

Article

# Comparing Assessment Methods of Attribute Importance in Teachers' Decisions: The Importance of Different Criteria for Tracking Recommendations after Primary School

Katrin Lintorf <sup>1,\*</sup>, Stefanie van Ophuysen <sup>2</sup> and Igor Osipov <sup>3</sup>

<sup>1</sup> Department Erziehungs- und Sozialwissenschaften, Universität zu Köln, Albertus-Magnus-Platz, 50923 Köln, Germany

<sup>2</sup> Institut für Erziehungswissenschaft, Westfälische Wilhelms-Universität Münster, Georgskommende 33, 48143 Münster, Germany; vanOphuysen@uni-muenster.de

<sup>3</sup> Institut für Bildungsforschung in der School of Education, Bergische Universität Wuppertal, Gaußstr. 20, 42119 Wuppertal, Germany; osipov@uni-wuppertal.de

\* Correspondence: klintorf@uni-koeln.de

**Abstract:** The importance of different criteria for tracking recommendations is usually inferred using regression weights as a cross-student measure. The few studies that have applied alternative approaches or differentiated between student groups sometimes reach different conclusions. According to research on judgment and decision making (JDM), different methods operationalize different facets of importance. Given this, we investigate whether the importance of criteria for tracking recommendations depends on a direct vs. indirect operationalization (regression weights vs. ratings). A total of 181 teachers selected four students from their most recent fourth-grade class using a 2 × 2 design (certain vs. uncertain qualification for the *Realschule* (vocational track) vs. the *Gymnasium* (academic track)). Then, they reported on the level and the importance of predetermined criteria for each student. Contrary to JDM research, we found few method-related differences, but striking differences between cases with a certain vs. an uncertain qualification. For the latter, the importance of the criteria is more homogeneous, the regression prediction is less successful and the importance varies with the dependent variable in the regression (actual recommendation vs. perceived qualification). We conclude that further research should focus on uncertain cases rather than method-related differences and suspect that, in uncertain cases, the formation of the recommendation is a multistage decision process.

**Keywords:** attribute importance; judgment and decision making; dominance analysis; tracking recommendation; school track



**Citation:** Lintorf, K.; van Ophuysen, S.; Osipov, I. Comparing Assessment Methods of Attribute Importance in Teachers' Decisions: The Importance of Different Criteria for Tracking Recommendations after Primary School. *Educ. Sci.* **2021**, *11*, 566. <https://doi.org/10.3390/educsci11100566>

Academic Editor: Randall S. Davies

Received: 30 July 2021

Accepted: 16 September 2021

Published: 23 September 2021

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Between-school tracking is one measure used to deal with heterogeneity among students that is employed worldwide. While tracked and comprehensive school systems have their proponents, evidence disfavors tracking for several reasons [1]. In particular, tracking maintains or increases social inequality, more so as the age at the time of transition decreases [2]. This can be attributed to two factors [3]: First, access to the different tracks depends on social background. Second, school tracks act as differential learning environments. So, the choice of a secondary school track is a significant decision for the child's future educational path. This is particularly true in Germany and, in international comparisons, Germany is characterized by strong social inequality [4], as well as by an early transition to secondary school [5]. Although the German school system explicitly allows for changes of school track in lower secondary education [6], this rarely happens and downward mobility dominates when it does [7].

Although teachers in all German federal states are obliged to give school track recommendations, in most states, the final choice is the parents' decision [8]. Nevertheless, the teachers' tracking recommendation is an important guide for parents. The recommendation, as well as the counseling that precedes the recommendation, informs parental

considerations. Thus, it determines the expectations for success that parents have regarding the future educational path of their child [9]. Furthermore, the recommendation is a central determinant of the parents' final choice. This is even true regardless of its binding nature [10]. This fact has motivated extensive research on the formation of the tracking recommendation and the criteria used by the teachers [11]. Characteristics that teachers should (or should not) take into account in their recommendations are only very roughly defined [8]. Therefore, non-performance-related characteristics (e.g., social background) make an additional contribution to the prediction of the recommendation when controlling legal criteria (e.g., school performance and work behavior) [11].

However, the predictive power of different characteristics and their relation to each other could also be a question of methodological approach. This is suggested by the results for attribute importance found in the research on judgment and decision making (JDM). This branch of research uses diverse methods to investigate attribute importance. In contrast, research on the importance of attributes for tracking recommendations is dominated by one method. In regression analyses, the teachers' recommendations are predicted by different criteria generalized to all students. Few studies have used different methods and/or differentiated between groups of students. Some of these studies come to different conclusions about the importance of attributes for tracking recommendations than the studies with the dominating method. Therefore, this paper examines whether the importance of the criteria of the tracking recommendation depends on the choice of method and the type of student. To this end, we bring together sociologically or educationally oriented research on criteria of the tracking recommendation (Section 3) with more economically oriented research on JDM (Section 2).

## 2. Judgment and Decision Making (JDM)

### 2.1. Theoretical Foundations of Attribute Importance in Research on JDM

Cognitive psychology considers decision-making against the background of normative, descriptive and prescriptive theories. Normative and prescriptive theories relate to ideal or optimal decisions. In contrast, descriptive decision analysis aims to reconstruct and understand actual human decisions [12]. Decisions are the result of choosing between at least two options with (different) values for specific attributes [13]. JDM research determines the importance of the attributes for the choice using a variety of methods. They are broadly distinguished into explicit vs. implicit weighting methods ([14], also decomposed vs. holistic [15] or direct vs. indirect [16]). In the first case, the decision-maker directly provides information on the importance of the attribute. In the second case, attribute importance is (statistically) inferred from an overall judgment of the alternatives.

Consumer research has long been concerned that different methods of measuring importance lead to different conclusions about the importance of product attributes [17,18]. In their review, van Ittersum et al. argue that these different measures represent different facets of the construct [19]. Drawing on the work of Myers and Alpert [20] they distinguish between three facets: "The *salience* of an attribute represents the *importance* of the attribute in memory. The *relevance* of an attribute represents the *importance* of the attribute to the individual based on personal values and desires. Finally, the *determinance* of an attribute represents the *importance* of the attribute in judgment and choice." [19] (p. 1180, original emphasis). As evidence for this distinction, the authors demonstrate that methods of the same facet correlate higher than methods of different facets.

Consequently, attribute-elicitation methods serve as measures of the *salience* facet. In these methods, researchers do not specify the criteria to be evaluated (e.g., free-elicitation techniques) and use the order of their nomination as the measure for importance. In contrast, methods measuring the *relevance* use a preselected set of attributes. They determine attribute importance by directly asking (e.g., direct-rating method) or inferring it from the information search behavior of the subjects (e.g., information-display-board method). Methods measuring *determinance* infer the importance of given attributes from the statistical relation between the attributes' values and the overall evaluation of the respective objects

(e.g., conjoint method) [19]. Therefore, even criteria that are used unconsciously can be identified using a determinance detection methodology.

Since determinance looks at the overall assessment or actual choice of an object, determinance may at first glance seem the most appropriate way to determine attribute importance. However, a multi-method approach to assessing the importance is advocated [13] since each facet/each group of methods has its specific (dis)advantages. Methods of *determinance* assessment that work with criteria provided by the researchers run the risk of overlooking important criteria. In turn, this can influence the estimation of the statistical relationship between the decision and the attributes under investigation (e.g., beta weights in regression analysis). Second, determinance is dependent on the variance of attributes. Small variances of an attribute (e.g., due to real-life situations) might lead to statistical underestimation of its importance. Third, a strong determinance of a criterion implies that it is statistically predictive, but this does not necessarily reflect how the information was actually processed by the decision-makers [21].

Like methods of determinance assessment, methods for measuring *relevance* also use criteria provided by the researchers. Again, there is a risk of overlooking important criteria. Furthermore, some relevance measures are direct or explicit measures and, as such, they are more susceptible to bias than indirect or implicit measures. It is conceivable that commonly known biases are caused by efforts to resolve cognitive dissonance [15] or to meet social desirability. In particular, relevance measures are susceptible to specific biases in weight assessment (e.g., proxy overweighting bias [16]).

As measures of *salience* are also usually direct measures, they are also susceptible to bias. Furthermore, these measures may be subject to memory distortions. A specific memory distortion might result from a small variance, as with determinance measures. In decision making, cognitive effort and time are mostly devoted to those attributes in which objects differ greatly. This should result in a higher quality and quantity of information processing. In turn, more intensively processed information is remembered as more important [22].

## 2.2. Application of Research on Judgment and Decision Making (JDM) to Tracking Recommendations

Research on JDM is strongly rooted in the economic sciences [23] and has a focus on consumer decisions. Comparably few studies investigate health, societal or educational decisions. Yet, consumer decisions and tracking recommendations—or broadly speaking: the professional decisions of educators—differ in at least three ways from consumer decisions. First, while consumers make decisions for themselves, the teachers are making surrogate decisions for their students. Second, consumers make a selection decision, while teachers make an allocation decision [24]. Third, as a result, the attributes considered belong to different entities. Consumers focus on the attributes of the decision options, whereas teachers focus on the attributes of the students for whom the decision is to be made.

Nevertheless, it seems reasonable that the results of research on JDM are also transferable to the professional decisions made by teachers. Linking research on school track recommendations to research on JDM was already successful when using the methodological paradigm MouseLab from Payne et al. [25,26]. Furthermore, at least one study based on JDM research compares different methods used with regard to tracking recommendations [27]. In this study, 16 first-year psychology students and 10 teachers provided recommendations to 60 fictitious pupils. The decisions were based on vignettes with information on the same characteristics (e.g., report marks, learning skills). The scores for the attributes were randomly assigned so that attribute values were uncorrelated. The authors determined the attribute importance for each of the participating subjects in terms of (1) regression weights, (2) importance ratings and (3) frequency of denomination in verbal protocols. For the participating teachers, they found only a medium correlation between the results from the three methods ( $r_{\text{Mdn}} = 0.54$ ). The correlation differed between individuals.

