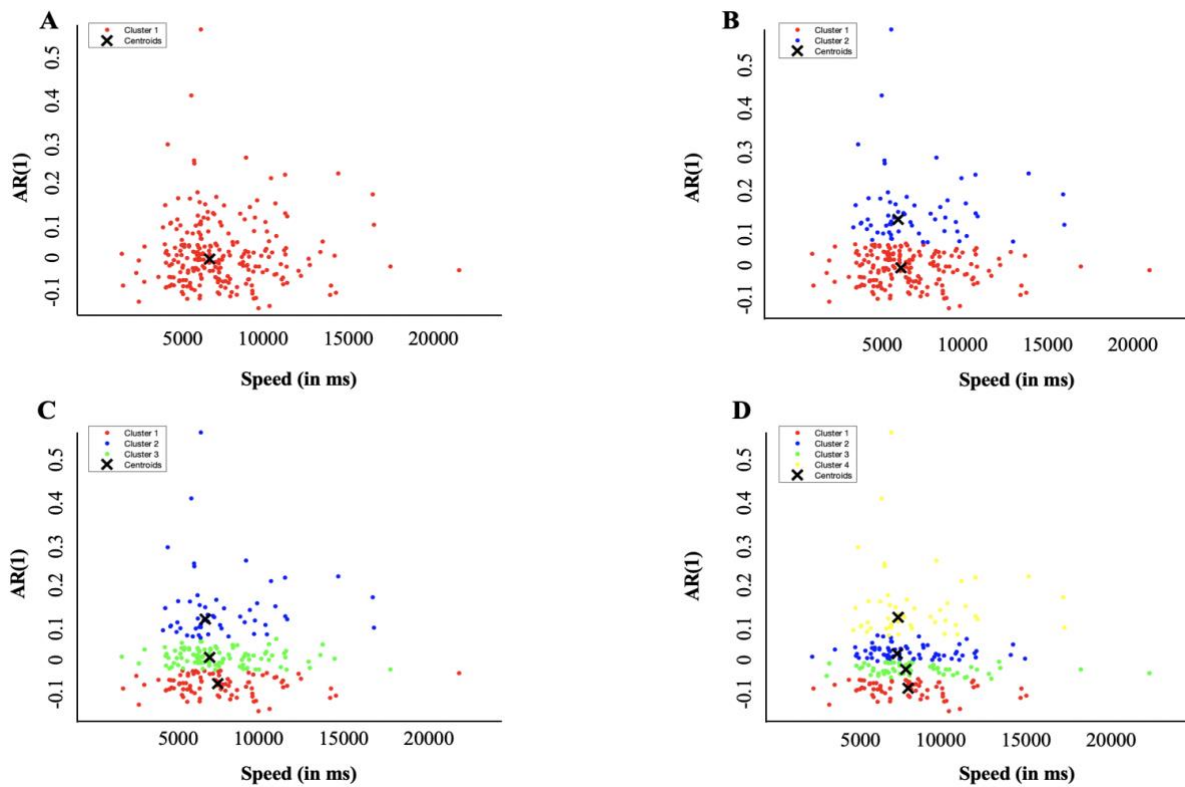
Furthermore, we tested whether graduates respond more consistently to different

facets within one trait compared to non-graduates. However, contrary to our expectation, we

did not find a significant difference in the recurrence rates for the NEO-PI-3 between the

graduates and the non-graduates ($t(246) = .90$, $p = .89$, $d = .16$) (Figure 2). This means that we

did not find a significant difference in the consistency of the answer patterns for specific

facets between graduates and non-graduates.

