Computers and Education Open 6 (2024) 100181

Contents lists available at ScienceDirect

Computers and Education Open

journal homepage: www.sciencedirect.com/journal/computers-and-education-open

Experiment on extraversion distribution in groups through a group formation algorithm

Adrienne Mueller [a,*], Anna Goeddeke [b], Petra Kneip [b], Johannes Konert [c], René Röpke [d], Henrik Bellhäuser [a]

[a] Johannes Gutenberg University Mainz, Psychology in the Educational Sciences, Binger Str. 14-16, 55122 Mainz, Germany
[b] ESB Business School, Hochschule Reutlingen, Alteburgstr. 150, 72762 Reutlingen, Germany
[c] Fulda University of Applied Sciences, Leipziger Str. 123, 36037 Fulda, Germany
[d] RWTH Aachen University, Computer Science 9 (Learning Technologies), Ahornstraße 55, 52074 Aachen, Germany

ARTICLE INFO

Keywords:
Educational technology
Higher education
Group formation algorithm, Media in education

ABSTRACT

Advances in technology have sparked a surge of interest in systematic group formation in educational contexts. The experimental study investigates group formation by extraversion distributions on group work outcomes, expected to influence group hierarchy. As an initial step in the experimental randomization process, an algorithmic group formation tool ensured an equal partitioning and aligned students into two experimental conditions with either consistent, homogeneous, or varied, heterogeneous, levels of extraversion. Over the course of one semester, a total of 114 students enrolled in several paralleled seminars, were surveyed on both subjective data (satisfaction with group work) and objective data (group performance) to evaluate the effect of the experimental intervention. The formation of extraversion at the group level contributed to the respective outcomes, emphasizing the value of collective social capital for both individuals and groups. Specifically, a homogeneous distribution of extraversion had a positive impact on group performance, as evident in improved grades on course group assignments and increased active participation in group meetings. Findings emphasize considering personality traits at group-level to enhance the success of groups.

1. Introduction

Working in groups is a widely recognized technique for promoting learning progress through social interaction [1,2]. It is commonly used in various educational and professional settings (National Education Association, 2014). However, despite the benefits of group work [3], studies have also highlighted potential downsides [4,5]. The underlying group dynamic processes and factors that lead to either success or failure in group work are not fully understood [6]. Therefore, research has focused on developing effective strategies for forming, shaping, and actively managing groups to ensure positive outcomes [7,8].

Optimizing group formation is essential for promoting effective group work in higher education, which is critical for improving students' learning experiences and outcomes, but still inconclusive in research results [9]. In previous research, experiments were conducted on group formation, primarily within the context of online group work [10,11]. In these studies, additional criteria were chosen for experimental group formation, and their effects on group work outcomes were investigated. However, it's important to note that these experiments were conducted in online settings, whereas the current study is centered on face-to-face interactions. Notably, the present study features successful longitudinal data collection with minimal dropout, setting it apart from other group work research, including some of our own work [12]. The unique experimental study design represents a distinctive feature, yielding valuable insights into the dynamics of group formation.

Building on this foundation, our research aims to delve specifically into the implications of extraversion distribution within groups. As noted by Mohammed and Angell [13], extraverted individuals, with their assertiveness and confidence, tend to dominate discussions this and inadvertently establish hierarchical structures. This phenomenon has been further elucidated by studies conducted by Taggar et al. [14] and Wilmot et al. [15], indicating that a heterogeneous extraversion distribution fosters leadership by those with higher extraversion, ultimately leading to the formation of a group hierarchy. Consequently, the present study endeavors to expand upon these observations by conducting a field experiment on of how heterogeneity of extraversion influences the

* Corresponding author at: Binger Str. 14-16, 55122 Mainz, Germany.
E-mail address: muadrien@uni-mainz.de (A. Mueller).

https://doi.org/10.1016/j.caeo.2024.100181
Received 17 March 2023; Received in revised form 29 February 2024; Accepted 9 April 2024
Available online 12 April 2024

dynamics and functioning of groups in face-to-face interactions.

1.1. From random to strategic: the didactic of group formation

To form groups effectively, teachers require adequate support to manage logistics and ensure timely execution. A systematic approach to group formation requires prior knowledge of the characteristics of the individuals to be assigned to groups, as well as the techniques to perform group formation, including technical support required [16]. However, this is not trivial, as the application of group formation techniques in collaborative learning contexts is a complex process and combinatorial challenge [17] that is influenced by various factors.

With the advent of technology-supported learning environments, algorithmic tools have gained popularity for their potential to facilitate objective and criteria-driven group formation [10,11]. The use of algorithmic group formation tools allows for experimental research to investigate relevant group formation criteria and their constellation, driven by the goal of providing the best possible group learning experience for each learner [16,18,19]. Additionally, algorithms can efficiently handle large datasets and intricate combinations of criteria, overcoming the limitations of manual allocation methods [20].

Current algorithmic approaches, such as semantic matchmakers and non-linear optimization techniques, mostly lack comparability in methods and equitable distribution of criteria, thus hindering fairness and acceptance among learners [16,21,22,23]. Likewise, studies relying on non-experimental designs and other methodological flaws [130,131], leaving the interpretative value of results to be questioned. A comparison of existing algorithmic solutions reveals no computational validation of algorithmic performance (24) or input criteria misalignment with requirements derived from educational psychology [132]. In instance, with studies applying unvalidated criteria such as psychometric constructs or learning styles, as well as poor consideration to individual differences [25,26,133], thereby limiting empirical validation of how individual differences effect group formation outcomes. To adequately address the complexity of group formation and fill the gaps in research, it is crucial to take an interdisciplinary and nuanced approach to develop robust and insightful methods for algorithmic group formation.