However, two factors could limit the generalizability of the results. First, the results may have a low ecological validity because they are based on fictitious students but see

also [28]. The effort to create randomly arranged vignettes, in particular, could have led to atypical student cases and, thus, atypical decision-making behavior. Furthermore, research on the tracking recommendation suggests that a distinction between different types of students might be fruitful because teachers seem to adapt their decision behavior to the respective student case (Section 3). Second, the authors were working with ad hoc generated student attributes. However, current research on the tracking recommendation allows for deriving characteristics theoretically and/or empirically (Section 3). This reduces the risk of neglecting important attributes (see the disadvantages of determinance and relevance in Section 2.1). Therefore, the following section reviews the research on the criteria for the tracking recommendation.

### 3. Criteria of the Tracking Recommendation at Primary School Transition

Research on the criteria for tracking recommendations can be divided into two groups: those with a classical approach and those with alternative approaches. We will compare their results after introducing the core features of both approaches.

In the *classical approach*, the teacher's tracking recommendation is predicted through a regression analysis based on various criteria that are predetermined by the researcher. The information is typically collected using self-reports from large samples of students or parents [29]. Teacher ratings of the characteristics are not used in these kinds of studies. The relative importance of the criteria is indirectly determined and inferred from the size of the regression weights. Thus, the classical approach relates teachers' judgments to the levels of the student characteristics. In terms of research on judgment and decision making (JDM), this reflects the *determinance* of the criteria.

The *alternative approaches* include interview-based and experimental studies. In the interview studies, researchers either explicitly ask about the criteria considered by the teachers [30] or implicitly deduce them from the teachers' case descriptions [31]. Thus, the data is based on retrospection. The importance of the criteria is determined on the basis of the order and frequency of nomination. This corresponds to the *salience* measure in research on JDM. The experimental studies follow a social cognition approach [32,33] and use dual process models as a theoretical basis [34]. According to these models, a judgment is the result of an automatic or a controlled strategy of information processing. The more inconsistent an information situation is, the more likely judges are to use a controlled strategy and consider more information. From a methodological point of view, the experimental studies draw on MouseLab, a process-tracing method from decision research [25] and ask teachers to give a tracking recommendation for fictional children. To this end, they are provided with covered information on grades, work and social behavior, as well as family background. Two properties of the information search behavior are used as indicators for the importance of the criteria: the frequency and order of information retrieval. Böhmer et al. combine this with a direct rating method in order to capture the subjectively assessed importance. Both may be classified as measures of *relevance* [19].

Although the alternative approaches make use of more than one facet of importance, they have common features that distinguish them from the determinance measure within the classical approach. They directly determine the importance of various criteria based on teachers' reports or criteria-related actions. Thereby, the data collection is sometimes based on specific student cases [31,35] and sometimes on global appraisals [30,36].

Regardless of the approach and operationalization, *performances in the main subjects* proved to be the strongest predictor of the recommendation [10,33,35–37]. Results for *work behavior*, which teachers should consider according to the official guidelines [8], are more differentiated. Studies taking an alternative approach emphasize the importance of work behavior [30,31,33], while, in contrast, studies using the classical approach come to inconclusive results. This might be due to the use of varying operationalizations. Work behavior is most likely to prove a significant predictor when operationalized by work virtues or affective aspects [29,38,39].

The discrepancy between the approaches in relation to non-school criteria is particularly striking. Studies using the classical approach have repeatedly shown that the *social background of families* is significantly connected to the recommendation, e.g., [10,40–42]. Among the various operationalizations, the educational background of the parents tends to be the most important factor. In combination with the ISEI (International Socio-Economic Index of Occupational Status), it had at least an incremental predictive power [43,44] and sometimes it was the only predictor associated with the social situation [45,46]. In contrast, the socio-economic background was not mentioned at all in the interview studies. Information of this kind was also accessed last and less frequently in the experimental studies [33].

Discrepancies in the results on *migration background* are less pronounced. Like social background, the migration background was not mentioned or retrieved in studies that took an alternative approach [33]. This corresponds largely to the findings from the classical approach, where no effects, e.g., [10,47] or reduced effects [39] were found when controlling for grades (but see also [42]). Again, the effects seem to depend on the operationalization. Insignificant findings occurred for family language [10,48,49], significant findings resulted for parental country of birth [45,50] and for a combined predictor [40], but see also [51].

Even if teachers do not directly name social background and migration background as relevant criteria in the studies taking an alternative approach, these criteria indirectly contribute to the recommendation. In these studies, teachers considered process characteristics associated with family background. This is especially true for *parental support*. Studies using the classical approach hardly considered this characteristic. In the rare cases, it could not compensate for the predictive power of structural characteristics of the family background [42,46]. In contrast, parental support is of great importance according to the studies taking an alternative approach, especially when a child has an unclear achievement profile [33,35]. In contrast, the classical approach does not even differentiate between different student cases.

A comparison of the results from both approaches leads to distinctly different conclusions about the importance of the criteria considered, especially with regard to social background. This may be an effect of the research method. So far, only Böhmer et al. [33] have tested the hypothesis of method dependence within one sample. They found a medium correlation between a direct rating of importance and two measures of information processing (retrieval frequency:  $r = .53$ ; retrieval order:  $r = -0.49$ ), indicating an influence of the research methods on the results. As the study was limited to an alternative approach, a comparison of the results between the two approaches is still pending.

#### 4. Research Aims

The criteria for tracking recommendations have been well studied. Most of the findings stem from the classical research method, but studies from alternative approaches come to partially different results. To our knowledge, there are only two studies that have applied several research methods on the same sample in order to estimate methodological influences on the results [27,33]. They worked with vignettes; therefore, we question the transferability on real-life contexts (but see [28]). This led us to pursue the following research question using real student cases: Does the importance of given criteria for the tracking recommendation depend on the operationalization of importance (determinance vs. relevance)?

In answering this question, we also address several shortcomings of the actual research on tracking recommendations. (1) In the classical approach, research is based on self-reports from students and parents. However, some of the characteristics considered are not known to the teachers (e.g., parental educational background, socio-economic status [52]). Thus, significant correlations do not necessarily imply that teachers take the social background of students into account when making their recommendations. Consequently, we will use teachers' estimations of the characteristics instead of students' or parents' self-reports. (2) In the studies of the alternative research approach, the decision/judgment processes in

uncertain student cases proved to be more complex than in consistent student cases [33]. The classical approach does not consider student-specific decision processes. Consequently, we will provide differential analyses for the determinance and relevance operationalization of importance. (3) The classical approach measures importance in terms of determinance. As explained above, this is associated with a risk of overlooking important criteria. This seems to be especially true for parental support. Consequently, we will examine the importance of this criterion as a supplement to the classical criteria.

## 5. Materials and Methods

### 5.1. Sample and Design

We questioned 181 teachers (91.2% female, 0.6% missing) from 68 elementary schools in the German federal state North Rhine-Westphalia (NRW). Participation in the study was voluntary. As an incentive, we raffled eight cash prizes of 50 euros to be used for funding for the teachers' classes. The respondents had  $M = 16.59$  years of professional experience ( $min = 0.5$ ;  $max = 43$ ) and had accompanied a transition  $M = 4.08$  times ( $min = 1$ ;  $max = 15$ ).

We asked each teacher to provide information about children in their most recent fourth-grade class. To limit their effort, we asked them to select only four children. Nevertheless, to ensure a certain range, we specified the following two factors as selection criteria (a) perceived qualification for a school track (*Gymnasium* (academic track) vs. *Realschule* (vocational track)) and (b) certainty of judgment (certain vs. uncertain). Thus, each teacher was to select one child who was a certain case for *Realschule* and one a certain case for *Gymnasium*, in addition to one who was an uncertain case for *Realschule* and one an uncertain case for *Gymnasium* (Table 1).

**Table 1.** Instruction on the four case types.

		Perceived Qualification for a School Track	
		<i>Realschule</i> <sup>1</sup>	<i>Gymnasium</i> <sup>1</sup>
Certainty of judgment	Certain	“Think of a child from your last fourth-grade class for whom you personally found it easy to make the tracking recommendation and quickly decided to go with a recommendation for <i>Realschule</i> .”	“Think of a child from your last fourth-grade class for whom you personally found it easy to make the tracking recommendation and quickly decided to go with a recommendation for <i>Gymnasium</i> .”
	Uncertain	“Think of a child from your last fourth-grade class for whom you personally found the tracking recommendation difficult and for whom you vacillated between a recommendation for <i>Gymnasium</i> or <i>Realschule</i> for a longer period. In such cases, you can specify both school tracks in the recommendation in NRW. Now select a child for whom you would have recommended the <i>Realschule</i> if you had had to choose between the two tracks.”	“Think of a child from your last fourth-grade class for whom you personally found the tracking recommendation difficult and for whom you vacillated between a recommendation for <i>Gymnasium</i> or a <i>Realschule</i> for a longer period. In such cases, you can specify both school tracks in the recommendation in NRW. Now select a child for whom you would have recommended the <i>Gymnasium</i> if you had had to choose between the two tracks.”

<sup>1</sup> In the German school system, the *Gymnasium* represents the academic track leading to a university entrance qualification. The *Realschule* is one of the tracks that leads to a vocational qualification first.

