

Finding the right (test) type: On the differences between type- vs. dimension-based personality tests and between statistics- vs. theory-based personality tests when deciding for or against a test in personnel selection

Janina Diekmann

Cornelius J. König

Thora Schadow

Yasmin Knab

Octavia V. Harrison

Saarland University

Current status: In press in Zeitschrift für Sozialmanagement/Journal of Social Management
---

#### Author Note

Janina Diekmann, Cornelius König, Thora Schadow, Yasmin Knab and Octavia Harrison, Department of Psychology, Saarland University, Saarbrücken, Germany. Janina Diekmann now works for Cemex Deutschland AG, Rüdersdorf bei Berlin, Germany. Yasmin Knab now works for Cosmos Lebensversicherungs-AG, Saarbrücken, Germany. Thora Schadow and Octavia Harrison are now at Goethe University Frankfurt, Germany.

Correspondence concerning this article should be addressed to Janina Diekmann, Department of Psychology, Saarland University, Campus A1 3, 66123 Saarbrücken, Germany. E-Mail: [j.diekmann@mx.uni-saarland.de](mailto:j.diekmann@mx.uni-saarland.de)

### Abstract

The personality test market offers a wider range of different tests that human resource (HR) practitioners may use for personnel selection. The decision for or against a specific test is likely affected by different criteria. The current three studies examine two such criteria that have previously been mentioned in the literature – whether a test is type-based or dimension-based, and whether it has a theoretical or a statistical development background. Using different versions of a fictitious personality test, we examined the attractiveness of these conditions with different subsamples of HR practitioners and business management and psychology students. We did not find differences between conditions in any of the studies. Implications of these null findings for selection scientists and practitioners are discussed.

### Zusammenfassung

Auf dem Markt für Persönlichkeitstests gibt es eine große Anzahl verschiedener Tests, die Personalmanagerinnen und -manager für die Personalauswahl verwenden können. Für welchen Test sie sich entscheiden, hängt vermutlich von verschiedenen Faktoren ab. Die vorliegenden drei Studien untersuchen zwei dieser Faktoren, die in der Literatur bereits erwähnt worden sind: 1. Ob der Test Personen in Typen klassifiziert oder auf Dimensionen positioniert und 2. ob bei der Testentwicklung auf eine Theorie zurückgegriffen wurde oder er auf statistischen Analysen beruht. Diese Faktoren wurden mithilfe verschiedener Versionen eines fiktiven Persönlichkeitstests und verschiedenen Stichproben von Personalmanagerinnen und -managern, Studierenden der Betriebswirtschaftslehre und Studierenden der Psychologie untersucht. Wir konnten keine Unterschiede zwischen den Versionen finden. Die Implikationen dieser Null-Ergebnisse für Personalauswahlforschung und -praxis werden diskutiert.

**Finding the right (test) type: On the differences between type- vs. dimension-based personality tests and between statistics- vs. theory-based personality tests when deciding for or against a test in personnel selection**

Practitioners use a wide range of different personality tests (Berchtold, 2005; Di Milia, 2004). Indeed, their choices are not always easy to understand, as they also use tests that do not seem to fit the purpose of selection well (Diekmann & König, 2015). For example, when the first author of this article applied for a job with a large German company in the automotive industry, a personality test was applied for preselection. Although the company was clearly using a Big Five personality inventory, its test only covered four of the Big Five – and the dimension missing was the one known to have the best predictive validity, conscientiousness (Barrick, Mount, & Judge, 2001). It seems that validity was not the only selection criterion.

Diekmann and König (2015) discussed some possible criteria that might affect practitioners' decision for or against a particular personality test beyond the criterion of validity. In particular, they discussed the attractiveness of type-based personality tests (a test that groups people into different classes) compared to dimension-based personality tests (a test that places people on continua), and considered a theoretical development background (a personality test based on a personality theory) compared to a statistical background (a personality test based on factor analytic procedures). They found that HR practitioners described type-based tests and a statistical background as more attractive. The purpose of the experiments reported here was to extend these results and to examine these two decision criteria experimentally.

### **Background**

Several surveys have provided an overview of practitioners' use of personnel selection methods. They show that some kind of interview is almost always used during the selection process, in management as well as non-management positions and with graduates as well as

experienced employees (e.g., Carless, 2007; Chartered Institute of Personnel and Development, 2009; Diekmann & König, 2015; Di Milia, 2004; Ryan, McFarland, Baron, & Page, 1999; Schuler, Hell, Trapmann, Schaar, & Boramir, 2007; Taylor, Keelty, & McDonnell, 2002). The use of all other selection methods varies considerably in different countries: Whereas personal references are quite often used in Sweden or Malaysia, they play only a minor role in the Netherlands or Spain (Ryan et al., 1999). While cognitive ability or personality tests are quite common in the United Kingdom or Australia, they are less popular in Germany and the USA (Carless, 2007; Chartered Institute of Personnel and Development, 2009; Diekmann & König, 2015; Ryan et al., 1999; Schuler et al., 2007).

The situation becomes even more complicated if one focuses on the use of personality tests in personnel selection. There are many different tests used in practice in the process of personnel selection (Berchtold, 2005; Diekmann & König, 2015; Di Milia, 2004), for example the 16 Personality Factor Questionnaire (16PF; Conn & Rieke, 1994), the Herrmann Brain Dominance Instrument (HBDI; Herrmann International, 2015), the Hogan Personality Inventory (HPI; Hogan, 1986), the Myers-Briggs Type Indicator (MBTI; Briggs Myers, McCaulley, Quenk, & Hammer, 1998), the DISC (Marston, 1979) and the Big Five Personality Inventory (NEO; Costa & McCrae, 1992). It is likely that only few of these tests were designed to be applied in selection processes. Indeed, some of the test publishers explicitly recommend not using them for selection purposes and instead stress the benefit of their use for personnel development (Herrmann International, 2015; The Insights Group Limited, 2015).