In sum, it's important to acknowledge that the effectiveness of algorithms depends on the quality and relevance of the criteria utilized for grouping, even though they may not account for nuanced contextual factors teachers consider [27], it still offers a fair and economical approach to group formation required in higher education.

Notwithstanding the above limitations, it is important to acknowledge certain findings that warrant attention, including research showing that algorithmic grouping according to personality traits outperforms student-selected groups, highlighting the importance of individual differences in group formation [28]. Again, supported by a meta-analysis conducted by [134].

Beyond this, genetic algorithms proved capable of creating more effective groups able to collaborate and improve performance compared to traditional approaches [28]. The positive contribution of systematic, algorithmic group formation [9,21,29,30], calls for research to explore the criteria for group formation in order to develop evidence-based best practices to group formation [31]. Against this background, individual differences in personality traits represent a valid criterion worth considering.

1.2. Interplay of individual differences and group-level resources in group formation

Building on the concept of individual and group-level factors, it's essential to consider various individual differences during the group formation process. These differences encompass demographics, personality traits, attitudes, and cognitive preconditions of group members [10,11,28,32,33]. The choice between homogeneous or heterogeneous distribution of these characteristics within learning groups can depend on the specific criteria used [34,35]. Furthermore, it is essential to test and distinguish the independent effects of group formation and individual traits. Model fit analysis reveals whether group- or individual-level factors play a more significant role in explaining group differences in outcomes within a specific context, assessing the relative impact of both group and individual-level factors [36].

Importantly, the configuration of trait expressions of specific characteristics among group members is a crucial factor in determining outcomes. It often holds more influence than the isolated trait expressions of individual members. This configuration implicitly triggers group dynamics, such as the formation of hierarchies, which, in turn, impact outcomes over time [37]. This interplay raises the question of whether group formation exerts an independent influence on outcomes, distinct from individual characteristics.

In the broader context, existing literature highlights the significance of distinguishing between group- and individual-level factors to gain a comprehensive understanding of group behavior [38]. Previous studies consistently demonstrate the discernible effects of group formation at the group level, influencing decision-making, problem-solving, and creativity [39–41]. Leadership dynamics and communication patterns, inherently linked to the constellation of individuals' personality traits, significantly impact group performance [42,43]. In essence, understanding the factors responsible for the variability in group work outcomes is paramount in the field of group formation research. It becomes increasingly evident that the outcomes of groups are significantly shaped by the dynamic interplay between individual and group-level factors over time [37], leading us to question whether it is reasonable to exclusively attribute group outcomes to an individual's isolated trait configuration. This emphasizes the necessity of adopting a holistic perspective that acknowledges the joint influence of the contextual environment in which groups operate [8].

1.3. Understanding the big five personality traits

The "Big Five" personality traits, also referred to as the Five-Factor Model, represent a well-established and widely utilized psychological framework [44,45]. This model provides a standardized and valid framework for comprehending and predicting personality [46,47]. It encompasses five fundamental personality dimensions: Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism [48,49]. The Big Five model is employed to depict an individual's personality based on these traits, each characterizing general personality features at varying levels of abstraction. Various tests and questionnaires, such as the NEO Personality Inventory-Revised (NEO-PI-R), the Sixteen Personality Factor Fifth Edition (16PF-5), and the Big Five Questionnaire (BFQ), have been developed to assess an individual's standing on each of the Big Five dimensions [46,47].

The Big Five traits, regarded as a construct-oriented approach, have been extensively studied, both within the context of individuals and groups, with the potential to influence thoughts, emotions, behaviors, and social interactions [13,48–53]. Numerous studies consistently highlight their significant impact on individual behavior and group outcomes [54]. Furthermore, the examination of specific trait constellations in group formation can have a positive influence on group hierarchy, power dynamics, and overall group functionality [30,55,56]. Nonetheless, comprehending the relationship between the personalities of individual group members and group outcomes remains an ongoing challenge [9,57]. While the Big Five model provides a standardized and effective framework for describing personality, achieving consensus on how to employ these traits as criteria for group formation and how their distribution among group members impacts overall group performance and social behavior continues to be a subject of ongoing discussion [50,58,59].

1.3.1. Group formation considering the expression of extraversion

Extraversion is characterized by sociability, assertiveness, and a tendency to seek out and enjoy social interactions. This trait indicates

how individuals may exhibit friendliness, approachability, talkativeness, and activity [60–62]. Individuals with high scores in extraversion tend to be sociable, enjoy interactions with others, possess an easy-going nature, and, moreover, often demonstrate leadership behaviors, maintain a positive attitude toward group interactions, and prefer engaging in social activities. Conversely, individuals with low scores in extraversion may exhibit more reserved behaviors and may lean towards solitary work [48]. Due to its interpersonal expression and direct association with social behavior, extraversion plays a pivotal role in understanding an individual's interactions within groups [60–62] and is related to the quality of group interaction [15]. Further, its association with leadership behavior and improved group communication, making it an interesting trait to consider when forming groups [34,62].

Literature on group formation suggests that the level of extraversion among group members can have a significant impact on group functioning [15,135], particularly social interactions and group role formation, such as the formation of hierarchical or non-hierarchical group structures [14,15]. In a group, people with high extraversion initiate discussions [62,13] and provide support to other group members [136], and because of that, they often take over the leadership role, while less extraverted individuals are more likely to follow their lead [63,14]. Thus, one can infer that a heterogeneous distribution in extraversion implicitly establishes a group hierarchy [64].