Figure 2

Recurrence rate for successful graduates (Grad) and non-graduates (Non-Grad).
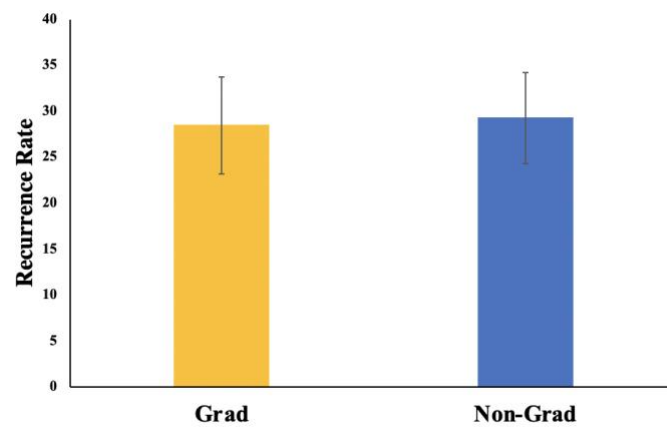
Discussion

The aim of the study was to assess *how* military recruits fill in personality questionnaires in terms of consistency and pace. Previous research has focused on the assessment of the Big Five personality traits and whether there are common patterns for people who are successful in the military. For instance, Braun and colleagues (1994) found that successful military personnel displays lower scores on neuroticism and agreeableness and higher scores on extraversion and conscientiousness. Additionally, Akrami and colleagues (2007) and Hall (2011) reasoned that successful military personnel displays traits like dedication and discipline (i.e., displaying a 'military culture').

Baumeister and colleagues (2007) pointed out that the overall score of a questionnaire does not resemble actual behavior. As current and past behavior are the best predictors of future behavior (Ouellette & Wood, 1998), the first step should be to assess the behavior of the recruits while filling in the questionnaires (i.e., *how* they fill in the questionnaires). Analyzing the actual behavior of the recruits while filling in the questionnaire reveals information about how disciplined and dedicated an individual is. Therefore, we investigated whether the response times as well as the answers to the individual facets of each of the Big

Five personality traits are different for graduates and non-graduates. In order to investigate this, we conducted multiple analyses (i.e., categorical recurrence quantification analysis, lag-1 autocorrelations, clustering analysis).

The first hypothesis states that specific clusters of response times and stability distinguish successful recruits from non-successful recruits. Therefore, we expected that the response times of the graduates are more consistent. Additionally, we also predicted that the response times of the graduates are neither too high or too low but more in a medium range. Hence, for the response times not only the stability of the timing was of interest but also the pace in general. We indeed found that the response times of the graduates are more consistent, however, the results are statistically non-significant (see Figure 2). The absence of the nonsignificant effects may be due to the fact that assessing temporal structures is more suited for assessing more global behavior or performance. In previous research, temporal structures were assessed in cognitive tasks (Van Orden et al., 2003; Wijnants et al., 2009, 2012), rowing performance (Den Hartigh et al., 2015), or even cardiovascular activity (Goldberger et al., 2002). As previously pointed out, these studies assessed structures that contained nested long-term and short-term processes. These structures may, however, be distorted when a system experiences fatigue (cf. Hill et al., 2020). Given that the recruits spent many hours filling in various questionnaires, the temporal structure may have become too similar between graduates and non-graduates to be distinguished.

The second hypothesis states that we expect the graduates to display higher consistency within their answer patterns to the individual facets of the Big Five personality traits. One would expect that graduates are more consistent as they may display traits known as important to the military (i.e., discipline and dedication). Assessing the answer patterns to the facets provides us with more detailed information about the underlying scores (i.e., temporal patterns) instead of just considering the total score of each trait (e.g., two recruits can have the same overall score with completely different underlying responses). As outlined