Nevertheless, research has focused almost exclusively on the Big Five model, and has even tried to analyze the relationship of other personality tests and models to the Big Five scheme (e.g., Furnham, Moutafi, & Crump, 2003; John, Naumann, & Soto, 2008). There is extensive research concentrating on the predictive validity of these five personality factors concerning different performance criteria such as academic success (Poropat, 2009; Trapmann, Hell, Hirn, & Schuler, 2007), job satisfaction (Judge, Heller, & Mount, 2002) or

job performance (Barrick & Mount, 1991; Hertz & Donovan, 2000; Salgado, 1997). A stable finding of these meta-analyses is the predictive validity of conscientiousness. The predictive quality of the other four factors differs depending on criteria or occupations, but seems to be generally lower than that of conscientiousness.

Due to the often lamented scientist-practitioner gap (e.g., Shapiro, Kirkman, & Courtney, 2007), practitioners' knowledge of these research findings remains questionable. In particular, of the American practitioners surveyed by Rynes, Colbert and Brown (2002), only half correctly disagreed with the statement that there are only four personality dimensions, as measured by the MBTI, and even fewer correctly disagreed with the statement that there is little difference in the predictive validity of different personality tests. These results have been replicated for Australia (Carless, Rasiah, & Irmer, 2009), the Netherlands (Sanders, van Riemsdijk, & Groen, 2008) as well as Finland, South Korea and Spain (Tenhiälä et al., 2016). Thus, there seems to be a persistent gap in practitioners' knowledge of research results in personnel selection, implying that reasons other than scientific findings influence the decision to use a particular selection method and a specific personality test.

To better understand practitioners and their use of personality tests for selection purposes, Diekmann and König (2015) took a closer look at the differences between personality tests in use. They reviewed the literature on the use of personality tests for personnel selection, focusing on practitioners' potential reasons for using one or the other test, and also complemented their arguments with a survey on practitioners' preferences. This led them to suggest several hypotheses, two of which we discuss and experimentally test in this article.

The first hypothesis refers to the question of whether dimension- or type-based personality tests are more appealing to practitioners, because the tests actually used by practitioners represent both types of test; for example, the MBTI and the DISC are type-based tests, while the 16PF and the NEO use dimensions. The survey conducted by Diekmann and

König (2015) provided first hints regarding the relevance of the type- vs. dimension-based differentiation: HR practitioners in their sample stated that they preferred the aggregation of traits to a type over the dimensional representation of traits. There are several reasons for this preference. A type-based personality test, in which the dimensional information of a scale is reduced to distinct types (e.g., whether a person is introverted or extroverted), may be more appealing because it is easier to interpret and to compare with a defined requirements specification (e.g., whether the applicant is supposed to be introverted or extroverted) or with other applicants. Moreover, the requirements specification itself may be easier: One only needs to decide whether the applicant is supposed to be one type or the other, while it is not necessary to determine the degree of (for example) introversion and extroversion. This corresponds to the limitation of cognitive capacity (Tversky & Kahneman, 1974) and the human tendency to think in schemata (Smith & Queller, 2008). Therefore, the appealing effect of types may positively influence practitioners' decision in favor of a respective personality test, although psychological researchers mostly adopt a more critical stance (e.g., concerning simplification, cutoff points and uniqueness of assignment to a type; Gangestad & Snyder, 1985; Robins, John, & Caspi, 1998; York & John, 1992) when it comes to the type-based personality tests. However, most HR practitioners are not psychologists (e.g., only 5% of the Diekmann & König, 2015, sample of German HR practitioners) and are thus unlikely to be aware of these critical issues, and they have often not been trained to apply and interpret personality tests. Accordingly, we hypothesize (and test in Experiments 1a and 1b) that:

*H1: A personality test based on types will be evaluated more positively than the same test based on dimensions.*

Our second hypothesis concerns the development of a given personality test, more precisely the question of whether a more theoretical or a more statistical background affects the decision for or against a personality test. The development of the various tests was influenced by different personality theories, for example the typology of C. G. Jung (Jung,

1960) or the behavioral types of William Marston (Marston, 1979). Jung's theory, consisting of two attitudes (extraversion and introversion) as well as two functions (sensing vs. intuitive perception and thinking vs. feeling judgment) was the starting point for the development of the MBTI (Briggs Myers et al., 1998; Briggs Myers & Myers, 1993), one of the most frequently used tests in business contexts (e.g., Brown, 1999; Furnham, 2008; Muñoz & Fernández-Hermida, 2010; Ryan & Sackett, 1987, 1992). Marston's types Dominance (D), Inducement (I), Submission (S), and Compliance (C) inspired the development of personality tests like the DISC Analysis. Other tests have a more statistical background, for example the NEO (e.g., Costa & McCrae, 1992; McCrae & Costa, 1997), which is based on the factor-analytically derived Five Factor Model of personality. Cattell's 16PF is also based on factor analytic methods (Cattell, 1978; Conn & Rieke, 1994). Often, personality tests use both approaches: Personality tests that were inspired by theory use statistical methods to evaluate, for example, factor structure and measurement properties, and personality tests based on factor analysis use theories, for example, to construct items. However, whereas some tests, like the NEO, stress the statistical background (Costa & McCrae, 1992), others, like the MBTI, focus on the personality theory (Briggs Myers et al., 1998).

Both approaches may have their attractive aspects for practitioners (Diekmann & König, 2015). On the one hand, the statistical methodology may be appealing as the empirical approach gives the test and its results the appearance of scientific and thus also legal legitimacy, which is also an important decision criterion (Klehe, 2004; König, Klehe, Berchtold, & Kleinmann, 2010). On the other hand, a test based on factor analysis of extensive universal personality traits does not focus on the work context and may therefore be too general for practitioners. A theory-based personality test may appear well-conceived: The mentioned theories are not new but have been known for years, they may be easily recognized and perceived as well proven, and they may serve the human need for explanation (Keil, 2006; Lombrozo, 2006; Malle, 2004) better than statistics. In the survey by Diekmann and König

(2015), practitioners preferred the statistics-based development. We wish to test this expressed preference experimentally (in Experiment 2), and therefore hypothesize that:

*H2: A statistics-based personality test will be evaluated more positively than the same test based on a personality theory.*

### **Experiment 1a (Type)**

#### **Method**

**Participants.** Participants were HR employees in German companies, found from different listings of companies with more than 500 employees, as we expected mainly large companies to have HR staff involved in personnel selection. We contacted HR departments by telephone and asked for the appropriate contact person concerning personnel selection. If this person agreed to participate, they were invited by e-mail to an online survey operated via Unipark (QuestBack GmbH). As compensation, they were offered a summary of the study results.