However, the influence of extraverted leadership is contingent on the context. While extraverted leadership can enhance group performance in situations where employees are passive, it may have a reverse effect when employees are proactive [65], with extraverts contributing more original solutions and comments than introverts [66]. This phenomenon aligns with the dominance complementarity theory, suggesting that interactions thrive when dominance and assertiveness are balanced by compliance and submissiveness [67]. In line with this, extraversion's ability to enhance social experiences might explain extraverts' greater subjective well–being [68] and higher leisure satisfaction and happiness [69]. This underscores the need for a nuanced understanding of how varying degrees of extraversion influence group dynamics and functioning in face-to-face group interactions. The literature provides mixed findings regarding the benefits of heterogeneous versus homogeneous distributions of extraversion for group functioning. Some studies have found that heterogeneity in extraversion is beneficial for task delegation and that heterogeneity in dominance, related to extraversion, enhances the management of relationship conflicts [70,71]. Conversely, others suggest that a homogeneous distribution of extraversion may lead to higher team innovation [72].

Despite this mixed evidence, there is a general assumption that a heterogeneous distribution of extraversion is superior to a homogeneous distribution for group functioning [50,73,74] as it can promote cooperative communication, clarify roles and responsibilities, and facilitate coordination and decision-making [73,75]. However, research shows inconclusive results on several group formation outcomes, and research concerning specific, relevant outcomes influenced by distributions of extraversion has not been thoroughly studied, leaving questions about the effectiveness of group formation strategies [16]. Our research aims to fill this gap by experimentally investigating group formation by the distribution of extraversion, hoping to generate more knowledge about its use for systematic group formation.

1.4. Research goals and justification for group formation by extraversion

The use of personality as a criterion for group formation has received relatively little attention in research, underscoring the necessity for further investigation [9]. The existing research on group formation reveals notable gaps, often characterized by a reliance on correlational designs that cannot establish causal relationships or the use of invalid criteria for group formation [76,77]. Additionally, there is limited research on the impact of various group formation criteria on group outcomes [78–81]. These limitations present significant challenges in ascertaining the most effective methods for selecting and weighing group formation criteria. Moreover, efforts to address the group formation problem in diverse contexts have yielded incomplete solutions, leaving significant voids in the existing body of literature [16]. Furthermore, comprehensive evaluations of current group formation strategies are conspicuously absent, and the assessments of implemented algorithms exhibit notable variations. In certain cases, the algorithmic tools employed for group formation lack clear justification [21]. Consequently, our understanding of this subject remains fragmented and inconclusive.

Previous research has predominantly focused on performance as the primary indicator of group success or failure [82,83]. However, for a more comprehensive understanding of successful group work, especially in higher education settings, it is crucial to consider additional outcome measures. Given the evidence suggesting that emotions can exert a notable impact on performance, the inclusion of measures assessing group emotions appears to be a valuable addition for a more profound understanding of group success [84]. Consequently, it becomes essential to explore other relevant outcome measures that can serve as indicators of successful group work, with a specific emphasis on factors such as group members' satisfaction or willingness to continue group work [52]. This broader approach leads to a more holistic and thorough understanding of group success.

Numerous studies have emphasized the crucial role of individual differences in determining group outcomes. Barrick et al. [60] highlighted the impact of member personality, including extraversion, on team processes and effectiveness. French & Kottke [85] explored the influence of teamwork interest and group extraversion on satisfaction. Lau & Jin [86] emphasized the significance of group personality composition on students' group work performance. Moreover, personality is an essential factor in group functioning and member satisfaction [137]. Regardless of how a person's personality trait level or distribution interacts with those of other group members, it inevitably affects the group's work process and outcome through the contextual setting in which the group operates [57,87]. Driven by group social capital theory [88,89], we predict that the associations between the distribution of extraversion and group functioning can be conceptualized as group-level resources related to outcomes. The outcome of group formation by the distribution of extraversion should be explained by the structure of the groups and not by the composition of the individuals within these groups [90]. This aligns with previous research suggesting that diverse group members can positively affect group outcomes, as it emphasizes the importance of considering various group-level resources in the context of group formation [138]. The intricate relationship between individual characteristics and group-level dynamics further underscores the need to explore the effects of extraversion distribution on group work outcomes, with a focus on the group as a whole. Therefore, we expect more variance to be explained at the group-level than at the individual-level for all outcome measures and formulate the first hypothesis as follows:

H1: In the models that include experimental manipulation of extraversion, the group-level will explain more variance in outcomes, namely satisfaction with group work (H1a), member involvement (H1b), and performance (H1c), compared to the individual level (e.g., lower model fit indicators).

This expectation is rooted in the dynamic interplay between individual differences and group-level resources during the group formation process. When individuals are brought together, their collective dynamics, interactions, and shared resources play a crucial role in shaping the outcomes. To explore this interplay, we have conducted an experiment where we deliberately manipulated the distribution of extraversion. This unique approach allows us to observe how the collective dynamics at the group level contribute to the amount of variance explained for the outcomes. Analyzing model fit aids in comprehending how group and individual factors influence outcomes in a specific setting. The distribution of traits among group members often has a

greater impact than individual traits alone, leading to group dynamics that impact outcomes over time.

Research suggests that while heterogeneous group characteristics complement each other [91–93], a homogeneous distribution leads to increased comfort and motivation to work together [94,95]. Here, factors such as social homophily [96,97] and the similarity-attraction paradigm [98] influence group similarity preferences, even when similarities may not significantly affect group outcomes [139].

Our study aims to investigate the impact of group formation on various outcome variables by manipulating the distribution of extraversion within groups, categorizing them as either homogeneous or heterogeneous. To offer a more comprehensive understanding, we experimentally test the hypothesis that a heterogeneous distribution of extraversion within a group is more advantageous than a homogeneous one, building upon inconsistent results reported in past research [10,11,72]. Within heterogeneous groups, individuals with higher extraversion levels, characterized by assertiveness and confidence, inadvertently establishing hierarchical structures and fostering leadership emergence. Still as other research suggests the superiority of groups with heterogeneously distributed extraversion [70,85,99]. By experimentally testing this hypothesis, we aim to contribute to the existing body of knowledge, utilizing a comprehensive range of outcome measures, encompassing both subjective and objective indicators in outcome measures satisfaction, participation, and performance. Consequently, we formulate the second hypothesis as follows:

H2: Groups with a heterogeneous distribution of extraversion will report greater satisfaction with their group work (H2a), show a higher degree of group member participation (H2b), and achieve better results (H2c) than groups with a homogeneous distribution of extraversion.