The sequence of the four cases was balanced over four versions of the questionnaire. Their shares in the responded questionnaires were between 20.4% and 30.4%. With 181 teacher participants, this design resulted in 724 potential students. However, 33 students had to be excluded from the analyses due to implausible data. Thus, data remained for 691 children (certain RS:  $n = 181$ , certain GY:  $n = 181$ , uncertain RS:  $n = 155$ , uncertain GY:  $n = 174$ ).

### 5.2. Treatment of Missing Values

The rate of item nonresponse across the analysis variables varied from 0.1% to 4.8%. Missing values were imputed using R package mice [53] with two-level predictive mean matching for interval scales and two-level logistic regression for binary variables. The results of all subsequent analyses were pooled across five data sets with imputed missing values.

### 5.3. Research Instruments and Operationalization

Our analyses are based on the student characteristics typically used to predict school track recommendations (Section 3). We asked teachers to indicate the level of each characteristic for each student. This was used to operationalize the facet *determinance*. The following characteristics were included in the analyses (descriptives and internal consistency see Table 2):

- School performance was measured in the form of report card grades in German (=average grade of language use, reading and orthography) and mathematics in the first semester of the fourth grade. For better interpretability, we reversed the polarity (1 = poor, 5 = very good; insufficient grades did not occur).
- Work behavior was measured using four items (example: “diligence in work behavior [e.g., neatness, orderliness, handwriting]”, 1 = low, 5 = high).
- We captured social background via two items on the child’s home-family environment (proximity to education, financial security; 1 = weak, 5 = high). To operationalize parental support, we asked about three items: to what extent parents could provide professional, organizational and financial support if needed (1 = not at all, 5 = very good). The two variables were highly correlated ( $r = 0.81$ ) and highly correlated, almost identically, with the other analysis variables (Tables A1–A3). To avoid collinearity, we, therefore, combined the two highly redundant characteristics into the variable family background [54]. All analyses are based on this combined variable.
- Migration background was operationalized by the family language (0 = only German, 1 = no German/German and other languages). Despite previous insignificant findings (Section 3), we preferred using language over the parents’ countries of birth because instead of students or parents we interviewed the teachers who have easier access to the former information.

**Table 2.** Descriptive data of the level of student attributes.

	<i>n</i> Items	All Cases ( <i>N</i> = 691)			Certain Cases ( <i>N</i> = 362)			Uncertain Cases ( <i>N</i> = 329)		
		<i>M</i>	<i>SD</i>	$\alpha$	<i>M</i>	<i>SD</i>	$\alpha$	<i>M</i>	<i>SD</i>	$\alpha$
Actual recommendation	1	0.69	-	-	0.55	-	-	0.84	-	-
Perceived qualification	1	0.51	-	-	0.50	-	-	0.53	-	-
German grade	3	3.76	0.70	0.83	3.94	0.83	0.90	3.56	0.44	0.49
Math grade	1	3.70	0.84	-	3.88	0.92	-	3.50	0.69	-
Work behavior	4	3.88	0.87	0.87	4.10	0.86	0.89	3.64	0.83	0.82
Family background	5	3.61	0.86	0.89	3.72	0.88	0.91	3.49	0.81	0.87
Migration background	1	0.20	0.40	-	0.18	0.38	-	0.23	0.42	-

For coding see the description in Section 5.3.

In line with research on judgment and decision making, we used the estimates from logistic regressions (Section 2.1), in order to depict two aspects of determinance. We predicted the actual recommendation—a public judgment—as well as the perceived qualification—a rather personal judgment (each 1 = *Gymnasium*). In North Rhine-Westphalia, teachers can also give a limited recommendation for the *Gymnasium*. These cases were combined with those who had an unqualified recommendation for the *Gymnasium*.

Perceived qualification and actual recommendation coincided highly, especially in the certain cases (94.53%). Teachers only recommended the *Gymnasium* in 5.47% of the certain cases where their questionnaire response classified the child as a *Realschule* case. In the

uncertain cases, the agreement was lower (68.63%). The teacher's recommendation for the *Gymnasium* differed from the perceived qualification for *Realschule* in 31.31% of uncertain cases. The opposite was rarely the case (0.06%).

The measure of *relevance* (Section 2.1) is based on the teachers' subjective assessments of importance. The teachers indicated how important each of the criteria described above was in forming their recommendation for the respective child (1 = unimportant, 4 = very important). For reasons of comparability, social background and parental support were again combined into one family support criteria. Descriptives and internal consistencies are presented in Table 3. Correlations can be found in Tables A4–A6.

**Table 3.** Descriptive data of teacher importance ratings (relevance).

	<i>n</i> Items	All Cases ( <i>N</i> = 691)			Certain Cases ( <i>N</i> = 362)			Uncertain Cases ( <i>N</i> = 329)		
		<i>M</i> <sup>1</sup>	<i>SD</i>	$\alpha$	<i>M</i> <sup>1</sup>	<i>SD</i>	$\alpha$	<i>M</i> <sup>1</sup>	<i>SD</i>	$\alpha$
German grade	3	3.34 <sup>a</sup>	0.50	0.83	3.37 <sup>a</sup>	0.48	0.83	3.31 <sup>a</sup>	0.51	0.82
Math grade	1	3.43 <sup>b</sup>	0.55	-	3.45 <sup>b</sup>	0.54	-	3.40 <sup>a</sup>	0.56	-
Work behavior	4	3.44 <sup>b</sup>	0.47	0.82	3.48 <sup>b</sup>	0.47	0.82	3.41 <sup>a</sup>	0.48	0.83
Family background	5	2.22 <sup>c</sup>	0.61	0.86	2.18 <sup>c</sup>	0.60	0.87	2.26 <sup>b</sup>	0.61	0.85
Migration background	1	2.25 <sup>c</sup>	0.92	-	2.21 <sup>c</sup>	0.91	-	2.29 <sup>b</sup>	0.94	-

Response format: 1 = unimportant, 4 = very important. <sup>1</sup> Different letters indicate significant differences in the ratings as indicated by post hoc comparisons (Table A7).

#### 5.4. Analysis Strategy

All analyses were performed in R [55] on multiply imputed data sets and accounted for the hierarchical structure of the data. All analyses were carried out for the total sample and separately for the two sub-samples (certain vs. uncertain student cases).

The multilevel logistic regressions were performed using R packages lme4 [56] and mitml [57]. Unfortunately, only one significant slope variance could be identified (Tables A10 and A11). So, the correlations between different methods could not be calculated for nearly all predictors and all cases. Instead, we relied on a descriptive comparison of the predictor ranks.

To determine the ranks of the ratings, we first averaged the ratings across the four students for each predictor. We then ranked the predictors across all teachers using post hoc tests in a mixed-effects ANOVA using R packages lme4 [56], mitml [57] and MKmisc [58].

For the regression weights, we determined importance based on a dominance analysis [59], using R package dominanceanalysis [60]. This analyzes the importance of each predictor compared to all other predictors across all possible subsets of the predictors. A predictor is dominant if its additional contribution to  $R^2$  is higher than that of the predictor it is being compared to. The dominance analysis distinguishes three forms. The strictest form, *complete dominance*, exists if a predictor makes a higher contribution in all models. A weaker form, *conditional dominance*, exists if this is not true for all models, but is true on average for all models with the same number of predictors. The weakest form, *general dominance*, exists if a predictor makes a higher contribution on average. In our analyses, we only relied on the strictest form. Additionally, we used bootstrap samples to further validate our conclusions. We interpreted that one predictor dominated another if this was true for 70% of the bootstrap samples [59] and then assigned this predictor the higher rank.

## 6. Results

### 6.1. Direct Measures: Relevance

Information on importance in terms of relevance can be found in Table 3. Mixed-effects ANOVA revealed significant differences in the importance of the different characteristics (all cases:  $F(4, 77,660.80) = 359.18, p < 0.001$ ; certain cases:  $F(4, 425,100.00) = 366.92, p < 0.001$ ; uncertain cases:  $F(4, 5434.94) = 231.98, p < 0.001$ ).



Based on the post hoc tests, it is evident that grades and work behavior are reported as the most important criteria. In contrast, similar to interview and experimental studies (Section 3), family background and migration background are rated as “rather unimportant” ( $M \approx 2.00$ ). This evaluation differs significantly from that of the other characteristics. While these findings apply to both certain and uncertain cases, the results on the legally permissible criteria differ between cases: math grade and work behavior come first in the certain cases, while work behavior and both grades are equally important in the uncertain cases.

## 6.2. Indirect Measures: Determinance

Information on importance in terms of determinance can be found in Table 4 (actual recommendation) and Table 5 (perceived qualification). Migration background is an insignificant predictor of the actual recommendation in all models. However, the significance of the other predictors is case-specific. In the certain cases, the German grade is, by far, the most dominant predictor, followed by the math grade which dominates the family background. While work behavior is an insignificant predictor in the certain cases, it is by far the most dominant predictor in the uncertain cases. No further dominances could be identified among the other predictors. Overall, the prediction in the certain cases ( $R^2 = 0.72$ ) is better than in the uncertain cases ( $R^2 = 0.21$ ).

**Table 4.** Logistic regressions of actual recommendations for school track (*Gymnasium* [GY] = 1): odds ratios (OR) and average contribution to McFaddens  $R^2$  ( $R^2.m$ , general dominance) (determinance).