In total, 115 HR experts (48.7% female, 47.8% male, 3.5% did not indicate their gender) participated in the first experiment. They had an average of about 13 years of professional experience in HR ( $M = 12.9$ ,  $SD = 9.2$ ). Almost all (92.2%) were personally involved in the selection process; 78.8% had the authority to decide which selection methods were to be used. Approximately two thirds (67.8%) had completed an academic education, mostly in business management ( $n = 37$ ), personnel management ( $n = 8$ ), and psychology ( $n = 8$ ). About one third (33.9%) had completed vocational training, in the majority of cases as industrial clerk ( $n = 9$ ), merchant ( $n = 7$ ), personnel officer ( $n = 5$ ), and office clerk ( $n = 4$ ). In 26.1 % of the companies, personality tests were used in selection procedures, which lies above the 15 % reported by Diekmann and König (2015).

Once they had begun the survey, participants were welcomed and given some information about informed consent. They were asked to think of a situation in which they wanted to supplement their selection procedure with a personality test and were randomly



assigned to the type ( $n = 53$ ) or the dimension ( $n = 62$ ) condition, which differed with regard to the subsequent material presented.

**Material.** The material consisted of a one-page personality test flyer. The design of this flyer was based on websites and flyers of popular personality tests, but the flyer was completely fictitious as we wished to avoid copyright issues and distortion based on connections to real test publishers that might be known by HR experts. Analogous to Diekmann, König and Alles (2015), we called the personality test the Personality at Work Inventory (P-WIN), and the structure of the flyer was also similar. The flyer was divided into three parts. On the left-hand side, it explained the benefits of using a personality test in personnel selection in general, on the right-hand side it clarified the advantages of this particular personality test, and in the middle it told the reader what P-WIN was supposed to measure. This part differed in the type and the dimension condition. The dimension condition emphasized the existence of eight basic personality dimensions. The word “dimension” was used twice, both times printed in bold. A graph showed the eight dimensions (see Figure 1). Similarly, in the type condition, the existence of eight basic personality types was pointed out and the word “type” was used twice, printed in bold. The graph showed the eight types, which were similar to the dimensions with respect to content (see Figure 2). The content of the eight traits was adapted from various personality tests, but the exact wording was completely invented for this study.

**Dependent variables.** After reading the flyer for P-WIN, participants were asked to evaluate this personality test. For this, we used seven items with a 7-point rating scale (1 = *strongly disagree* to 7 = *strongly agree*) as well as one open-format item in which participants were able to comment on which aspects concerning content or design of the flyer had affected their ratings. The seven rating items concerned the suitability of the instrument for personnel selection (item 1), contribution to the objectivity of the selection decision (item 2), attractiveness of the content (item 3), extent of gain of information compared with other

selection methods (item 4), interest in more information about P-WIN (item 5), whether the participant would use this test in his or her company (item 6) and the overall liking of the personality test (item 7). As these seven items showed a very good reliability (Cronbach's  $\alpha = .91$ ), they were combined to form a scale of overall test rating.

### **Results and Discussion**

We used a t-test for independent samples to test our hypothesis that a type-based test is more attractive to HR experts than a dimension-based personality test. Contrary to expectation, we found no significant effect,  $t(114) = -0.18$ , between the type ( $M = 3.72$ ,  $SD = 1.29$ ) and dimension ( $M = 3.72$ ,  $SD = 1.18$ ) condition. Hypothesis 1 was therefore not confirmed.

A possible explanation may lie in the potency of our manipulation. As the two versions of our flyer were fairly similar, our participants may not have been aware of the fact that they were evaluating a type-based personality test or a dimension-based personality test. Overall, 58.3% of the respondents used the possibility to write a comment using the open-format item, but only 3.5% mentioned the word "type" and only 2.6% the word "dimension" (both values increase marginally to 4.3% when including the word "category" for type and the words "component" and "personality factor" for dimension). On the other hand, in the study by Diekmann et al. (2015), a comparable manipulation did work. However, to enhance the strength of our manipulation, we designed a second experiment in which we did not use a flyer, but a results report with two different versions. Our hypothesis that type-based tests are generally more attractive than dimension-based tests remained the same and was tested with business management students as participants.

### **Experiment 1b (Type)**

#### **Method**

**Participants.** The recruitment of participants was twofold: One part (27.5%) was contacted during two lectures at a local technical college in Germany. They were given a

paper-and-pencil version of the material and answered the questions during the lectures. No compensation was offered. Another part (72.5%) was contacted by email distribution lists comprising business management students of different German universities. They were offered the chance to win one of five 10€ Amazon gift cards. Of the initial sample, only those studying business management or similar were included in the analysis.

A total of 160 students (66.3% male, 33.8% female) participated in this experiment. Of these, 78.1% were studying at a technical college and 20.6% at a university. Most of them had a study focus on personnel management (16.3%), controlling (15.0%) or logistics (13.1%) and were heading for a Bachelor degree (79.4%). One third (36.9%) had experience in personnel selection (for example by having been an applicant in a selection process) and of these, 3.8% stated having experience with personality tests.

Participants using the paper-and-pencil version were verbally welcomed and told about informed consent conditions. On the first page of the material, they were asked to imagine being a human resources employee who wants to supplement their selection procedure with a personality test. On the next page, they were randomly assigned to the type or the dimension condition. Participants of the online version were welcomed and told about informed consent conditions in written form. The subsequent information matched the paper-and-pencil version. Altogether,  $n = 78$  participants were assigned to the type condition and  $n = 82$  participants to the dimension condition.

**Material.** For this study, we designed a fictitious results report of the P-WIN personality test in a type-based version and a dimension-based version. The report consisted of two pages. On the first page, there was a logo of P-WIN, the information that this was the “results profile of John Doe, sample company”, as well as some basic information about P-WIN and the “fact” that there are eight basic dimensions or types of personality, respectively. On the second page, there was a description of the eight dimensions or types (which were almost the same as in Experiment 1a – we merely replaced structuredness with composure and

spirit of discovery with performance orientation) as well as a pictorial representation of the resulting personality profile. This differed markedly in the two versions: In the dimension-based report version, there was a point on a scale representing the manifestation on each of the eight dimensions (see Figure 3); in the type-based version, there were only two of the eight types marked as primary and secondary preference (see Figure 4). Therefore, there was a stark difference between the two results reports. In the dimension-based version, John Doe was self-evaluated on all eight dimensions; in the type-based version, there were preferences for only two types.