2. Method

2.1. Sampling

We recruited participants from two undergraduate classes at two public universities in Germany: Microeconomics at University Reutlingen and Educational Studies at University Mainz. Participation in the courses and groupwork was mandatory. We obtained written consent from all participants and matched international business students (N = 65) from University Reutlingen and teacher education students (N = 58) from University Mainz into groups of three (N(groups) = 38). We excluded a total of 29 students from our initial sample, who had previously attended the seminar and might biases the outcomes due to prior course knowledge. Among these, 16 students belonged to University Reutlingen, while 8 students were from University Mainz. Additionally, students responding carelessly in the questionnaire for experimental group formation were additional excluded (N = 5). Those students were excluded in random groups and, consequently, not included in the final sample for analysis. Table 1 displays sample and respective group conditions.

In Condition 1, we established 20 groups with a heterogeneous distribution of extraversion, and in Condition 2, we established 18 groups with a homogeneous distribution of extraversion. We ensured equal motivation and prior knowledge levels across all groups by employing a rigorous randomization and group formation procedure. This meticulous approach ensured that motivation and prior knowledge levels remained constant across all groups. After the algorithm had assigned individuals to groups, the students worked on problem sets in face-to-face-groups throughout the term, completing assignments and evaluations on the quality of their groupwork. Three assignments received grades, and every student completed three evaluations.

The study at University Reutlingen was conducted in an introductory Microeconomics-class during the first term of the freshman year. Students were assigned to two groups (A and B) to allow for an engaging and supportive learning environment. Each class had approximately 35 students, who were taught for a total of 180 minutes weekly. To pass the course, students had to sit a final exam (scoring a minimum of 51 out of 100) and the opportunity to collect up to 20 bonus points throughout the term. The teacher also awarded up to five bonus points for class participation and up to 15 for three group work submissions, making group work mandatory for passing the course.

Table 1
Sample Characteristics and Group Conditions by University after Dropout.

	Total N	Group k
Overall	114	38
Homogeneous in Extraversion	60	20
Heterogeneous in Extraversion	54	18
By University		
University Reutlingen	54	18
Homogeneous in Extraversion	30	10
Heterogeneous in Extraversion	24	8
University Mainz	65	20
Homogeneous in Extraversion	30	10
Heterogeneous in Extraversion	30	10
Exclusion Criteria	**University Reutlingen**	**University Mainz**
Prior Knowledge (Previous Attended Seminar)	16	8
Careless Response in Questionnaires	1	4

Note. Number of participants and groups by university and conditions, after the initial survey. Exclusion Criteria Prior Knowledge (Attended Seminar) N = 29, University Reutlingen N = 16 before group formation, University Y N = 8. Exclusion Criteria Careless Response in Questionnaires N = 5.

At University Mainz, participants were educational students in the bachelor's program for the teaching profession. Like University Reutlingen, the course was mandatory for graduation and required participation (allowance for a maximum of two missed dates), preparation of meetings at home (literature, slides, podcasts), active participation in discussions, and group work, including turning in three group assignments. The teachers evaluated the group assignments three times, according to previously defined evaluating criteria, with each assignment stimulating group work.

2.1.1. Exclusion criteria and data elimination procedures

To ensure data quality, we implemented exclusion criteria. We excluded students retaking the course due to their prior knowledge and eliminated participants' data when detecting careless responses or incomplete or missing data, using case deletion. We also excluded data from participants who had not filled in the questionnaire before experimental grouping, resulting in the algorithm placing them in random groups with other participants with missing data. In instances where participants forgot their code name or misspelled it in the post-test, we were unable to match data over time, leading to the exclusion of these participants from the analyses. To further ensure data quality, we scanned questionnaire data for traces of careless responses and eliminated any data that was deemed unreliable. Finally, the number of students excluded using each exclusion criterion was N = 29, for University Reutlingen N = 16 and University Mainz N = 8, additionally N = 5 students were grouped randomly and therefore not included in the final sample.

2.1.2. Justification of sample size

Following suggestions by Lakens [100], researchers needed to consider the resources available to conduct a study. The limiting factor in our study was the availability of students. We recruited as many students as possible, but resources limited our sample size. We also performed an a priori statistical power analysis for sample size estimation, based on data from a previous study (N(groups) = 60) (Glimmpse software[1]). The effect size in this study was for outcome variable

[1] Glimmpse 3.0.0 (samplesizeshop.org) accessed on June 26, 2022.

performanced = 0.03, *participationd* = 0.07, and *satisfactiond* = 0.04, considered extremely small using Cohen's (1988) criteria [12]. With α = 0.05 and power *d* = 0.90, the projected required sample size approximately reflects outcome measures: *performance N*(groups) = 24, *participation N*(groups) = 36, and *satisfaction N*(groups) = 44 to perform a between-group comparison. Thus, our proposed sample size of *N* (groups) = 38 was adequate. It should also allow for expected attrition and our additional objectives of controlling for possible mediating or moderating factors and interpreting results.

2.2. Study design and experimental conditions

The study employed a longitudinal experimental design, with one factor (extraversion) at two levels (heterogeneous and homogeneous) with three evaluation time points. The university where we conducted the study utilized Moodle's online learning management system (LMS). The Moodle[2] platform played a central role as it enabled the implementation of the grouping plugin MoodlePeers developed for this purpose, which made adjustments to ensure equal or unequal distribution of mean values for selected criteria across groups.