by Baumeister et al. (2007) the way how people fill in questionnaires may be more predictive of their actual behavior than the overall score. The graduates indeed show slightly more consistent answer patterns to the individual facets of the Big Five personality traits, compared to the non-graduates. However, the p-value was rather high and therefore we have to conclude that we did not find statistical support for our hypothesis. The absence of the nonsignificant effects may be due to how the questionnaire (i.e., NEO-PI-3) is constructed. The individual facets are assessed by six questions presented in regular intervals (30 steps away, Hoekstra & De Fruyt, 2014). Therefore, there is not one single structure that is being assessed, but instead there are thirty overlapping structures within the same sequence. Previous research has typically focused the analyses on a single structure (that may be made up of overlapping elements). For example, the study by Den Hartigh and colleagues (2015) analyzed the sequence of rowing strokes to differentiate advanced rowers from beginners. Although it may be argued that the rowing stroke consists of various movements, all of these occur in a fixed temporal sequence in order to manifest a single behavioral structure. Thus, assessing 30 different structures at once may mask the specific dynamics of the individual structures. Hence, it may be necessary to assess the facets of the Big Five differently or focus on questionnaires that assess a single construct.

Theoretical Implications

In order to not only rely on a single overall score, it is valuable to assess the temporal patterns of response times and the answers to the facets of the Big Five personality questionnaires (Baumeister et al., 2007). Notably, temporal patterns allow for more insight on the underlying pattern than just the overall results of each personality trait (De Ruiter et al., 2015; Den Hartigh et al., 2015; Goldberger et al., 2002; Hill et al., 2020). However, questionnaires are merely sign-based approaches and it may be that the recruits of the Korps Commandotroepen are too homogeneous for such an approach (Wernimont & Campbell, 1968). Although these particular signs may be useful according to previous literature,

zooming in on the temporal structure may suggest a behavioral sample which may be better suited for homogeneous groups (Baumeister et al., 2007; Goldberger et al., 2002). Even though we were not able to find significant differences between the graduates and the non-graduates it does not mean that our approach was not appropriate. One reason why the results are non-significant is that the group of recruits is extremely homogeneous. A main theoretical implication may be that this approach can be suited for clinical populations, where the groups are more heterogeneous (e.g., when people display deficits to concentrate).

Limitations

A key limitation of the current study may be that the individuals who acquire a career within the special forces display certain traits already before starting the program. This implies that recruits who are admitted to the training of the Korps Commandotroepen are rather homogenous. This may especially be the case because it is a necessity to follow basic military training before entering the selection program of the Korps Commandotroepen (i.e., homogeneity). Another key limitation may be that the recruits were not aware that the response times were being measured. On the one hand, this is a strength as we were able to measure their true focus and response times, but on the other hand, they may have behaved differently if they would have known (e.g., no breaks). Additionally, more than one question was presented per page and hence, we cannot be sure that the recruits were reading the first question, choosing the answer, reading the next question, choosing the answer and so on. It may be the case that a recruit was reading all of the questions before answering them and hence, displaying a rather high response time for the first question and lower ones for the following questions. Furthermore, the differences in sample size (i.e., 35 graduates and 213 non-graduates) pose a limitation to our statistical analyses (Cohen, 1988). Especially the small sample size of the graduates *(n = 35)* can be problematic for specific analyses we did (i.e., chi-square test).

Additionally, Thunholm (2009) found that Swedish military personnel were responding in a socially desirable way to questionnaires. Even though the recruits of the KCT were informed beforehand that their scores on the questionnaires have no influence on whether they will be selected or not, it is not unlikely that some of them tried to respond in a favorable way. Responding to the questionnaire in a favorable way (e.g., appearing less neurotic than one is) might have impacted the consistency of the answers on the individual facets. Hence, the scores on a specific facet of the Big Five may seem more consistent than they actually are. This is in line with research by Niessen and colleagues (2017), who found that self-presentation is a problem in high-stakes contexts regardless of whether people are informed that the final decision is based on those scores or not. This is indeed likely because the recruits are highly motivated to make it through the program and if they believe that the instructors take a look at the data they probably want to appear as the perfect candidate. Besides the findings of Thunholm (2009), self-report questionnaires are commonly known as object to faking (van de Mortel, 2008). Furthermore, while the lag-1 autocorrelation provides a standardized indicator of the temporal stability, it does not yield much information on the fluctuations within a signal (Riley & Greenhall, 2004).