	All Cases ( $N = 691$ , $n_{GY} = 477$ )		Certain Cases ( $N = 362$ , $n_{GY} = 201$ )		Uncertain Cases ( $N = 329$ , $n_{GY} = 276$ )	
	OR	$R^2.m$ <sup>1</sup>	OR	$R^2.m$ <sup>1</sup>	OR	$R^2.m$ <sup>1</sup>
German grade	6.14 *	0.14 <sup>a</sup>	16.45 *	0.29 <sup>a</sup>	3.14 *	0.05 <sup>b</sup>
Math grade	2.80 *	0.09 <sup>b</sup>	4.50 *	0.19 <sup>b</sup>	2.21 *	0.04 <sup>b</sup>
Work behavior	1.46 *	0.06 <sup>b,c</sup>	1.57	0.14 <sup>b,c,d</sup>	2.60 *	0.09 <sup>a</sup>
Family background	1.80 *	0.05 <sup>c</sup>	2.79 *	0.10 <sup>c</sup>	1.61 *	0.03 <sup>b</sup>
Migration background	0.84	<0.01 <sup>d</sup>	0.96	<0.01 <sup>d</sup>	0.47	0.01 <sup>b</sup>
$R^2$ McFadden	0.34		0.72		0.21	

\*  $p < 0.05$ ; <sup>1</sup> Different letters indicate significant differences in the ratings as indicated by dominance analysis (Table A8).

**Table 5.** Logistic regressions of perceived qualifications for school track (*Gymnasium* [GY] = 1): odds ratios (OR) and average contribution to McFaddens  $R^2$  ( $R^2.m$ , general dominance) (determinance).

	All Cases ( $N = 691$ , $n_{GY} = 355$ )		Certain Cases ( $N = 362$ , $n_{GY} = 181$ )		Uncertain Cases ( $N = 329$ , $n_{GY} = 174$ )	
	OR	$R^2.m$ <sup>1</sup>	OR	$R^2.m$ <sup>1</sup>	OR	$R^2.m$ <sup>1</sup>
German grade	7.54 *	0.18 <sup>a</sup>	196.23 *	0.36 <sup>a</sup>	3.81 *	0.05 <sup>a</sup>
Math grade	3.37 *	0.12 <sup>a,b</sup>	21.00 *	0.25 <sup>a,b</sup>	2.35 *	0.04 <sup>a,b</sup>
Work behavior	1.74 *	0.08 <sup>b,c</sup>	7.36 *	0.20 <sup>b,c</sup>	1.79 *	0.05 <sup>a,b</sup>
Family background	1.80 *	0.05 <sup>c</sup>	3.73 *	0.10 <sup>c</sup>	1.57 *	0.02 <sup>b</sup>
Migration background	1.06	<0.01 <sup>d</sup>	2.06	<0.01 <sup>d</sup>	0.93	<0.01 <sup>c</sup>
$R^2$ McFadden	0.44		0.90		0.17	

\*  $p < 0.05$ ; <sup>1</sup> Different letters indicate significant differences in the ratings as indicated by dominance analysis (Table A9).

The results for perceived qualification (Table 5) largely agree with the results for the actual recommendation in the certain cases. This was to be expected given the high degree of agreement between the two dependent variables (Section 5.3). Only work behavior proves to be an additional significant predictor and the German grade is no longer dominant over the math grade. For the uncertain cases, the picture is partly different from the actual recommendation. Grades and work behavior rank equally high here. The German grade dominates the family background. Migration background is dominated by all other variables. Similar to the actual recommendation, all characteristics appear to be equally

important compared to the certain cases. Again, the prediction of the recommendation is better in the certain cases ( $R^2 = 0.90$ ) than in the uncertain cases ( $R^2 = 0.17$ ). Overall, the prediction of the perceived qualification succeeds slightly less for the uncertain cases and is clearly better for the certain cases than the prediction of the actual recommendation.

### 6.3. Comparison of Relevance and Determinance

Taken together, the results for both measures of importance are similar. First, the order of the criteria hardly changes. In both forms of measurement, the legally relevant criteria receive the highest importance, while family background and migration background are at most of secondary importance. Method-related differences are found only within the group of legally permissible criteria. Among the relevance measures, the math grade and work behavior are more important than the German grade. In the determinance measures, the two grades are generally of the highest importance, and sometimes more important than work behavior. The only exception is the dominance of work behavior in the prediction of the actual recommendation for the uncertain cases. Second, regardless of the method, the analyses differentiate between the importance of the criteria more strongly for the certain cases than for the uncertain criteria. Here, all criteria seem to be almost equally important.

However, two striking differences appear. First, no criterion is rated as unimportant in terms of relevance. This is true even for the migration background (all cases:  $t(174.58) = 37.75$ ,  $p < 0.001$ ; certain cases:  $t(177.90) = 35.75$ ,  $p < 0.001$ ; uncertain cases:  $t(140.63) = 33.56$ ,  $p < 0.001$ ), which does not make a significant predictive contribution in the logistic regressions. Second, the criteria can be divided into distinct groups according to their relevance. This is not true for their determinance: here, the groups overlap.

## 7. Discussion

The starting point of this study was the question of whether the importance of given criteria for the tracking recommendation depends on the type of operationalization. To answer the question, we compared ratings on the importance of predetermined criteria (direct measure: relevance) with weights determined in logistic regressions to predict actual school track recommendation vs. personally perceived qualification for a particular school track (indirect measure: determinance). Furthermore, we distinguished between cases with certain vs. uncertain school track recommendations.

Overall, the results from all analyses were quite similar. We found hardly any method-related differences. Thus, our results, obtained using real student cases, are not consistent with the results from studies based on vignettes [27,33]. Consistent with previous research findings (Section 3), grades and work behavior received the highest weight, while family background and migration background were less important.

However, our analyses also revealed differences between the measures of importance. Although the migration background turned out to be an insignificant predictor in the logistic regressions (i.e., low determinance), it was attributed some level of importance by the teachers in the ratings (i.e., medium relevance). This was unexpected as biases in the sense of social desirability or effects of cognitive dissonance were assumed to cause a devaluation of the relevance of non-performance-related characteristics, especially in direct assessment methods (Section 2.1). Obviously, these biases hardly apply here. However, the lack of effect in the logistic regressions could still be due to peculiarities of the methods used to assess determinance. Migration background has a comparatively low variance (Table 2), so its importance may be underestimated (Section 2.1). In addition, in a regression, weights are estimated by taking into account joint predictive contributions with other characteristics. Ratings, on the other hand, consider the sole importance of the queried characteristics.

According to our results, it is hardly necessary to capture importance (relevance vs. determinance) using multiple methods as is required in consumer research [19,61]. In the case of transition research, it seems to be much more important to consider different student cases (certain vs. uncertain cases) and to carefully choose the operationalization of the

recommendation (perceived qualification vs. actual recommendation). Both variations led to different results in the logistic regressions.

First, while the specified characteristics led to very high variance explanations for the certain cases, the opposite was true for the uncertain student cases. Although we attempted to supplement the criteria typically studied, it remains largely unclear on which characteristics the recommendations are based. However, in contrast to the few studies of the classical approach that took parental background into account [42,46], we had to combine parental support and family background because of collinearity (Section 5.3). This deviation from the previous state of research presumably occurred because of the change in the methodological approach. Previous studies relied on student and parent responses, whereas we used teacher assessments exclusively. Either teachers do not differentiate between the two constructs or they might infer one from the other due to a lack of information [52]. Overall, the low predictive performance in the models for the uncertain cases illustrates the disadvantage of relying on predetermined characteristics (Section 2.1).

Second, the results for the uncertain cases differ depending on whether the actual recommendation or perceived qualification is predicted. In predicting the actual recommendation, work behavior dominates all other predictors. However, in the prediction of the perceived qualification, the legally relevant criteria dominate as a group. This data pattern could indicate that the formation of a school track recommendation is a multi-stage decision process (for sequential strategies in diagnostics see [24]). The formation of a recommendation is a complex process of decision-making. In comparison to consumer research, there are only a few alternatives to choose from when making a school track recommendation, but there is a wide range of information that must be taken into account. Moreover, information is often vague and fraught with uncertainty [31]. Teachers may try to reduce the complexity of this decision-making situation by breaking down the decision. First, similar to the strategy “elimination by aspects” [62], they may consider which school track they *perceive* a child is fundamentally *qualified* for based on the most stable and predictive characteristics (i.e., achievement). In a further step, more variable and less predictive characteristics (e.g., work behavior) would then be considered for the *actual recommendation*. Their influence should be visible, above all, in the uncertain cases because in these cases no certain qualification could be derived on the basis of performance.

Even if the largely method-independent results seem unexpected against the background of previous research, the result is, nevertheless, a desirable one with regard to practice. Teachers do indeed judge as they say: their recommendation seems to be the result of a process in which information is weighed against each other in a deliberate and reasoned manner. This corresponds to the requirements of professional practice [63] and the controlled system of dual-process models [34,64]. The results may be due to the high-stakes nature of school track recommendations. Therefore, results may not be transferable to other situations, because teachers do indeed act heuristically and less analytically when decisions are less important and have fewer consequences [65].

Accordingly, diagnostics-related teacher education could and should sensitize future teachers to the properties of human decision making. In general, this includes the distinction into two systems of thought, the controlled system and the automatic system. In particular, teachers should be informed that the easily retrievable salient characteristics or aspects that are personally perceived as important are not necessarily those that are guiding the ultimate educational decision. While this is acceptable for easily revisable, short-term decisions, it is not true for long-term, hard-to-revise decisions such as the school track choice made after the teachers’ recommendation.

In addition, limitations must be taken into account when interpreting the findings. First, our conclusions are based on descriptive analyses alone. An inferential statistical analysis, as found in Harte and Koele [27], was not possible in the absence of significant slope variances (Section 5.4). Second, our results, based on a survey at only one point in time and in only one federal state, might not be generalizable. It is well known from consumer research that decisions are highly context dependent [66]. Transition research

also addresses such effects. For example, the importance of social background for the recommendation varies depending on its binding nature [51] and the advantage of certain groups of students changes over time [67] (but see also [10]). This leads to different levels of importance for the same criterion. However, it is not clear whether and to what extent this context dependency influences the answer to the research question pursued here, because context or time dependence is not the same as method dependence.