Students initially completed the T1 questionnaire directly in the Moodle platform. Afterwards, the plugin randomly divided the entire sample into two comparable halves and ensured a balanced distribution of the chosen criteria by testing whether both halves had equal mean values in the group formation criterion extraversion and in the control variables prior knowledge and motivation. Afterwards, the algorithm performed the group formation in both halves separately from each other.

In one half, the experimental group, the algorithm was programmed to create groups that were heterogeneous in extraversion, while trying to achieve a similar extent of heterogeneity in all groups in this experimental condition. In the other half, the control group, the algorithm created groups being homogeneous in extraversion, again aiming to achieve the same extent of homogeneity in all groups in this experimental condition. To control for the impact of prior knowledge and motivation, in both halves the algorithm tried to create groups that were heterogeneous in the control variables simultaneously with the respective manipulation of the distribution of extraversion. Thereby, all resulting groups in both halves had similar standard deviations in prior knowledge and motivation which necessarily also leads to similar mean values in both control variables. With respect to the experimental condition, all groups in the experimental group had similarly high standard deviations in extraversion, while all groups in the control condition had similarly low standard deviations in extraversion.

Following the group formation process, students engaged in coursework tailored to their respective seminars throughout the semester, culminating in three graded group work submissions, each complemented by evaluation questionnaires (T2-T4). At the end of the semester, students completed the T4 final evaluation questionnaire.

Fig. 1 offers an illustrative representation of the steps involved in the group formation process as well as data collection for the study. The methodology utilized an experimental, longitudinal design, with distinct measurement time points represented as T1 -T4.

2.3. Measurement instruments

2.3.1. Control variables

Algorithmic support. To form groups, we developed a Moodle plugin called MoodlePeers, for optimizing group formation. This plugin provides a user interface for instructors to set up group formation within a course, as well as the administration of questionnaires and an overview of the status of group formation (e.g., not yet started, open for answers, and groups formed). In addition, the plugin includes an implementation of the optimization algorithm. The algorithm was used to determine how to group the participants and apply a different set of matching criteria to each part. By minimizing (or maximizing) the distance between all three group members simultaneously, it maintains the same fitness level (i.e., prior knowledge and motivation) in the matched overall groups [140,141]. The algorithm was specifically designed to optimize the formation of learning groups within Moodle. It emphasizes achieving balanced group composition in terms of relevant criteria chosen. The algorithm endeavors to ensure an equally good fit between all possible pairs within a group, considering both homogeneous and heterogeneous optimization criteria.

In our study, the criteria for homogeneous distribution were student motivation and self-assessed initial prior knowledge level. The algorithm employed the distribution of extraversion, either homogeneously or heterogeneously within groups, as another set of optimization criteria. It is important to clarify that "homogeneous criteria" refer to characteristics that the algorithm aims to make ideally equal within a group, while "heterogeneous criteria" refer to characteristics that the algorithm aims to make ideally different within a group. For further details on the algorithm and relevant factors in learning group formation, please refer to the provided references [22,23].

Personality. We measured extraversion using the German short version of the Big Five Inventory (BFI-K) [101]. The personality questionnaire had robust reliabilities in this setting (extraversion: eight items, α = 0.89, conscientiousness: nine items, α = 0.83, openness: five items, α = 0.70, neuroticism: four items, α = 0.79, agreeableness: four items, α = 0.64). We measured prior knowledge as an average subjective rating ("How do you judge your knowledge about the course content," etc.) ranging from 0 to 100 points.

Motivation. We used four subscales to measure motivation: expectations (four items, e.g.," I know that I can learn the contents of the preliminary course," $\alpha = 0.86$), use (five items, e.g.," I understand how important the preliminary course is for my future," $\alpha = 0.78$), cost (six items, e.g.," The time required for the preliminary course seems great to me," $\alpha = 0.83$), and interest (seven items, e.g.," I'm looking forward to the preliminary course," $\alpha = 0.80$). Reliabilities of the motivation scales were high.

Team Orientation. We measured attitudes toward teamwork (team orientation), using three questions (e.g., "If I have a choice, I'd rather work in a team than alone," $\alpha = 0.86$). Reliabilities of the scale were high. Participants rated all questions online, using a scale from 1 ("does not apply") to 6 ("does completely apply").

2.3.2. Outcome measures

We collected data on dependent variables by administering online evaluation surveys at regular intervals, three times throughout the semester, following the submission of each group assignment. The surveys were conducted at the same intervals in both universities. The dependent variables can be broadly categorized into satisfaction, participation of members, and performance-related variables.

Satisfaction. Within the short evaluation questionnaire, participants self-rated their satisfaction concerning group work based on 4 items (e.g.," *I am satisfied how we work together as a group*"). Participants rated questions on a six-point Likert scale ranging from "1 = *does not apply*" to "6 = *applies*". Satisfaction comprises the average of the items, whereby higher values indicate higher satisfaction. Reliability analysis showed a good internal consistency ($\alpha = 0.93$).

Participation. Participation in group work was assessed through self-report of the attendee number using an online evaluation questionnaire. The questionnaire was distributed to participants immediately after each group work session, which was mandatory for all group members and took place on the same day. Participants were asked to report the number of group members actively participating in solving the week's homework, including themselves ("How many people in your group have actively participated in solving this week's homework?"). Since each group comprised three members, the possible response options ranged from 1 (= only me) to 3 (= everyone). If no member answered, participation was rated to be 0. This method allowed for the individual assessment of participation and collaboration quality,