Directions for Future Research

Besides looking into the response times for personality traits and their underlying patterns, it may be interesting to look into response times of intelligence questionnaires as well as their underlying constructs (i.e. numerical and spatial ability). This may be interesting as there is a lot of existing literature on how to assess cognitive ability with response times (cf. Kyllonen & Zu, 2016). If a recruit responds fast to most of the questions and has a high total score (i.e., correct answers), it may imply that this recruit is able to master the demanding cognitive tasks a commando is facing during the career (e.g., finding safe houses without maps and GPS, coming up with strategies when encountering unpredictable enemy contact). Additionally, future research should conduct categorical recurrence quantification

analysis with other samples to assess if the military recruits indeed score very consistent. This should be done because we can only compare the groups in the study and hence, there is no reference to the general population or specifically stable or erratic clinical groups. Furthermore, as we found slightly higher response times for graduates compared to non-graduates (note that they were not statistically significant), future research should assess whether those differences are also displayed in different units of the military (i.e., at a lower level than the special forces).

Conclusion

The main purpose of this thesis was to extend previous research and to look into *how* military recruits fill in personality questionnaires in terms of consistency and pace. In order to do this, we examined the response times to each question as well as the answers to the individual facets of the Big Five personality questionnaires. Contrary to our expectations, we did not find significant differences in the response times between graduates and non-graduates. Additionally, there was also no statistical support for differences in the consistency of the responses to the personality questionnaires. However, that the response times of the graduates are slightly higher compared to the response times of the non-graduates seems promising for future research. Therefore, we may conclude that further research is needed to support our findings.

References

Akrami, N., Hedlund, L. E., & Ekehammar, B. (2007). Personality scale response latencies as

    self-schema indicators: The inverted-U effect revisited. *Personality and Individual*

    Differences, 43, 611-618. https://doi.org/10.1016/j.paid.2006.12.005

Baumeister, R. F., Vohs, K. D., & Funder, D. C. (2007). Psychology as the science of self-

    reports and finger movements: Whatever happened to actual behavior?. *Perspectives*

    on Psychological Science, 2, 396-403. https://doi.org/10.1111/j.1745-

    6916.2007.00051.x

Beebe-Mocilac, V. R. (2007). Beyond the call of duty: A study of factors that influence

    soldier attitude and behavior with respect to the concepts of duty, loyalty, and

    commitment (Doctoral dissertation, Capella University).

Boe, O., Nilsen, F. A., Kristiansen, O., Krogdahl, P., and Bang, H. (2017). "Measuring

    important character strengths in Norwegian Special Forces officers," in Proceedinga of

    the International Conference on Education and New Learning Technologies,

    Barcelona.

Braun, D. E., Prusaczyk, W. K., & Pratt, N. C. (1994). Personality Profiles of US Navy Sea-

    Air-Land (SEAL) Personnel. *Naval Health Research Center San Diego*.

Buss, A. H. (1989). Personality as traits. *American Psychologist, 44,* 1378-1388.

    https://doi.org/10.1037/0003-066X.44.11.1378

Campbell, J. S., Castaneda, M., & Pulos, S. (2010). Meta-analysis of personality assessments

    as predictors of military aviation training success. *The International Journal of*

    Aviation Psychology, 20, 92-109. https://doi.org/10.1080/10508410903415872

Coco, M. I., & Dale, R. (2014). Cross-recurrence Quantification Analysis of Categorical and

    Continuous Time Series: an R package. *Frontiers in Psychology, 5*, 510-532.

    https://doi.org/10.3389/fpsyg.2014.00510

Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Costa Jr, P. T., & McCrae, R. R. (1992). The five-factor model of personality and its relevance to personality disorders. *Journal of Personality Disorders, 6*, 343-359. https://doi.org/10.1521/pedi.1992.6.4.343

Covaleskie, J. F. (1992). Discipline and morality: Beyond rules and consequences. *The Educational Forum, 56*, 173-183. https://doi.org/10.1080/00131729209335193