**Dependent variables.** The evaluation of the results reports began immediately after reading the two pages of report. We shortened the number of items to five, concerning the suitability of the instrument for personnel selection (item 1), contribution to the objectivity of the selection decision (item 2), attractiveness of the results report (item 3), interest in more information about P-WIN (item 4), and the overall liking of the personality test (item 5). Again, we used a 7-point rating scale (1 = *strongly disagree* to 7 = *strongly agree*) and added the same open-format question as in the first experiment. The five items showed an acceptable reliability (Cronbach's  $\alpha = .79$ ) and were combined to form a scale of overall test rating.

### Results and Discussion

Again, we used a t-test for independent samples to test our hypothesis that the type-based report is more attractive than the dimension-based personality report. However, we found no significant effect,  $t(158) = 0.92$ , between the type ( $M = 4.57$ ,  $SD = 1.07$ ) and dimension ( $M = 4.42$ ,  $SD = 0.99$ ) condition. Only 15.6% of the participants used the chance to explain their ratings in the open-format question, and just 3.1% mentioned the word "type" and 0.6% the word "dimension". When including the words "category" and "aspect" for type, and "trait", "component", "aspect" or "points" for dimension, 10.0% commented on the type aspect and 3.8% on the dimension aspect. Despite the more obvious difference between the two conditions, our hypothesis was again not confirmed. [Furthermore, taking the different

modalities (paper-and-pencil vs. online) into account did not change these results, with the corresponding ANCOVA showing an  $F(157) = 1.17$ .]

Therefore, the results of Experiment 1a were replicated: In both experiments, participants did not show a preference for a type-based test over a dimension-based test.

### **Experiment 2 (Theory)**

In our second experiment we wanted to test our hypothesis that a statistics-based personality test will be evaluated more positively than its theory-driven counterpart. We decided to use the flyer manipulation again, as it did work in the study by Diekmann et al. (2015) and we could not transfer the results report design to this hypothesis. Furthermore, we decided to work with three groups of participants. As in Experiment 1, we asked HR experts and business management students, as most of the HR experts had studied business management and we wanted to see whether the difference in experience accounted for differences in evaluations. As a third group, we included psychology student because they should know personality tests from their studies, and we wished to analyze possible differences between the different training backgrounds.

### **Method**

#### **Participants**

**HR experts.** The recruitment procedure was the same as in Experiment 1a. A total of 89 German HR experts (48.3% female, 41.6% male, 10.1% did not indicate their gender) participated in the second experiment. They had been working as HR employees for an average of 11.4 years ( $SD = 9.4$ ). Overall, 82.0% were personally involved in the selection process, and 61.8% had decision-making power concerning the choice of selection methods. The majority (71.9%) had completed an academic education, mostly having studied business management ( $n = 27$ ) or psychology ( $n = 18$ ). About one third (34.8%) had completed vocational training, mostly as industrial clerk ( $n = 14$ ) or as banker ( $n = 6$ ). In 27.0 % of the

companies, personality tests were used in selection procedures. Participants were randomly assigned to the theory ( $n = 44$ ) or the statistics ( $n = 45$ ) condition.

***Business management students.*** Business management students were recruited during a business management lecture at a local university in Germany. They were given a paper-and-pencil version of the material and answered the questions during the lecture (without any compensation). Of the initial sample, only those studying business management or similar were included.

Overall, 97 students (37.1% male, 62.9% female) participated. Most of them focused on personnel management (18.6%) and marketing (12.4%) and were aiming for a Master degree (71.5%). Almost half (42.3%) stated that they had experience in personnel selection; 12.5% of these had experience with personality tests. Participants were verbally welcomed and told about informed consent conditions. On the first page of the material, they were asked to imagine being an employee in human resources who wants to supplement the selection process with a personality test. On the next page, they were randomly assigned to the theory ( $n = 44$ ) or the statistics ( $n = 53$ ) condition.

***Psychology students.*** Psychology students were recruited via e-mail distribution lists with psychology students of different German universities. They were offered the chance to win one of five 10€ Amazon gift cards. Of the initial sample, only those studying psychology were included in the analysis.

A total of 93 psychology students (80.6% female, 18.3% male, 1.1% did not indicate their gender) participated. They were mostly heading for a Bachelor (67.7%) or Master degree (20.4%) with a focus on clinical psychology (33.3%) or work and organizational psychology (21.5%). Only 23.7% had experience in personnel selection, and of these, 22.7% had experience with personality tests. The procedure was the same as in Experiment 1. Participants were again randomly assigned to the theory ( $n = 48$ ) or the statistics ( $n = 45$ ) condition.

**Overall sample.** Altogether, 279 individuals participated in Experiment 2, with 48.7% in the theory and 51.3% in the statistics condition.

**Material.** Building on Diekmann et al. (2015), we used the flyer material from Experiment 1. The structure was the same, with the benefits of generally using a personality test explained on the left-hand side and the advantages of using this personality test in particular outlined on the right-hand side. The middle part informed the reader about the construction background of P-WIN, which differed in the theory and the statistics condition. Both cases involved a fictitious American professor named Smith. In the statistics condition, the reader was informed that this professor discovered the existence of eight personality traits using statistical methods. The use of statistics was emphasized three times. In the theory condition, it was stated that the same professor developed a personality theory instead. Again, the word “theory” was used three times. The personality traits were the same in both conditions, and printed below this information.

**Dependent variables.** We used the same six items as in Diekmann et al. (2015), with the 7-point rating scale (1 = *strongly disagree* to 7 = *strongly agree*) plus the open-format item. The items referred to the suitability of the instrument for personnel selection (item 1), contribution to the objectivity of the selection contexts (item 2), attractiveness of content (item 3) and design (item 4) of the flyer, interest in additional information about P-WIN (item 5) and overall test liking (item 6). As the six items showed a good reliability (Cronbach’s  $\alpha = .87$ ; also in the three subsamples: HR experts  $\alpha = .86$ , business management students  $\alpha = .83$ ; psychology students  $\alpha = .88$ ), they were combined to form a scale of overall test rating.