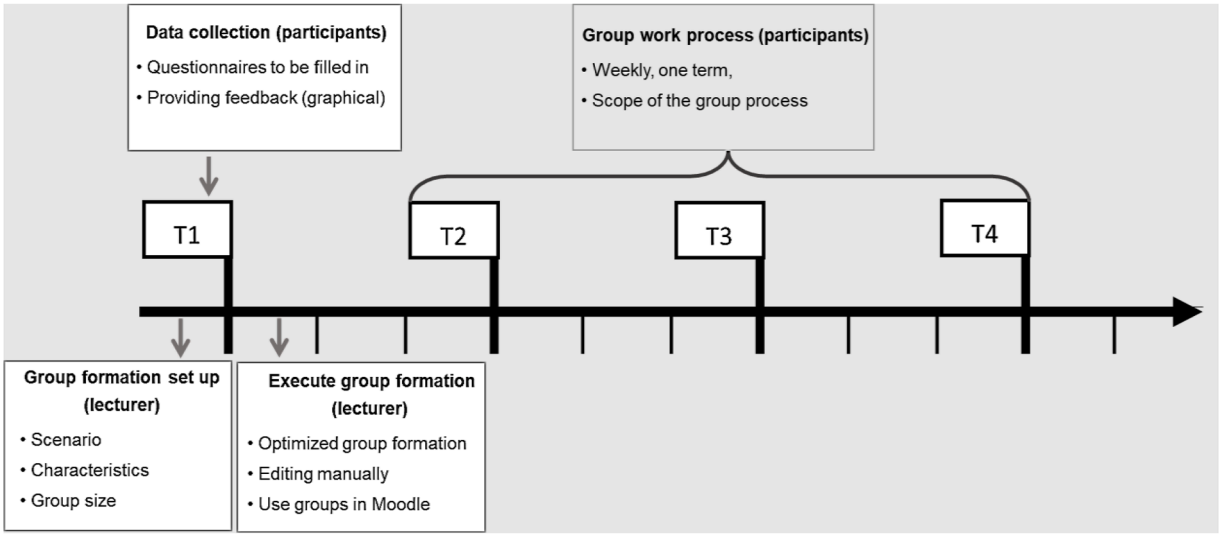

Fig. 1. *Group Formation and Coursework*
Note. Overview of Study Process by T1 = experimental group formation, T2-T4 group work progress thereafter.

capturing the perspectives of all group members.

Performance. Serving as an indicator for the outcome performance, three homework assignments were graded by the tutors. To transfer the values of the performance scores to be merged for the different universities in a consistent manner, percentages were calculated based on the maximum score. The scale ranged from 0 to 100, with higher scores indicating better performance.

2.4. Transparency and openness

Before commencing the study, we informed participants about the general experiment, including the group formation process and obtained written consent prior to participation, to maintain transparency and ethical standards. However, we intentionally withheld specific details such as the criterion (extraversion) for group formation and the study's hypotheses to prevent potential bias in participant behavior and evaluation. To address the sensitive nature of the questions asked in the survey, participants were assured of the confidentiality of their responses, and were informed that their raw data would not be shared with any third parties. Data were analyzed using R, version 4.0.0 [102] and the package nlme (v3.1–152; [103]). We describe our sampling plan, all data exclusions (if any), all manipulations, and all measures in the study.

2.5. Statistical models

The manipulated categorical independent variable was personality trait extraversion (either homogeneous or heterogeneous within each group). We measured the dependent variable based on the study questionnaires, the submitted homework, and the grades obtained for the project. We conducted a multilevel analysis (MLM) due to the research design structure and accounting for the nested data. Thereby, we could test *hypothesis 1*, if group-level or individual-level predictors explained most of the variances. By applying random intercept and random slope models separately and in combination, the model fit indicators AIC and BIC will answer the question of level. Model fit indicators are used to assess the influence of group formation on outcomes over time, as highlighted by Hitt et al. [36]. A model with a lower AIC and BIC provides a reasonable fit [104]. If the group-level could explain the variance, we could test if extraversion manipulation supported *hypothesis 2* for the outcome measures, *satisfaction (H2a), participation (H2b),* and *performance (H2c)*. The random slope model captures how extraversion influences outcome variables differently across groups.

3. Results

3.1. Data preparation and analysis

We ran analyses using SPSS 23.2 and R. Before the analysis, an epsilon test was performed to assess whether the data of the two universities were equivalent and could be combined. The epsilon test is a widely used procedure in multilevel modeling that evaluates the comparability of data from different levels, such as the data from the two universities in this study. This test examines the similarity of the variance-covariance matrices of the different levels, and if the result is not significant, it implies that the data can be pooled and analyzed together. Since the analyses revealed that the data sets from the two universities were comparable and based on that, the necessary conditions were met to proceed with the data analysis of the merged data, which was then carried out. All variables were grand mean-centered to allow a more straightforward interpretation of results. We used the standard $p = 0.05$ criterion to determine if a predictor on any given level explained variance. The post hoc test suggests that the results are significantly different from those expected if the null hypothesis is retained.

3.2. Distribution of individual differences across groups

3.2.1. Results of t-test to compare experimental groups divided by extraversion level

Before conducting the main analyses, we ensured the comparability of the experimental conditions. To achieve this, we performed a *t*-test to compare the experimental groups divided by their level of extraversion in the grand mean values of the initial survey, which included gender, age, team orientation, and personality traits. We found that there were no significant differences between the groups in terms of gender, $t(110) = 0.30$, $p = 0.77$, or age $t(110) = -0.28$, $p = 0.78$. However, we did find a significant difference in team orientation between the scores for the heterogeneous ($M = -0.29$, $SD = 1.49$) and homogeneous ($M = 0.32$, $SD = 1.08$) conditions, with $t(111) = -2.48$, $p = 0.02$. In addition, the *t*-test validated the intended purpose of the implemented algorithm, because the standard deviation of extraversion within groups in the scores for heterogeneous ($M = 0.92$, $SD = 0.58$) and homogeneous ($M = 0.47$, $SD = 0.40$) conditions differ significantly ($t(109) = 4.74$, $p < 0.01$).