Dale, R., & Spivey, M. J. (2005). Categorical recurrence analysis of child language. *Proceedings of the Annual Meeting of the Cognitive Science Society, 27.*

De Ruiter, N. M., Den Hartigh, R. J., Cox, R. F., Van Geert, P. L., & Kunnen, E. S. (2015). The temporal structure of state self-esteem variability during parent–adolescent interactions: More than random fluctuations. *Self and Identity, 14*, 314-333. https://doi.org/10.1080/15298868.2014.994026

Den Hartigh, R. J., Cox, R. F., Gernigon, C., Van Yperen, N. W., & Van Geert, P. L. (2015). Pink noise in rowing ergometer performance and the role of skill level. *Motor Control, 19*, 355-369. https://doi.org/10.1123/mc.2014-0071

Goldberger, A. L., Peng, C. K., & Lipsitz, L. A. (2002). What is physiologic complexity and how does it change with aging and disease?. *Neurobiology of Aging, 23*, 23-26.

Greenwald, A. G., McGhee, D. E., & Schwartz, J. L. (1998). Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology, 74*, 1464-1480.

Griffith, J. (2009). After 9/11, what kind of reserve soldier? Considerations given to emerging demands, organizational orientation, and individual commitment. *Armed Forces & Society, 35*, 214-240. https://doi.org/10.1177/0095327X07312490

Halfhill, T., Nielsen, T. M., Sundstrom, E., & Weilbaecher, A. (2005). Group personality

      composition and performance in military service teams. *Military Psychology, 17*, 41-

      54. https://doi.org/10.1207/s15327876mp1701_4

Hall, L. K. (2011). The importance of understanding military culture. *Social Work in Health*

      Care, 50, 4-18. https://doi.org/ 10.1080/00981389.2010.513914

Hartmann, E., Sunde, T., Kristensen, W., & Martinussen, M. (2003). Psychological measures

      as predictors of military training performance. *Journal of Personality Assessment, 80*,

      87-98. https://doi.org/10.1207/S15327752JPA8001_17

Hedlund, L. E. (2010). Response time as self-schema indicator: Implications for personality

      assessment (Doctoral dissertation, Acta Universitatis Upsaliensis).

Hill, Y., Den Hartigh, R. J., Cox, R. F., De Jonge, P., & Van Yperen, N. W. (2020).

      Predicting resilience losses in dyadic team performance. *Nonlinear Dynamics,*

      Psychology, and Life Sciences, 24, 327-351.

Hill, P. L., & Roberts, B. W. (2011). The role of adherence in the relationship between

      conscientiousness and perceived health. *Health Psychology, 30*, 797-804.

      https://doi.org/10.1037/a0023860

Hills, P., & Argyle, M. (2001). Emotional stability as a major dimension of happiness.

      Personality and Individual Differences, 31, 1357-1364. https://doi.org/10.1016/S0191-

      8869(00)00229-4

Hoekstra, H., & De Fruyt, F. (2014). NEO-PI-3 en NEO-FFI-3: Persoonlijksheidsvragen-

      lijsten: handleiding. Hogrefe.

Johnson, J. L., & Hill, W. R. (2009). Personality Traits and Military Leadership. *Individual*

      Differences Research, 7, 1-13.

Jung, K. R., Zhou, A. Q., & Lee, R. M. (2017). Self-efficacy, self-discipline and academic

      performance: Testing a context-specific mediation model. *Learning and Individual*

      Differences, 60, 33-39. https://doi.org/10.1016/j.lindif.2017.10.004

Kyllonen, P. C., & Zu, J. (2016). Use of response time for measuring cognitive ability. *Journal of Intelligence, 4*, 14. https://doi.org/10.3390/jintelligence4040014

Mart, C. T. (2013). A passionate teacher: Teacher commitment and dedication to student learning. *International Journal of Academic Research in Progressive Education and Development, 2*, 437-442.