## Results and Discussion

To analyze whether people would prefer a test developed by statistical methods or a test designed on the basis of a psychological theory, we used a 2 (statistics vs. theory)  $\times$  3 (HR experts vs. business management students vs. psychology students) analysis of variance.

Levene's test for equality of variances indicated that the variability of each group was approximately equal,  $F(5,273) = 1.10$ . Means and standard deviations of all groups can be found in Table 1. Contrary to our expectations, an ANOVA showed no significant main effect of test version,  $F(1,273) = 0.58$ . Thus, there was no support for our hypothesis that the perception of tests with a statistical background differed from that of tests with a theoretical background. There was, however, a significant main effect of subsample,  $F(2,273) = 14.85$ ,  $p < .01$ , indicating a difference of ratings between subsamples. Descriptive results indicated that psychology students ( $M = 4.46$ ,  $SD = 1.18$ ) generally rated our test better than business management students ( $M = 3.86$ ,  $SD = 0.98$ ) or HR practitioners ( $M = 3.60$ ,  $SD = 1.10$ ). A post hoc Scheffé test revealed significant differences between psychology students and business management students,  $p < .01$ , and between psychology students and HR practitioners,  $p < .01$ . No interaction between subsample and test version could be found,  $F(2,273) = 0.17$ .

### **Overall Discussion**

In the present studies, we analyzed two possible decision criteria of practitioners when choosing a personality test for specific use in employee selection. Contrary to our first hypothesis, there was no significant difference between a type-based and a dimension-based personality test in our experimental design, although Diekmann and König (2015) found a preference for types in their survey. There may be different explanations for this: First, although we performed our experimental design with a second test material stressing the difference between types and dimensions (the results report in contrast to the flyer), the manipulation may still not have been sufficiently strong. Second, HR managers in the survey by Diekmann and König (2015) likely had particular tests in mind when thinking about dimension- or type-based tests. Possibly, some of them worked with (or at least knew) some particular tests in personnel selection or personnel development which were type-based, therefore triggering a preference for types. Our experiment was not based on a specific personality test (either type-based or dimension-based) and therefore eliminated this bias.



Likewise, we did not find a significant difference between the evaluations of theory-based and statistics-based personality tests, although Diekmann and König (2015) found a (self-stated) preference for the statistics-based development of personality tests among practitioners. Again, the survey by Diekmann and König (2015) may have entailed a bias based on knowledge: The HR experts surveyed may have heard that a statistics-based development is more scientific and answered accordingly due to social desirability in this study conducted by university researchers.

Although there were no significant differences between conditions, we should not underestimate the implications of our results – finding no effects also matters for this field. Psychological science pays great attention to dimension-based personality tests. Research about types is limited, and the most widespread type-based model focuses on three types of children (resilient, under- and overcontroller; e.g. Van den Akker, Deković, Asscher, Shiner, & Prinzie, 2013). Researchers should ask themselves whether they are meeting practitioners' needs when they choose to ignore type-based personality models in organizational contexts. If a considerable number of practitioners like (or at least do not dislike) a type-based approach to personality, ignoring types might create another research-practice gap. Instead, researchers should try to find out what causes this interest and what practitioners really want when using a personality test. Researchers could either adjust their focus and try to develop more scientifically sound type models, or increase their efforts to give good advice concerning the benefit of dimension-based personality tests.

Similarly, researchers should reconsider a purely statistical approach to the construction of personality tests. More effort should be undertaken to develop a theoretical grounding of our personality models and/or it should be argued more explicitly what the importance of empirically founded personality traits is. Furthermore, Diekmann et al. (2015) recommend more tailored solutions to the research-practice gap and its specific aspects – in this case the use of personality tests in personnel selection. Our results suggest the same:

Researchers should be more open to the needs of practitioners, and conscientiously develop new and agreeable solutions to establish an emotionally stable relationship with our practitioner partners.

Thus, these studies show that there are also merits in null findings. In our case, there were no significant differences between conditions, but this also means a lack of effect in a direction science may desire (e.g., a preference for dimensions over types). Although psychology seems to have a bias towards publishing significant results (e.g., Fanelli, 2010a, 2010b, 2011), this study shows the advantages of the recent trend to overcome this bias (e.g., Landis, James, Lance, Pierce, & Rogelberg, 2014; van Dick, 2015; see also Kundoor & Mueen, 2010).

As in every study, there are limitations that need to be addressed. The experiments in this paper were simulations using a fictitious personality test in a laboratory framework. The fact that we used a fictitious test is based on practical reasons: We could not compare a real dimension-based personality test to a real type-based personality test because to our knowledge, there is no such test existing in both versions. However, since Diekmann et al. (2015) already used this design and found effects, we could be fairly certain that the design of the flyer would work. The laboratory design was used to enable us to experimentally test our hypotheses without confounding variables affecting our results. In order to draw causal conclusions, we did not simulate a real-world decision, but rather simplified the design and reduced the variables. Although this might constitute a limitation to the external validity of our results, it is crucial to keep in mind that laboratory and field effects correlate rather highly (Mitchell, 2012).

Our experiments are hopefully just the beginning of a new line of research (see also Diekmann & König, 2015): To better understand which criteria have an impact on practitioners' decisions for or against a personality test, further research is clearly needed. This

should be accomplished with an application-oriented objective – to deliver better solutions to HR practitioners that meet scientific standards as well as practical requirements.

### References

- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1–26. <http://doi.org/10.1111/j.1744-6570.1991.tb00688.x>
- Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment, 9*, 9–30. <http://doi.org/10.1111/1468-2389.00160>
- Berchtold, M. (2005). Häufigste Auswahlverfahren in der Personalselektion [Most frequently used selection methods in personnel selection]. *HR Today, 12/2005*, 37–39.
- Briggs Myers, I., McCaulley, M. H., Quenk, N. L., & Hammer, A. L. (1998). *MBTI manual: A guide to the development and use of the Myers-Briggs Type Indicator* (3rd ed). Palo Alto, CA: Consulting Psychologists Press.
- Briggs Myers, I., & Myers, P. B. (1993). *Gifts differing: Understanding personality type*. Palo Alto, CA: Consulting Psychologists Press.
- Brown, R. (1999). The use of personality tests: A survey of usage and practice in the UK. *Selection and Development Review, 15*, 3–8.
- Carless, S. A. (2007). Graduate recruitment and selection in Australia. *International Journal of Selection and Assessment, 15*, 153–166. <http://doi.org/10.1111/j.1468-2389.2007.00377.x>
- Carless, S. A., Rasiah, J., & Irmer, B. E. (2009). Discrepancy between human resource research and practice: Comparison of industrial/organisational psychologists and human resource practitioners' beliefs. *Australian Psychologist, 44*, 105–111. <http://doi.org/10.1080/00050060802630015>
- Cattell, R. B. (1978). *The scientific use of factor analysis in behavioral and life sciences*. New York, NY: Plenum Press.