3.2.2. Descriptive statistics in outcomes between experimental group conditions

In the present study, we examined the effects of group formation, considering the variance of extraversion, heterogeneously structured, or homogeneously structured in groups, on different outcomes. To provide an initial overview of the data, we first present descriptive analyses of the main outcome measures of *satisfaction, participation*, and *performance* for the two experimental groups, divided by extraversion in the heterogeneously (1) and homogeneously (2) structured group conditions. The results are presented in Table 2.

3.3. Randomizing at the individual-level for improved precision to test for the effects of group formation

To prepare for the testing of hypotheses 1 and 2, we constructed a three-level MLM with time (level 1) nested in individuals (level 2) and individuals nested within groups (level 3). We use the term 'nested' as each student only learns in one group, and each group is doing evaluations and group tasks over time (three events to be nested). MLMs offer information about which of the levels should be used for randomized experimental conditions. In terms of statistical precision and power, it is usually best to randomize at the lowest level possible, in our case level 2 [105]. Therefore, we chose the individual-level (level 2) of a three-level MLM.

MLM analysis also identifies the unexplained variance at each model level. For example, in the case of the influence of other personality traits at the individual level, we can assume that some elements not considered in any given questionnaire represent unexplained variances at any level. Specifically, by not including information about the groups, we may miss important variables at the group level that might explain performance at the individual level. Therefore, we also developed an incorrect model to understand the outcome variable of interest beyond the known problem with underestimating standard errors. In the context of MLMs, including variables at each level is relatively simple, as are interactions among variables at different levels. To sum up, the greater the model complexity, the greater the possibility of understanding the phenomenon of interest [106]. The empty or null level model, initially set up without any explanatory variables, describes the partition (ICC) of variance between the student and group levels. Since it can be assumed that all 3 group members should have the same score in *participation* and in *performance*, there was no variance on the individual level expected. The resulting ICCs for both individual-level and group-level are listed in Table 3. All outcome measures were found to provide more explanation at the group level.

The ICCs do not add up to 100 % because it is a measure of the proportion of total variance that is explained by the group-level or individual-level factors. It represents the percentage of the total variability in the outcome that is due to the grouping variable (e.g., group membership) or individual-level factors (e.g., personality traits, motivation, other individual traits). However, there may still be other sources of variability that are not accounted for by these factors (i.e., residual variance) and contribute to the remaining percentage of variance.

3.3.1. Impact of distribution of extraversion on outcomes

We initially established different models to test them against each other before choosing the best model fit. Following the initial step of building an intercept-only model or null model, we included the variations among participants in the model. Next, we created the models to compare variances across groups alone and individuals within these groups. We tested each model against the null model. Afterward, we added our experimental condition, criterion extraversion—responsible for the structure of extraversion within groups—to the model to determine if it could explain the amount of variance. In doing so, we explicitly included the influence of experimental group-level manipulation (distribution of extraversion within the group, heterogeneous or homogeneous). To assess potential multicollinearity within the dataset, we expanded our analysis by employing additional models. These models, introduced further predictors at the individual level including an array of personality traits, namely extraversion, conscientiousness, neuroticism, agreeableness, openness, and team orientation, as independent variables (Model 3) as well as the inclusion of those predictors, namely personality traits and team orientation, in their squared form (Model 4). We describe the procedure in detail for different dependent variables in the following.

Satisfaction. We specified a model with the dependent variable *satisfaction*. In the model, extraversion did not explain significant variance for the group level, thereby rejecting hypothesis 1a. Furthermore, the model did not show a significant effect of the experimental grouping, thus rejecting hypothesis 2a. In Model 3, none of the additional variables showed statistical significance in this context. Model 4 specifically focused on the quadratic effects of personality traits and team orientation on participants' satisfaction levels, revealing a significant negative curvilinear effect for extraversion^2. This indicates that satisfaction is highest at moderate levels of extraversion. In other words, students with moderate levels of extraversion reported the highest levels of satisfaction, with satisfaction decreasing for students with both lower and higher levels of extraversion. Additionally, there was a positive curvilinear effect for the quadratic term of team orientation on satisfaction. This implies that satisfaction was highest at moderate levels of team orientation. Table 4 displays the model results for the dependent variable satisfaction, including the experimental condition variable of the study.

Participation. For the dependent variable *participation*, we

Table 3
Results of intercept-only model per individual and group-level.

	Individual-Level	Group-Level
Satisfaction	11 %	53 %
Participation	0 %	63 %
Performance	0 %	64 %

Note. The intraclass correlation coefficient (ICC) explained variance at the individual- and group-level.

Table 2
Descriptive measures of main dependent variables as divided by the algorithm.

Dependent Variable	Experimental Group	1		2		3	
		N	M (SD)	N	M (SD)	N	M (SD)
Performance (Assignment)	Heterogeneous	58	42.96 (23.36)	46	56.20 (17.76)	22	70 (18.09)
	Homogeneous	54	44.97 (28.85)	45	67.39 (17.62)	30	69 (10.21)
Satisfaction	Heterogeneous	58	3.59 (1.69)	53	3.33 (1.58)	51	3.34 (1.60)
	Homogeneous	54	3.40 (1.78)	45	3.76 (1.67)	50	3.66 (1.84)
Participation	Heterogeneous	60	2.14 (0.54)	60	2.21 (0.46)	60	1.43 (0.99)
	Homogeneous	54	2.37 (0.49)	54	2.34 (0.55)	54	1.44 (1.22)

Note. N, M, and (SD) represent the number of observations, the mean, and the standard deviation, respectively. 1, 2, 3 = Time (Measurement time points).

established an intercept-only model and compared it to the random intercept model of individuals to determine the best fit. Next, we compared a model of random intercepts of groups to the intercept-only model and found it to be the best fit. The best-fitting model for *participation* was the model for the intercept of individuals per group and, therefore, hypothesis 1b can be accepted, which confirms that the *participation* is explained at the group-level. Considering hypothesis 2b, we added the study's experimental variable in the Random Intercept Model and Random Slope Model to see the variance explained. The distribution of extraversion has a significant effect on the *participation* of group members. But contrary to hypothesis 2, homogeneous groups had higher member *participation* than heterogeneous in extraversion formed groups. Therefore, we must reject hypothesis 2b. Model 3 revealed a significant, positive effect for consciousness. This finding suggests that higher levels of consciousness were positively associated with increased participation in group activities. In contrast, among the additional variables considered, none exhibited statistical significance. In Model 4, there were no observable curvilinear effects to report. This implies that participation was not significantly influenced by the quadratic effects of the included variables. It suggests that the linear effects alone are sufficient to explain the relationship between these traits and participation.