Third, we have concluded that the results on the importance of attributes in the case of school track recommendation are robust to the use of different operationalizations (and, thus, different facets of importance). This conclusion might be somewhat hasty, as we did not consider the facet of salience. However, results similar to those reported above can be expected for salience. Characteristics that make a student a “certain case” are those relating to consistently good or poor achievement. In free-elicitation methods, these characteristics should, therefore, be salient in memory and should come up first and more frequently than other characteristics. In the uncertain cases, achievement alone is not the determining characteristic. In terms of the multistage decision process outlined above, we expect that a great amount of attention would first be paid to the either inconsistent or consistently medium achievements. Given its low decision support, other features would then be given a similar amount of attention. Therefore, equal naming frequencies but distinct naming orders would be expected.

Future research could well continue to address methodological issues in capturing the importance of teachers’ decision criteria as relates to tracking recommendations. So far, it remains open whether the finding of method independence can be generalized to the multitude of other pedagogical decisions (e.g., assignment of grades, determination of special educational needs, repetition vs. continuation in the subject matter). As expressed above, method independence could depend on the degree of bindingness of the pedagogical decision in question. This needs to be examined further. However, with regard to the school track recommendations, other research goals seem more urgent, especially the identification of recommendation criteria in uncertain cases and the search for causes of the differences in private vs. public expressions of opinion. In particular, further studies could explore the conjecture formulated above that our dependent variables mark different stages in a decision process.

**Author Contributions:** Conceptualization, S.v.O. and K.L.; methodology, K.L., S.v.O. and I.O.; formal analysis, I.O.; investigation, K.L.; resources, K.L. and S.v.O.; data curation, K.L. and I.O.; writing—original draft preparation, K.L.; writing—review and editing, K.L., S.v.O. and I.O.; visualization, K.L.; project administration, K.L.; funding acquisition, S.v.O. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Data are available upon request.

**Acknowledgments:** The authors would like to thank the participating teachers and the students of a course taught by Katrin Lintorf who assisted with the data collection.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

**Table A1.** Correlations of the level of student attributes (all cases,  $N = 691$ ).

	(1)	(2)	(3)	(4)	(5)	(6a)	(6b)	(6)	(7)
(1) Actual recommendation	-								
(2) Perceived qualification	0.97 *	-							
(3) German grade	0.51 *	0.61 *	-						
(4) Math grade	0.42 *	0.52 *	0.52 *	-					
(5) Work behavior	0.40 *	0.50 *	0.62 *	0.39 *	-				
(6a) Social background	0.36 *	0.38 *	0.42 *	0.26 *	0.32 *	-			
(6b) Parental support	0.33 *	0.38 *	0.42 *	0.26 *	0.35 *	0.81 *	-		
(6) Family background	0.36 *	0.40 *	0.44 *	0.28 *	0.35 *	0.93 *	0.97 *	-	
(7) Migration background	-0.15 *	-0.14 *	-0.08	-0.05	0.01	-0.30 *	-0.28 *	-0.30 *	-

\*  $p < 0.05$ . Correlations between two dichotomous variables are tetrachoric correlations.

**Table A2.** Correlations of the level of student attributes (certain cases,  $N = 362$ ).

	(1)	(2)	(3)	(4)	(5)	(6a)	(6b)	(6)	(7)
(1) Actual recommendation	-								
(2) Perceived qualification	0.99 *	-							
(3) German grade	0.80 *	0.85 *	-						
(4) Math grade	0.69 *	0.75 *	0.66 *	-					
(5) Work behavior	0.68 *	0.73 *	0.76	0.54 *	-				
(6a) Social background	0.55 *	0.55 *	0.54 *	0.41 *	0.47 *	-			
(6b) Parental support	0.54 *	0.55 *	0.54 *	0.41 *	0.48 *	0.84 *	-		
(6) Family background	0.57 *	0.57 *	0.56 *	0.42 *	0.49 *	0.94 *	0.97 *	-	
(7) Migration background	-0.22 *	-0.21 *	-0.09	-0.08	-0.02	-0.28 *	-0.27 *	-0.28 *	-

\*  $p < 0.05$ . Correlations between two dichotomous variables are tetrachoric correlations.

**Table A3.** Correlations of the level of student attributes (uncertain cases,  $N = 329$ ).

	(1)	(2)	(3)	(4)	(5)	(6a)	(6b)	(6)	(7)
(1) Actual recommendation	-								
(2) Perceived qualification	0.86 *	-							
(3) German grade	0.24 *	0.31 *	-						
(4) Math grade	0.19 *	0.24 *	0.05	-					
(5) Work behavior	0.31 *	0.29 *	0.28 *	0.07	-				
(6a) Social background	0.19 *	0.19 *	0.17 *	0.00	0.12 *	-			
(6b) Parental support	0.17 *	0.20 *	0.14 *	-0.02	0.14 *	0.78 *	-		
(6) Family background	0.19 *	0.21 *	0.16 *	-0.01	0.14 *	0.92 *	0.96 *	-	
(7) Migration background	-0.20 *	-0.09	0.00	0.01	0.07	-0.32 *	-0.29 *	-0.32 *	-

\*  $p < 0.05$ . Correlations between two dichotomous variables are tetrachoric correlations.

**Table A4.** Correlations of teacher importance ratings (all cases,  $N = 691$ ).

	(1)	(2)	(3)	(4)	(5)	(6a)	(6b)	(6)	(7)
(1) Actual recommendation	-								
(2) Perceived qualification	0.97 *	-							
(3) German grade	0.10 *	0.10 *	-						
(4) Math grade	0.09 *	0.06	0.72 *	-					
(5) Work behavior	0.16 *	0.16 *	0.42 *	0.35 *	-				
(6a) Social background	0.03	0.03	0.06	0.06	0.08	-			
(6b) Parental support	0.00	-0.03	0.12 *	0.11 *	0.11	0.67 *	-		
(6) Family background	0.01	-0.01	0.11 *	0.10 *	0.11 *	0.88 *	0.95 *	-	
(7) Migration background	-0.02	-0.02	0.12 *	0.10 *	0.00	0.46 *	0.49 *	0.52 *	-

\*  $p < 0.05$ . Correlations between two dichotomous variables are tetrachoric correlations.

**Table A5.** Correlations of teacher importance ratings (certain cases,  $N = 362$ ).

	(1)	(2)	(3)	(4)	(5)	(6a)	(6b)	(6)	(7)
(1) Actual recommendation	-								
(2) Perceived qualification	0.99 *	-							
(3) German grade	0.14 *	0.12 *	-						
(4) Math grade	0.11 *	0.10	0.74 *	-					
(5) Work behavior	0.23 *	0.26 *	0.48 *	0.39 *	-				
(6a) Social background	0.03	-0.02	0.08	0.10	0.07	-			
(6b) Parental support	-0.05	-0.09	0.10	0.13 *	0.10	0.69 *	-		
(6) Family background	-0.02	-0.06	0.10	0.13 *	0.10	0.89 *	0.95 *	-	
(7) Migration background	-0.02	-0.05	0.10	0.12 *	0.04	0.49 *	0.53 *	0.56 *	-

\*  $p < 0.05$ . Correlations between two dichotomous variables are tetrachoric correlations.

**Table A6.** Correlations of teacher importance ratings (uncertain cases,  $N = 329$ ).

	(1)	(2)	(3)	(4)	(5)	(6a)	(6b)	(6)	(7)
(1) Actual recommendation	-								
(2) Perceived qualification	0.86 *	-							
(3) German grade	0.12 *	0.08	-						
(4) Math grade	0.10	0.02	0.71 *	-					
(5) Work behavior	0.16 *	0.07	0.35 *	0.30 *	-				
(6a) Social background	0.01	0.07	0.06	0.02	0.09	-			
(6b) Parental support	0.01	0.03	0.14 *	0.11	0.14 *	0.65 *	-		
(6) Family background	0.01	0.06	0.12 *	0.08	0.13 *	0.87 *	0.94 *	-	
(7) Migration background	-0.05	0.01	0.15 *	0.09	-0.02	0.43 *	0.44 *	0.48 *	-

\*  $p < 0.05$ . Correlations between two dichotomous variables are tetrachoric correlations.

**Table A7.** Results of the post hoc-tests in the mixed-effects ANOVA.

	All Cases ( $N = 691$ )			Certain Cases ( $N = 362$ )			Uncertain Cases ( $N = 329$ )		
	$M_{diff}$	$SD_{diff}$	$d$	$M_{diff}$	$SD_{diff}$	$d$	$M_{diff}$	$SD_{diff}$	$d$
<b>MB-FB</b>	0.03	0.64	0.04	0.03	0.68	0.04	0.02	0.80	0.03
<b>Ger-FB</b>	1.13 *	0.61	1.84	1.19 *	0.64	1.86	1.06 *	0.69	1.54
<b>Mat-FB</b>	1.21 *	0.64	1.89	1.27 *	0.65	1.93	1.15 *	0.74	1.56
<b>WB-FB</b>	1.22 *	0.60	2.05	1.29 *	0.62	2.08	1.15 *	0.67	1.71
<b>Ger-MB</b>	1.10 *	0.85	1.30	1.16 *	0.90	1.29	1.04 *	0.97	1.08
<b>Mat-MB</b>	1.18 *	0.87	1.37	1.24 *	0.89	1.39	1.13 *	1.01	1.12
<b>WB-MB</b>	1.20 *	0.88	1.36	1.26 *	0.90	1.40	1.12 *	1.01	1.11
<b>Mat-Ger</b>	0.08 *	0.27	0.30	0.08 *	0.30	0.25	0.09 *	0.36	0.25
<b>WB-Ger</b>	0.09 *	0.37	0.25	0.10 *	0.39	0.26	0.08	0.48	0.18
<b>WB-Mat</b>	0.01	0.45	0.03	0.03	0.47	0.06	0.00	0.56	-0.01

\*  $p < 0.05$ . Abbreviations: Ger = German grade, Mat = Math grade, WB = Work behavior, FB = Family background, MB = Migration background.