McDonald, D., Norton, J., & Hodgdon, J. (1990). Training success in US navy special forces. *Aviation, Space, and Environmental Medicine.*

Niessen, A. S. M., Meijer, R. R., & Tendeiro, J. N. (2017). Measuring non-cognitive predictors in high-stakes contexts: The effect of self-presentation on self-report instruments used in admission to higher education. *Personality and Individual Differences, 106,* 183-189. https://doi.org/10.1016/j.paid.2016.11.014

Obringer, L. A., & Guzman, F. (2006) How the Navy SEALs Work. *HowStuffWorks.com.* https://science.howstuffworks.com/navy-seal.htm. 29 December 2021.

Ouellette, J. A., & Wood, W. (1998). Habit and intention in everyday life: The multiple processes by which past behavior predicts future behavior. *Psychological Bulletin*, *124*, 54-74. https://doi.org/10.1037/0033-2909.124.1.54

Riley, W. J., & Greenhall, C. A. (2004). Power law noise identification using the lag 1 autocorrelation. *In 2004 18th European Frequency and Time Forum*, 576-580. https://doi.org/10.1049/cp:20040932

Rothmann, S., & Coetzer, E. P. (2003). The big five personality dimensions and job performance. *SA Journal of Industrial Psychology*, *29*, 68-74. https://doi.org/10.4102/sajip.v29i1.88

Skoglund, T. H., Brekke, T. H., Steder, F. B., & Boe, O. (2020). Big Five personality profiles in the Norwegian special operations forces. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.00747

Toepoel, V., Das, M., & Van Soest, A. (2009). Design of web questionnaires: The effects of

　　the number of items per screen. *Field Methods*, *21*, 200-213.

　　https://doi.org/10.1177/1525822X08330261

Thunholm, P. (2009). Social desirability in personality testing of military officers. *Military*

　　Psychology. https://doi.org/10.1207/S15327876MP1304_3

Vaara, J. P., Eränen, L., Ojanen, T., Pihlainen, K., Nykänen, T., Kallinen, K., & Kyröläinen,

　　H. (2020). Can Physiological and Psychological Factors Predict Dropout from Intense

　　10-Day Winter Military Survival Training? *International Journal of Environmental*

　　Research and Public Health, *17*, 9064. https://doi.org/10.3390/ijerph17239064

Van de Mortel, T. F. (2008). Faking it: social desirability response bias in self-report research.

　　The Australian Journal of Advanced Nursing, *25*, 40-48.

Van Orden, G. C., Holden, J. G., & Turvey, M. T. (2003). Self-organization of cognitive

　　performance. *Journal of experimental psychology*, *132*, 331-350.

　　https://doi.org/10.1037/0096-3445.132.3.331

Von Neumann, J. (1941). Distribution of the ratio of the mean square successive difference to

　　the variance. *The Annals of Mathematical Statistics*, *12*, 367-395.

Wernimont, P. F., & Campbell, J. P. (1968). Signs, samples, and criteria. *Journal of Applied*

　　Psychology, *52*, 372-376.

Wijnants, M. L., Bosman, A. M., Hasselman, F., Cox, R. F., & Van Orden, G. C. (2009). 1/f

　　scaling in movement time changes with practice in precision. *Nonlinear Dynamics,*

　　Psychology, and Life Sciences, *13*, 75-94.

Wijnants, M. L., Hasselman, F., Cox, R. F. A., Bosman, A. M. T., & Van Orden, G. (2012).

　　An interaction-dominant perspective on reading fluency and dyslexia. *Annals of*

　　Dyslexia, *62*, 100-119. https://doi.org/10.1007/s11881-012-0067-3

Yan, T., & Tourangeau, R. (2008). Fast times and easy questions: The effects of age,

　　experience and question complexity on web survey response times. *Applied Cognitive*

Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition, *22*, 51-68. https://doi.org/10.1002/acp.1331

Zheng, R., & Cook, A. (2012). Solving complex problems: A convergent approach to cognitive load measurement. *British Journal of Educational Technology*, *43*, 233-246. https://doi.org/10.1111/j.1467-8535.2010.01169.x