- Chartered Institute of Personnel and Development. (2009). *Recruitment, retention and turnover. Annual survey report 2009* (pp. 1–42).
- Conn, S. R., & Rieke, M. L. (Eds.). (1994). *The 16PF fifth edition technical manual*. Champaign, IL: Institute for Personality and Ability Testing.
- Costa, P. T., & McCrae, R. R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.
- Diekmann, J., & König, C. J. (2015). Personality testing in personnel selection: Love it? Leave it? Change it! In I. Nikolaou & J. K. Oostrom (Eds.), *Employee recruitment, selection, and assessment: Contemporary issues for theory and practice* (pp. 117–135). Hove, UK: Psychology Press.
- Diekmann, J., König, C. J., & Alles, J. (2015). The role of neuroscience information for choosing a personality test: Not as seductive as expected. *International Journal of Selection and Assessment*, 23, 99–108. <http://doi.org/10.1111/ijsa.12099>
- Di Milia, L. (2004). Australian management selection practices: Closing the gap between research findings and practice. *Asia Pacific Journal of Human Resources*, 42, 214–228. <http://doi.org/10.1177/1038411104045364>
- Fanelli, D. (2010a). Do pressures to publish increase scientists' bias? An empirical support from US States data. *PLoS ONE*, 5, e10271. <http://doi.org/10.1371/journal.pone.0010271>
- Fanelli, D. (2010b). “Positive” results increase down the hierarchy of the sciences. *PLoS ONE*, 5, e10068. <http://doi.org/10.1371/journal.pone.0010068>
- Fanelli, D. (2011). Negative results are disappearing from most disciplines and countries. *Scientometrics*, 90, 891–904. <http://doi.org/10.1007/s11192-011-0494-7>

- Furnham, A. (2008). HR professionals' beliefs about, and knowledge of, assessment techniques and psychometric tests. *International Journal of Selection and Assessment*, *16*, 300–305. <http://doi.org/10.1111/j.1468-2389.2008.00436.x>
- Furnham, A., Moutafi, J., & Crump, J. (2003). The relationship between the revised NEO-personality inventory and the Myers-Briggs Type Indicator. *Social Behavior and Personality: An International Journal*, *31*, 577–584. <http://doi.org/10.2224/sbp.2003.31.6.577>
- Gangestad, S., & Snyder, M. (1985). “To carve nature at its joints”: On the existence of discrete classes in personality. *Psychological Review*, *92*, 317–349. <http://doi.org/10.1037/0033-295X.92.3.317>
- Herrmann International. (2015). HBDI Assessment. Retrieved October 20, 2015, from <http://www.herrmannsolutions.com/hbdi-assessment/>
- Hogan, R. (1986). *Manual for the Hogan Personality Inventory*. Minneapolis, MN: National Computer Systems.
- Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology*, *85*, 869–879. <http://doi.org/10.1037/0021-9010.85.6.869>
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114–156). New York, NY: Guilford Press.
- Judge, T. A., Heller, D., & Mount, M. K. (2002). Five-factor model of personality and job satisfaction: A meta-analysis. *Journal of Applied Psychology*, *87*, 530–541. <http://doi.org/10.1037/0021-9010.87.3.530>
- Jung, C. G. (1960). *Psychologische Typen [Psychological types]* (9th ed.). Zürich, Switzerland: Rascher.

- Keil, F. C. (2006). Explanation and understanding. *Annual Review of Psychology*, *57*, 227–254. <http://doi.org/10.1146/annurev.psych.57.102904.190100>
- Klehe, U.-C. (2004). Choosing how to choose: Institutional pressures affecting the adoption of personnel selection procedures. *International Journal of Selection and Assessment*, *12*, 327–342. <http://doi.org/10.1111/j.0965-075X.2004.00288.x>
- König, C. J., Klehe, U., Berchtold, M., & Kleinmann, M. (2010). Reasons for being selective when choosing personnel selection procedures. *International Journal of Selection and Assessment*, *18*, 17–27. <http://doi.org/10.1111/j.1468-2389.2010.00485.x>
- Kundoor, V., & Mueen, A. K. K. (2010). Uncovering negative results: Introducing an open access journal “Journal of Pharmaceutical Negative Results.” *Pharmacognosy Journal*, *2*, 365–367. [http://doi.org/10.1016/S0975-3575\(10\)80016-1](http://doi.org/10.1016/S0975-3575(10)80016-1)
- Landis, R. S., James, L. R., Lance, C. E., Pierce, C. A., & Rogelberg, S. G. (2014). When is nothing something? Editorial for the null results special issue of Journal of Business and Psychology. *Journal of Business and Psychology*, *29*, 163–167. <http://doi.org/10.1007/s10869-014-9347-8>
- Lombrozo, T. (2006). The structure and function of explanations. *Trends in Cognitive Sciences*, *10*, 464–470. <http://doi.org/10.1016/j.tics.2006.08.004>
- Malle, B. F. (2004). *How the mind explains behavior: Folk explanations, meaning, and social interaction*. Cambridge, MA: MIT Press.
- Marston, W. M. (1979). *Emotions of normal people*. Minneapolis, MN: Persona Press.
- McCrae, R. R., & Costa, P. T. J. (1997). Personality trait structure as a human universal. *American Psychologist*, *52*, 509–516. <http://doi.org/10.1037/0003-066X.52.5.509>
- Mitchell, G. (2012). Revisiting truth or triviality: The external validity of research in the psychological laboratory. *Perspectives on Psychological Science*, *7*, 109–117. <http://doi.org/10.1177/1745691611432343>