**Table A8.** Results from the dominance analysis for the actual recommendation (complete dominance).

I-j	All Cases (N = 691, n <sub>GY</sub> = 477)					Certain Cases (N = 362, n <sub>GY</sub> = 201)					Uncertain Cases (N = 329, n <sub>GY</sub> = 276)				
	D <sub>ij</sub>	SE <sub>Dij</sub>	p <sub>ij</sub>	p <sub>ji</sub>	p <sub>noij</sub>	D <sub>ij</sub>	SE <sub>Dij</sub>	p <sub>ij</sub>	p <sub>ji</sub>	p <sub>noij</sub>	D <sub>ij</sub>	SE <sub>Dij</sub>	p <sub>ij</sub>	p <sub>ji</sub>	p <sub>noij</sub>
Ger-Mat	0.88	0.21	0.76	0.00	0.24	0.87	0.22	0.73	0.00	0.27	0.57	0.39	0.38	0.24	0.38
Ger-WB	1.00	0.02	1.00	0.00	0.00	1.00	0.00	1.00	0.00	0.00	0.15	0.32	0.10	0.80	0.11
Ger-FB	0.99	0.05	0.99	0.00	0.01	0.98	0.11	0.95	0.00	0.05	0.66	0.37	0.49	0.17	0.34
Ger-MB	1.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	0.00	0.82	0.25	0.66	0.01	0.33
Mat-WB	0.80	0.25	0.60	0.00	0.40	0.80	0.24	0.60	0.00	0.40	0.12	0.24	0.03	0.80	0.18
Mat-FB	0.91	0.23	0.85	0.03	0.13	0.91	0.20	0.82	0.01	0.18	0.62	0.39	0.45	0.21	0.34
Mat-MB	1.00	0.02	1.00	0.00	0.00	1.00	0.00	1.00	0.00	0.00	0.81	0.28	0.66	0.04	0.30
WB-FB	0.50	0.20	0.08	0.08	0.84	0.51	0.09	0.03	0.01	0.97	0.93	0.18	0.87	0.00	0.13
WB-BM	0.91	0.19	0.83	0.00	0.17	0.74	0.25	0.48	0.00	0.52	0.98	0.10	0.96	0.00	0.04
FB-MB	0.99	0.07	0.98	0.00	0.02	0.99	0.08	0.97	0.00	0.03	0.73	0.29	0.51	0.04	0.44

D<sub>ij</sub> = relation between pairs of predictors according to the dominance matrices in three values (0, 0.5 and 1), p<sub>ij</sub> = proportion of bootstrap samples where i dominates j, p<sub>ji</sub> = proportion of bootstrap samples where j dominates i, p<sub>noij</sub> = proportion of samples where no dominance can be asserted. Abbreviations: Ger = German grade, Mat = Math grade, WB = Work behavior, FB = Family background, MB = Migration background.

**Table A9.** Results from the dominance analysis for the perceived qualification (complete dominance).

I-j	All Cases (N = 691, n <sub>GY</sub> = 355)					Certain Cases (N = 362, n <sub>GY</sub> = 181)					Uncertain Cases (N = 329, n <sub>GY</sub> = 174)				
	D <sub>ij</sub>	SE <sub>Dij</sub>	p <sub>ij</sub>	p <sub>ji</sub>	p <sub>noij</sub>	D <sub>ij</sub>	SE <sub>Dij</sub>	p <sub>ij</sub>	p <sub>ji</sub>	p <sub>noij</sub>	D <sub>ij</sub>	SE <sub>Dij</sub>	p <sub>ij</sub>	p <sub>ji</sub>	p <sub>noij</sub>
Ger-Mat	0.77	0.25	0.55	0.00	0.46	0.78	0.25	0.57	0.00	0.43	0.55	0.37	0.34	0.23	0.43
Ger-WB	1.00	0.04	0.99	0.00	0.01	1.00	0.04	0.99	0.00	0.01	0.61	0.44	0.52	0.30	0.19
Ger-FB	1.00	0.04	1.00	0.00	0.01	1.00	0.00	1.00	0.00	0.00	0.85	0.29	0.75	0.06	0.19
Ger-MB	1.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	0.00	0.99	0.06	0.98	0.00	0.02
Mat-WB	0.83	0.24	0.66	0.00	0.34	0.76	0.25	0.53	0.01	0.46	0.54	0.32	0.25	0.17	0.58
Mat-FB	1.00	0.05	0.99	0.00	0.01	1.00	0.03	1.00	0.00	0.00	0.79	0.31	0.65	0.07	0.29
Mat-MB	1.00	0.00	1.00	0.00	0.00	1.00	0.00	1.00	0.00	0.00	0.99	0.06	0.98	0.00	0.02
WB-FB	0.63	0.22	0.26	0.00	0.74	0.75	0.25	0.49	0.00	0.51	0.74	0.34	0.58	0.10	0.31
WB-BM	0.99	0.08	0.97	0.00	0.03	0.93	0.17	0.86	0.00	0.14	0.98	0.10	0.96	0.00	0.04
FB-MB	1.00	0.04	0.99	0.00	0.01	0.92	0.19	0.84	0.00	0.17	0.95	0.16	0.90	0.00	0.10

D<sub>ij</sub> = relation between pairs of predictors according to the dominance matrices in three values (0, 0.5 and 1), p<sub>ij</sub> = proportion of bootstrap samples where i dominates j, p<sub>ji</sub> = proportion of bootstrap samples where j dominates i, p<sub>noij</sub> = proportion of samples where no dominance can be asserted. Abbreviations: Ger = German grade, Mat = Math grade, WB = Work behavior, FB = Family background, MB = Migration background.

**Table A10.** Pooled random parameters from multilevel logistic regressions of the actual recommendation.

Random Parameter	All Cases (N = 691, n <sub>GY</sub> = 477)		Certain Cases (N = 362, n <sub>GY</sub> = 201)		Uncertain Cases (N = 329, n <sub>GY</sub> = 276)	
	τ	p	τ	p	τ	p
Intercept <sup>a</sup>	0.00	1.00	0.00	1.00	0.00	1.00
German grade <sup>b</sup>	0.19	0.90	n/c		0.00	1.00
Math grade <sup>b</sup>	0.36	0.70	0.00	1.00	0.00	1.00
Work behavior <sup>b</sup>	0.33	0.47	0.00	1.00	0.00	1.00
Family background <sup>b</sup>	0.00	1.00	0.00	1.00	0.00	1.00
Migration background <sup>b</sup>	0.00	1.00	0.00	1.00	0.00	1.00

<sup>a</sup> random intercept only models; <sup>b</sup> random slope only models; n/c = no convergence; p = p-value from the Likelihood Ratio Test where the model w/o random effects served as reference.

**Table A11.** Pooled random parameters from multilevel logistic regressions of the perceived qualification.

Random Parameter	All Cases (N = 691, n <sub>GY</sub> = 355)		Certain Cases (N = 362, n <sub>GY</sub> = 181)		Uncertain Cases (N = 329, n <sub>GY</sub> = 174)	
	$\tau$	$p$	$\tau$	$p$	$\tau$	$p$
Intercept <sup>a</sup>	0.00	1.00	0.00	1.00	0.00	1.00
German grade <sup>b</sup>	0.00	1.00	0.00	1.00	0.00	1.00
Math grade <sup>b</sup>	1.25	0.26	0.00	1.00	0.00	1.00
Work behavior <sup>b</sup>	0.21	0.90	0.00	1.00	0.00	1.00
Family background <sup>b</sup>	1.69	0.03	0.00	1.00	0.00	1.00
Migration background <sup>b</sup>	0.00	1.00	0.00	1.00	0.00	1.00

<sup>a</sup> random intercept only models; <sup>b</sup> random slope only models;  $p = p$ -value from the Likelihood Ratio Test where the model w/o random effects served as reference.