- Muñiz, J., & Fernández-Hermida, J. R. (2010). La opinión de los psicólogos españoles sobre el uso de los tests [Spanish psychologists opinions on test use]. *Papeles Del Psicólogo*, *31*, 108–121.
- Poropat, A. E. (2009). A meta-analysis of the Five-Factor model of personality and academic performance. *Psychological Bulletin*, *135*, 322–338. <http://doi.org/10.1037/a0014996>
- Robins, R. W., John, O. P., & Caspi, A. (1998). The typological approach to studying personality. In R. B. Cairns, L. R. Bergman, & J. Kagan (Eds.), *Methods and models for studying the individual* (pp. 135–160). Thousand Oaks, CA: Sage.
- Ryan, A. M., McFarland, L., Baron, H., & Page, R. (1999). An international look at selection practices: Nation and culture as explanations for variability in practice. *Personnel Psychology*, *52*, 359–391. <http://doi.org/10.1111/j.1744-6570.1999.tb00165.x>
- Ryan, A. M., & Sackett, P. R. (1987). A survey of individual assessment practices by I/O psychologists. *Personnel Psychology*, *40*, 455–488. <http://doi.org/10.1111/j.1744-6570.1987.tb00610.x>
- Ryan, A. M., & Sackett, P. R. (1992). Relationships between graduate training, professional affiliation, and individual psychological assessment practices for personnel decisions. *Personnel Psychology*, *45*, 363–387. <http://doi.org/10.1111/j.1744-6570.1992.tb00854.x>
- Rynes, S. L., Colbert, A. E., & Brown, K. G. (2002). HR Professionals' beliefs about effective human resource practices: Correspondence between research and practice. *Human Resource Management*, *41*, 149–174. <http://doi.org/10.1002/hrm.10029>
- Salgado, J. F. (1997). The Five Factor model of personality and job performance in the European Community. *Journal of Applied Psychology*, *82*, 30–43. <http://doi.org/10.1037/0021-9010.82.1.30>
- Sanders, K., van Riemsdijk, M., & Groen, B. (2008). The gap between research and practice: A replication study on the HR professionals' beliefs about effective human resource



practices. *International Journal of Human Resource Management*, 19, 1976–1988.

<http://doi.org/10.1080/09585190802324304>

Schuler, H., Hell, B., Trapmann, S., Schaar, H., & Boramir, I. (2007). Die Nutzung psychologischer Verfahren der externen Personalauswahl in deutschen Unternehmen: Ein Vergleich über 20 Jahre [Use of personnel selection instruments in German organizations during the last 20 years]. *Zeitschrift für Personalpsychologie*, 6, 60–70. <http://doi.org/10.1026/1617-6391.6.2.60>

Shapiro, D. L., Kirkman, B. L., & Courtney, H. G. (2007). Perceived causes and solutions of the translation problem in management research. *Academy of Management Journal*, 50, 249–266. <http://doi.org/10.5465/AMJ.2007.24634433>

Smith, E. R., & Queller, S. (2008). Mental representations. In A. Tesser & N. Schwarz (Eds.), *Blackwell handbook of social psychology: Intraindividual processes* (pp. 111–133). New York, NY: Wiley.

Taylor, P., Keelty, Y., & McDonnell, B. (2002). Evolving personnel selection practices in New Zealand organisations and recruitment firms. *New Zealand Journal of Psychology*, 31, 8–18.

Tenhiälä, A., Giluk, T. L., Kepes, S., Simón, C., Oh, I.-S., & Kim, S. (2016). The research-practice gap in human resource management: A cross-cultural study. *Human Resource Management*, 55, 179-200. <http://doi.org/10.1002/hrm.21656>

The Insights Group Limited. (2015). Insights Discovery Personal Profile. Retrieved October 20, 2015, from <https://www.insights.com/845/insights-discovery-personal-profile.html>

Trapmann, S., Hell, B., Hirn, J.-O. W., & Schuler, H. (2007). Meta-analysis of the relationship between the Big Five and academic success at university. *Zeitschrift für Psychologie/ Journal of Psychology*, 215, 132–151. <http://doi.org/10.1027/0044-3409.215.2.132>

- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. *Science*, *185*, 1124–1131. <http://doi.org/10.1126/science.185.4157.1124>
- Van den Akker, A. L., Deković, M., Asscher, J. J., Shiner, R. L., & Prinzie, P. (2013). Personality types in childhood: Relations to latent trajectory classes of problem behavior and overreactive parenting across the transition into adolescence. *Journal of Personality and Social Psychology*, *104*, 750–764. <http://doi.org/10.1037/a0031184>
- van Dick, R. (2015). Registered reports, advance articles online, and the way ahead. *Journal of Personnel Psychology*, *14*, 1–3. <http://doi.org/10.1027/1866-5888/a000140>
- York, K. L., & John, O. P. (1992). The four faces of Eve: A typological analysis of women's personality at midlife. *Journal of Personality and Social Psychology*, *63*, 494–508. <http://doi.org/10.1037/0022-3514.63.3.494>

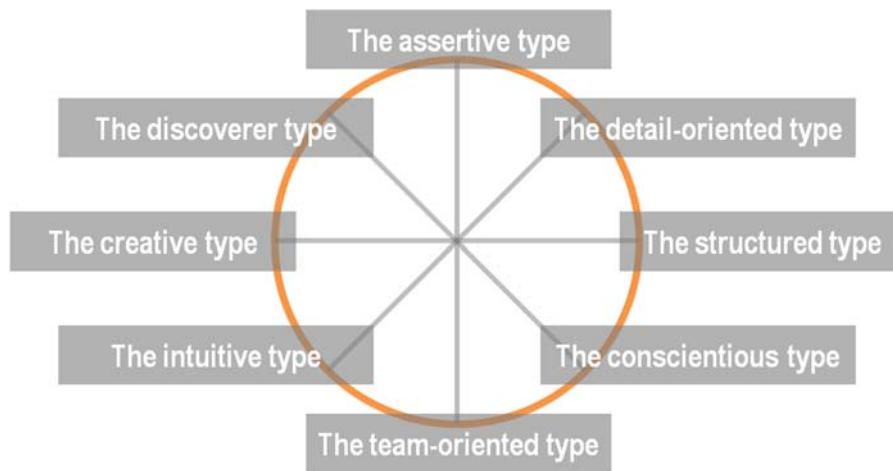
**Tables**

Table 1

*Overall test rating, standard deviations, and effect sizes for sample and subsamples.*

Sample	Test version	<i>n</i>	<i>M</i>	<i>SD</i>	<i>d</i>
Overall	Statistics	142	3.92	1.06	0.10
	Theory	137	4.03	1.22	
HR experts	Statistics	44	3.58	1.06	0.04
	Theory	45	3.62	1.15	
Business management students	Statistics	53	3.83	0.95	0.05
	Theory	44	3.88	1.02	
Psychology students	Statistics	45	4.35	1.06	0.18
	Theory	48	4.56	1.29	

**Figures**



*Figure 1.* Representation of traits in the flyer of the type-based condition.