## References

1. Ansalone, G. Tracking: Educational differentiation or defective strategy. *Educ. Res. Q.* **2010**, *34*, 3–17.
2. Wößmann, L. International evidence on school tracking: A review. *CESifo DICE Rep.* **2009**, *7*, 26–34.
3. Maaz, K.; Trautwein, U.; Lüdtke, O.; Baumert, J. Educational transitions and differential learning environments: How explicit between-school tracking contributes to social inequality in educational outcomes. *Child Dev. Perspect* **2008**, *2*, 99–106. [CrossRef]
4. Organisation of Economic Cooperation and Development. *Equity in Education: Breaking down Barriers to Social Mobility*; PISA: Paris, France, 2018. [CrossRef]
5. Organisation of Economic Cooperation and Development. Education Policy Outlook: Germany. Available online: <https://www.oecd.org/education/policy-outlook/country-profile-germany-2020.pdf> (accessed on 9 July 2021).
6. *The Education System in the Federal Republic of Germany 2017/2018: A Description of the Responsibilities, Structures and Developments in Education Policy for the Exchange of Information in Europe*; Secretariat of the Standing Conference of the Ministers of Education and Cultural Affairs of the Länder in the Federal Republic of Germany: Bonn, Germany, 2019.
7. Blossfeld, P. Social background and between-track mobility in the general education system in West Germany and in East Germany after German Unification. *Z. Soziol.* **2018**, *47*, 255–269. [CrossRef]
8. Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. Übergang von der Grundschule in Schulen des Sekundarbereichs I und Förderung, Beobachtung und Orientierung in den Jahrgangsstufen 5 und 6 (sog. Orientierungsstufe) [Transition from Elementary School to Lower Secondary Schools and Support, Observation and Orientation in Grades 5 and 6 (So-Called Orientierungsstufe)]. Available online: [http://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2015/2015\\_02\\_19-Uebergang\\_Grundschule-SI-Orientierungsstufe.pdf](http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2015/2015_02_19-Uebergang_Grundschule-SI-Orientierungsstufe.pdf) (accessed on 9 July 2021).
9. Klinge, D. *Die Elterliche Übergangsentscheidung nach der Grundschule [The Parental Transition Decision after Elementary School]*; Springer: Wiesbaden, Germany, 2016. [CrossRef]
10. Baeriswyl, F.; Wandeler, C.; Trautwein, U.; Oswald, K. Leistungstest, offenheit von Bildungsgängen und obligatorische Beratung der Eltern: Reduziert das Deutschfreiburger Übergangsmodell die Effekte des sozialen Hintergrunds bei Übergangsentscheidungen? [Performance tests, open-access courses and obligatory counseling of parents – Does the Freiburg transition model reduce the effects of social background on transition decisions?]. *Z. Erziehungswiss.* **2006**, *9*, 373–392. [CrossRef]
11. Glock, S.; Krolak-Schwerdt, S.; Klapproth, F.; Böhmer, M. Prädiktoren der Schullaufbahnpflichtung für die Schulzweige des Sekundarbereichs I [Predictors of school track recommendation for lower secondary school tracks.]. *Pädag. Rundsch.* **2013**, *67*, 349–367.
12. McFall, J.P. Rational, normative, descriptive, prescriptive, or choice behavior? The search for integrative metatheory of decision making. *Behav. Dev. Bull.* **2015**, *20*, 45–59. [CrossRef]
13. Harte, J.M.; Koele, P. Modelling and describing human judgement processes: The multiattribute evaluation case. *Think Reason* **2001**, *7*, 29–49. [CrossRef]
14. Pena, J.; Nápoles, G.; Salgueiro, Y. Explicit methods for attribute weighting in multi-attribute decision-making: A review study. *Artif. Intell. Rev.* **2020**, *53*, 3127–3152. [CrossRef]
15. Borchering, K.; Schmeer, S.; Weber, M. Biases in multiattribute weight elicitation. In *Contributions to Decision Making*; Caverni, J.-P., Bar-Hillel, M., Barron, F.H., Jungermann, H., Eds.; Elsevier: Amsterdam, The Netherlands, 1995; pp. 3–28.
16. Pfister, H.-R.; Jungermann, H.; Fischer, K. *Die Psychologie der Entscheidung [The Psychology of Decision]*; Springer: Berlin, Germany, 2017. [CrossRef]
17. Heeler, R.M.; Okechuku, C.; Reid, S. Attribute importance: Contrasting measurements. *J. Mark. Res.* **1979**, *16*, 60–63. [CrossRef]
18. Jaccard, J.; Brinberg, D.; Ackerman, L.J. Assessing attribute importance: A comparison of six methods. *J. Consum. Res.* **1986**, *12*, 463. [CrossRef]
19. van Ittersum, K.; Pennings, J.M.; Wansink, B.; van Trijp, H.C. The validity of attribute-importance measurement: A review. *J. Bus. Res.* **2007**, *60*, 1177–1190. [CrossRef]
20. Myers, J.H.; Alpert, M.I. Determinant buying attitudes: Meaning and measurement. *J. Mark.* **1968**, *32*, 13–20. [CrossRef]



21. Westenberg, M.R.; Koele, P. Multi-attribute evaluation processes: Methodological and conceptual issues. *Acta Psychol.* **1994**, *87*, 65–84. [[CrossRef](#)]
22. Alba, J.W.; Hutchinson, J.W.; Lynch, J.G. Memory and decision making. In *Handbook of Consumer Behavior*; Robertson, T.S., Kassarjian, H.H., Eds.; Prentice Hall: Englewood Cliffs, NJ, USA, 1991; pp. 1–49.
23. Jungermann, H. Entscheiden aus psychologischer Sicht [Decision making from a psychological point of view]. In *Entscheidung und Urteil. [Decision and Judgment]*; Jungermann, H., Lütge, C., Eds.; Vandenhoeck & Ruprecht: Göttingen, Germany, 2009.
24. Cronbach, L.J.; Gleser, G.C. *Psychological Tests and Personnel Decisions*, 2nd ed.; University of Illinois Press: Urbana, IL, USA, 1965.
25. Payne, J.W.; Bettman, J.R.; Johnson, E.J. *The Adaptive Decision Maker*; Cambridge University Press: Cambridge, UK, 1993. [[CrossRef](#)]
26. Böhmer, I.; Gräsel, C.; Krolak-Schwerdt, S.; Hörstermann, T.; Glock, S. Teachers' school tracking decisions. In *Competence Assessment in Education: Research, Models and Instruments*; Leutner, D., Fleischer, J., Grünkorn, J., Klieme, E., Eds.; Springer: Cham, Switzerland, 2017; pp. 131–147.
27. Harte, J.M.; Koele, P. A comparison of different methods for the elicitation of attribute weights: Structural modeling, process tracing, and self-reports. *Organ. Behav. Hum. Decis. Process.* **1995**, *64*, 49–64. [[CrossRef](#)]
28. Krolak-Schwerdt, S.; Hörstermann, T.; Glock, S.; Böhmer, I. Teachers' assessments of students' achievements: The ecological validity of studies using case vignettes. *J. Exp. Educ.* **2018**, *86*, 515–529. [[CrossRef](#)]
29. Arnold, K.-H.; Bos, W.; Richert, P.; Stubbe, T. Schullaufbahnpräferenzen am Ende der vierten Klassenstufe [School track preferences at the end of the fourth grade]. In *IGLU 2006: Lesekompetenzen von Grundschulkindern in Deutschland im Internationalen Vergleich. [IGLU 2006. Reading Competencies of Primary School Children in Germany in International Comparison]*; Bos, W., Hornberg, S., Arnold, K.-H., Faust, G., Fried, L., Lankes, E.-M., Schwippert, K., Valtin, R., Eds.; Waxmann: Münster, Germany, 2007; pp. 271–297.
30. Pohlmann, S. *Der Übergang am Ende der Grundschulzeit [The Transition at the End of Primary School]*; Waxmann: Münster, Germany, 2009.
31. Diebig, K. Die Relevanz der Kriterien für die Übergangsempfehlung: Ein Vergleich Unterschiedlicher Erhebungsmethoden [The Relevance of Criteria of the Tracking Recommendation: A Comparison of Different Methods]. Dissertation, Westfälische Wilhelms-Universität, Münster, Germany, 2016.
32. Böhmer, I.; Gräsel, C.; Hörstermann, T.; Krolak-Schwerdt, S. Die Informationssuche bei der Erstellung der Übergangsempfehlung: Die Rolle von Fallkonsistenz und Expertise [The search for information when preparing the tracking recommendation: The role of case consistency and expertise]. *Unterrichtswissenschaft* **2012**, *40*, 140–155.
33. Böhmer, I.; Hörstermann, T.; Gräsel, C.; Krolak-Schwerdt, S.; Glock, S. Eine Analyse der Informationssuche bei der Erstellung der Übergangsempfehlung: Welcher Urteilsregel folgen Lehrkräfte? [An analysis of the search for information when preparing the tracking recommendation. What decision rule do teachers follow?]. *J. Educ. Res. Online* **2015**, *7*, 59–81.
34. *Dual-Process Theories in Social Psychology*; Chaiken, S.; Trope, Y. (Eds.) Guilford Press: New York, NY, USA, 1999.
35. Hollstein, B. Der Anteil der Lehrer an der Reproduktion sozialer Ungleichheit: Grundschulempfehlungen und soziale Selektion in verschiedenen Berliner Sozialräumen [The teachers' share in the reproduction of social inequality: Tracking recommendations and social selection in different social spaces in Berlin]. In *Die Natur der Gesellschaft: Verhandlungen des 33. Kongresses der Deutschen Gesellschaft für Soziologie in Kassel 2006. [The Nature of Society: Negotiations of the 33rd Congress of the German Society for Sociology in Kassel 2006]*; Rehberg, K.-S., Ed.; Campus: Frankfurt/Main, Germany, 2008; pp. 2605–2613.
36. Nölle, I.; Hörstermann, T.; Krolak-Schwerdt, S.; Gräsel, C. Relevante diagnostische Informationen bei der Übergangsempfehlung—Die Perspektive der Lehrkräfte [Relevant diagnostic information in the transition recommendation—The teachers' perspective]. *Unterrichtswissenschaft* **2009**, *37*, 294–310. [[CrossRef](#)]
37. Ditton, H.; Krüsen, J.; Schauenberg, M. Bildungsungleichheit—Der Beitrag von Familie und Schule [Educational inequality—The contribution of family and school]. *Z. Erziehungswiss.* **2005**, *8*, 285–304. [[CrossRef](#)]
38. Stubbe, T.; Bos, W. Schullaufbahnempfehlungen von Lehrkräften und Schullaufbahnentscheidungen von Eltern am Ende der vierten Jahrgangsstufe [Tracking recommendations by teachers and school career decisions by parents at the end of the fourth grade]. *Empir. Pädag.* **2008**, *22*, 49–63.
39. Klapproth, F.; Glock, S.; Krolak-Schwerdt, S.; Martin, R.; Böhmer, M. Prädiktoren der Sekundarschulempfehlung in Luxemburg: Ergebnisse einer Large-Scale-Untersuchung [Predictors of the tracking recommendation in Luxembourg: Results of a large-scale study]. *Z. Erziehungswiss.* **2013**, *16*, 355–379. [[CrossRef](#)]
40. Caro, D.H.; Lenkeit, J.; Lehmann, R.; Schwippert, K. The role of academic achievement growth in school track recommendations. *Stud. Educ Eval.* **2009**, *35*, 183–192. [[CrossRef](#)]
41. Gröhlich, C.; Guill, K. Wie stabil sind Bezugsgruppeneffekte der Grundschulempfehlung für die Schulformzugehörigkeit in der Sekundarstufe? [How stable are reference group effects of secondary school track recommendations?]. *J. Educ. Res. Online* **2009**, *1*, 154–171.
42. Schneider, T. Die Bedeutung der sozialen Herkunft und des Migrationshintergrundes für Lehrerurteile am Beispiel der Grundschulempfehlung [The importance of social and migrant backgrounds for school recommendations in Germany at the end of primary school]. *Z. Erziehungswiss.* **2011**, *14*, 371–396. [[CrossRef](#)]
43. Ditton, H. Wer geht auf die Hauptschule? Primäre und sekundäre Effekte der sozialen Herkunft beim Übergang nach der Grundschule [Who goes to the German "Hauptschule"? Primary and secondary effects of social origin in the transition to secondary school]. *Z. Erziehungswiss.* **2013**, *16*, 731–749. [[CrossRef](#)]

44. Stahl, N. Schülerwahrnehmung und -beurteilung durch Lehrkräfte [Student perception and assessment by teachers]. In *Kompetenzaufbau und Laufbahnen im Schulsystem: Ergebnisse einer Längsschnittuntersuchung*. [Competence Development and Careers in the School System: Results of a Longitudinal Study]; Ditton, H., Ed.; Waxmann: Münster, Germany, 2007; pp. 171–198.
45. Neumann, M.; Milek, A.; Maaz, K.; Gresch, C. Zum Einfluss der Klassenzusammensetzung auf den Übergang von der Grundschule in die weiterführenden Schulen [The influence of class composition on the transition from primary to secondary school]. In *Der Übergang von der Grundschule in Die Weiterführende Schule: Leistungsgerechtigkeit und Regionale, Soziale und Ethnisch-Kulturelle Disparitäten*. [The Transition from Primary to Secondary School: Achievement Equity and Regional, Social and Ethnic-Cultural Disparities]; Maaz, K., Baumert, J., Gresch, C., McElvany, N., Eds.; BMBF: Bonn, Germany, 2010; pp. 229–251.
46. Schmitt, M. Die Bedeutung von sozialer Herkunft und bundeslandspezifischen Übergangsregelungen für die Grundschulempfehlung [The importance of social origin and federal state specific transitional regulations for the elementary school recommendation]. In *Pädagogische Professionalität als Gegenstand Empirischer Forschung*. [Pedagogical Professionalism as an Object of Empirical Research]; Lankes, E.-M., Ed.; Waxmann: Münster, Germany, 2008; pp. 111–121.
47. Kristen, C. Ethnische Diskriminierung in der Grundschule?: Die Vergabe von Noten und Bildungsempfehlungen [Ethnic discrimination in elementary school? The allocation of grades and tracking recommendations]. *Köln. Z. Soziol. Sozpsychol.* **2006**, *58*, 79–97. [CrossRef]
48. Tiedemann, J.; Billmann-Mahecha, E. Zum Einfluss von Migration und Schulklassenzugehörigkeit auf die Übergangsempfehlung für die Sekundarstufe I [The influence of ethnic criteria and frame of reference effects on teachers' recommendations]. *Z. Erziehungswiss.* **2007**, *10*, 108–120. [CrossRef]
49. Wagner, W.; Helmke, A.; Schrader, F.-W. Die Rekonstruktion der Übergangsempfehlung für die Sekundarstufe I und der Wahl des Bildungsgangs auf der Basis des Migrationsstatus, der sozialen Herkunft, der Schulleistung und schulklassenspezifischer Merkmale [Reconstruction of recommendations and choices at the transition to secondary education: Ethnic criteria, social background, achievement, and class composition]. In *Bildungsentscheidungen: Zeitschrift für Erziehungswissenschaft Sonderheft 12 | 2009*; [Educational decisions]; Baumert, J., Maaz, K., Trautwein, U., Eds.; VS: Wiesbaden, Germany, 2010; pp. 183–204. [CrossRef]
50. Schulze, A.; Wolter, F.; Unger, R. Bildungschancen von Grundschulern: Die Bedeutung des Klassen- und Schulkontextes am Übergang auf die Sekundarstufe I [Elementary school pupils' educational opportunities: The contribution of class and school contexts at the transition to secondary schools]. *Köln. Z. Soziol. Sozpsychol.* **2009**, *61*, 411–435. [CrossRef]
51. Kleine, L.; Birnbaum, N.; Zielonka, M.; Doll, J.; Blossfeld, H.-P. Auswirkungen institutioneller Rahmenbedingungen auf das Bildungsstreben der Eltern und die Bedeutung der Lehrerempfehlung [The effect of varying institutional settings on parental aspirations and teachers' recommendations]. *J. Educ. Res. Online* **2010**, *2*, 72–93.
52. Riek, K.; van Ophuysen, S. Nicht immer zählt nur Leistung: Schulformabhängige Prädiktoren der Übergangsempfehlung [Performance is not always the only thing that counts—School track dependent predictors of the transition recommendation]. In *Facetten Grundschulpädagogischer und -Didaktischer Forschung* [Facets of Primary School Pedagogical and Didactical Research]; Liebers, K., Landwehr, B., Reinhold, S., Riegler, S., Schmidt, R., Eds.; VS Verlag für Sozialwissenschaften: Wiesbaden, Germany, 2016; pp. 13–18. [CrossRef]
53. van Buuren, S.; Groothuis-Oudshoorn, K. Mice: Multivariate imputation by chained equations in R. *J. Stat. Softw.* **2011**, *45*, 1–67. [CrossRef]
54. Schneider, H. Nachweis und Behandlung von Multikollinearität [Detection and treatment of collinearity]. In *Methodik der Empirischen Forschung*. [Methodology of Empirical Research]; Albers, S., Klapper, D., Konradt, U., Walter, A., Wolf, J., Eds.; Gabler: Wiesbaden, Germany, 2007; pp. 183–198. [CrossRef]
55. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2021; Available online: <https://www.R-project.org/> (accessed on 17 September 2021).
56. Bates, D.; Mächler, M.; Bolker, B.; Walker, S. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **2015**, *67*, 1–48. [CrossRef]
57. Grund, S.; Robitzsch, A.; Lüdtke, O. mitml: Tools for Multiple Imputation in Multilevel Modeling. 2021. Available online: <https://CRAN.R-project.org/package=mitml> (accessed on 17 September 2021).
58. Kohl, M. MKmisc: Miscellaneous Functions from M. Kohl. 2019. Available online: <http://www.stamats.de> (accessed on 17 September 2021).
59. Azen, R.; Traxel, N. Using dominance analysis to determine predictor importance in logistic regression. *J. Educ. Behav. Stat.* **2009**, *34*, 319–347. [CrossRef]
60. Bustos Navarrete, C.; Coutinho Soares, F. Dominanceanalysis: Dominance Analysis. 2020. Available online: <https://CRAN.R-project.org/package=dominanceanalysis> (accessed on 17 September 2021).
61. Gustafsson, A.; Johnson, M.D. Determining attribute importance in a service satisfaction model. *J. Serv. Res.* **2004**, *7*, 124–141. [CrossRef]
62. Tversky, A. Elimination by aspects: A theory of choice. *Psychol. Rev.* **1972**, *79*, 281–299. [CrossRef]
63. Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland. Standards für die Lehrerbildung: Bildungswissenschaften: Beschluss der Kultusministerkonferenz vom 16.12.2004 i. d. F. vom 16.05.2019 [Standards for teacher education: Educational Sciences. Resolution of the Conference of Ministers of Education and Cultural Affairs of 16.12.2004 in the Version of 16.05.2019]. Available online: [https://www.kmk.org/fileadmin/veroeffentlichungen\\_beschluesse/2004/2004\\_12\\_16-Standards-Lehrerbildung-Bildungswissenschaften.pdf](https://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2004/2004_12_16-Standards-Lehrerbildung-Bildungswissenschaften.pdf) (accessed on 27 August 2021).
64. *In Two Minds: Dual Processes and Beyond*; Evans, J.S.B.T.; Frankish, K. (Eds.) Oxford University Press: Oxford, UK, 2009. [CrossRef]

- 
65. Krolak-Schwerdt, S.; Böhmer, M.; Gräsel, C. Leistungsbeurteilungen von Schulkindern: Welche Rolle spielen Ziele und Expertise der Lehrkraft? [Performance assessments of school children. What is the role of the teacher's goals and expertise?]. *Z. Entwicklungspsychol. Pädagog. Psychol.* **2012**, *44*, 111–122. [[CrossRef](#)]
  66. Shafir, E.; LeBoeuf, R.A. Context and conflict in multiattribute choice. In *Blackwell Handbook of Judgment and Decision Making*; Koehler, D.J., Harvey, N., Eds.; Blackwell Publishing: Malden, MA, USA, 2004; pp. 341–359.
  67. Timmermans, A.C.; de Boer, H.; Amsing, H.T.A.; van der Werf, M.P.C. Track recommendation bias: Gender, migration background and SES bias over a 20-year period in the Dutch context. *Br. Educ. Res. J.* **2018**, *44*, 847–874. [[CrossRef](#)]