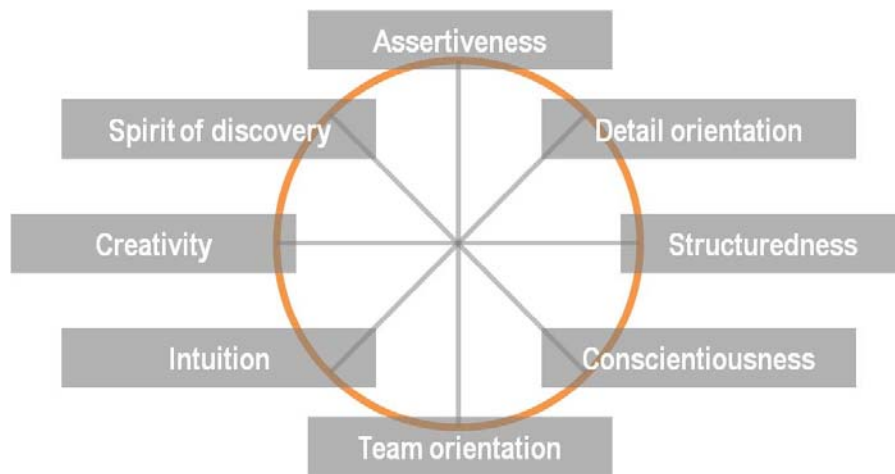


Figure 2. Representation of traits in the flyer of the dimension-based condition.

### The P-WIN<sup>®</sup> profile of John Doe

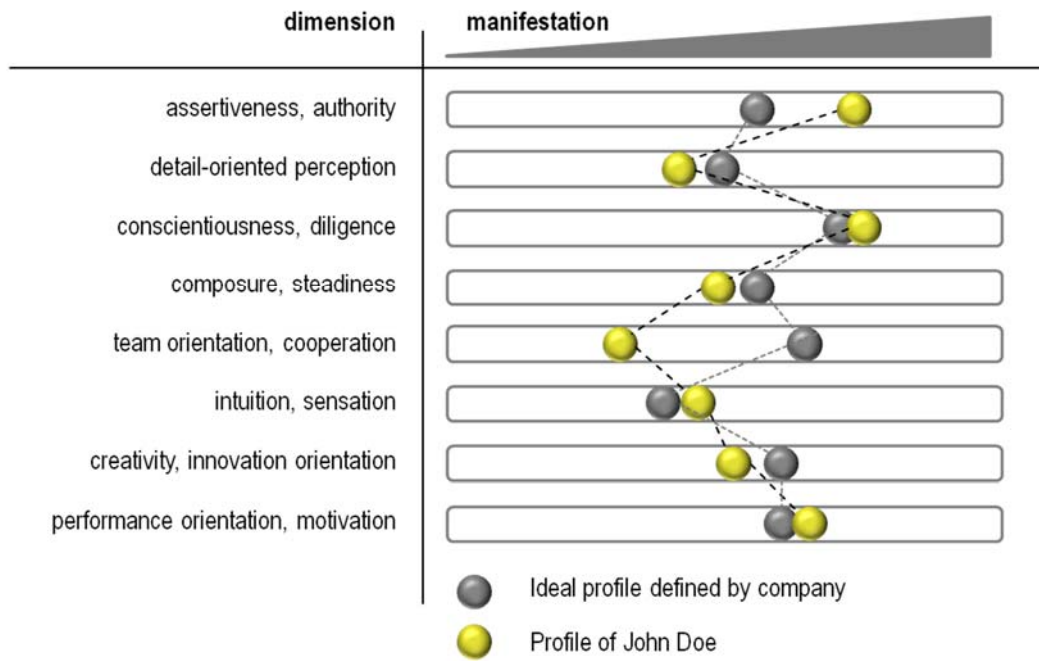
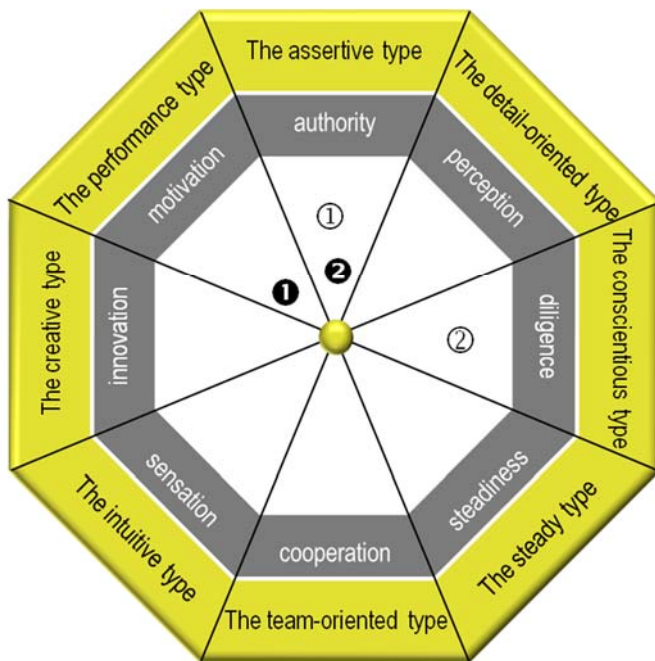


Figure 3. Representation of traits in the results report of the dimension-based condition.

### The P-WIN® profile of John Doe



① Primary preference: This type conforms most with the personality; it symbolizes the natural, inherent and learned behavior patterns.

② Secondary preference: This type characterizes the behavior in stressful situations. There may be a difference to the primary preference (but it does not have to be).

❶ / ❷ Ideal preferences defined by company.

Figure 4. Representation of traits in the results report of the type-based condition.