Table 4
Individual-level and group-level predictors of dependent variable satisfaction.

	Dependent variable: Satisfaction			
	(1)	(2)	(3)	(4)
Constant	3.643***	3.644***	3.744***	3.827***
	(0.323)	(0.209)	(0.276)	(0.348)
Experimental Condition	−0.197	−0.190	−0.307	−0.425
	(0.442)	(0.290)	(0.304)	(0.301)
Extraversion			−0.053	−0.199
			(0.216)	(0.222)
Consciousness			0.026	−0.046
			(0.228)	(0.234)
Neuroticism			−0.021	−0.047
			(0.191)	(0.204)
Agreeableness			0.056	0.348
			(0.179)	(0.280)
Openness			−0.001	0.073
			(0.169)	(0.183)
Team orientation			−0.116	0.017
			(0.128)	(0.143)
Extraversion2				−0.232*
				(0.124)
Conscientiousness2				−0.331
				(0.246)
Agreeability2				−0.012
				(0.136)
Neuroticism2				0.155
				(0.118)
Openess2				0.118
				(0.144)
Teamorientation2				0.143**
				(0.072)
Observations	322	322	319	319
Log Likelihood	−497.049	−530.803	−524.224	−518.997
Akaike Inf. Crit.	1006.097	1077.605	1076.448	1077.994
Bayesian Inf. Crit.	1028.744	1107.802	1129.160	1153.298

Note. Treatment Effect: Extraversion homogeneous = 0, heterogeneous = 1. Unstandardized coefficients are reported. Robust standard errors are in parentheses. Missing data handled with case deletion. *p**p***$p < 0.01$.

Detailed results can be found in Table 5.

Performance. For the dependent variable *performance*, we established an intercept-only model and compared it to the random intercept model of individuals to determine the best fit. The random intercepts of groups proved the best fit, thus, hypothesis 1c can be accepted, which confirms that the *performance* is explained by the group-level. We added the study's experimental variable to the models to determine whether it could explain the variance, and it significantly did. Hereby, groups with the homogeneous distribution in extraversion outperformed the heterogeneous distributed ones, thus not confirming Hypothesis 2c, despite the significant effect of extraversion distribution on member *performance*. In Model 3, the results showed that none of the variables had a significant effect on performance. However, in Model 4, a significant negative curvilinear effect for neuroticism on performance was observed. This implies that performance reaches its peak at moderate levels of neuroticism and decreases for students with both lower and higher levels of this trait. Furthermore, agreeableness was found to have a negative effect on performance in a linear manner, but the quadratic effect of agreeableness (squared agreeableness) was not significant, suggesting a more straightforward linear relationship between agreeableness and performance. The results are shown in Table 6.

We could not create a model with an advantageous fit for the other dependent variables, including criterion extraversion. Additionally, it is important to highlight that we calculated the same models for a homogeneous fit of extraversion in groups, showing the same positively

Table 5
Individual-level and group-level predictors of dependent variable participation.

	Dependent variable: Participation			
	(1)	(2)	(3)	(4)
Constant	2.317***	2.675***	2.695***	2.776***
	(0.094)	(0.044)	(0.057)	(0.074)
Experimental Condition	−0.111	−0.160**	−0.147**	−0.155**
	(0.129)	(0.062)	(0.063)	(0.064)
Extraversion			0.008	−0.002
			(0.044)	(0.047)
Consciousness			0.082*	0.066
			(0.047)	(0.050)
Neuroticism			−0.056	−0.057
			(0.040)	(0.044)
Agreeableness			0.009	0.011
			(0.037)	(0.060)
Openness			0.042	0.037
			(0.035)	(0.039)
Team orientation			0.008	−0.005
			(0.026)	(0.031)
Extraversion2				−0.015
				(0.026)
Conscientiousness2				−0.065
				(0.052)
Agreeability2				0.001
				(0.029)
Neuroticism2				0.0005
				(0.025)
Openess2				−0.016
				(0.031)
Teamorientation2				−0.010
				(0.015)
Observations	321	321	318	318
Log Likelihood	−330.613	−304.417	−298.424	−296.559
Akaike Inf. Crit.	673.225	624.835	624.848	633.119
Bayesian Inf. Crit.	695.854	655.006	677.517	708.360

Note. Treatment Effect: Extraversion homogeneous = 0, heterogeneous = 1. Unstandardized coefficients are reported. Robust standard errors are in parentheses. Missing data handled with case deletion. *p**p***$p < 0.01$.

significant values as seen in the tables above.

Additionally, the results are summarized graphically below, with each figure displaying one of the three dependent variables (*satisfaction, participation,* and *performance)* and highlighting the group formation (heterogeneous = red colored, or homogeneous = blue colored in extraversion) over three measurement time points. This enables us to identify which group condition is more beneficial for each outcome. The results are displayed in Fig. 2–4.

Fig. 2 shows boxplots of the satisfaction outcome over three time points divided by the experimental group conditions. The boxplots display minor changes. No uniform results can be reported. However, the mean values for the homogeneous groups are more fluctuating than for the heterogeneous groups.

Fig. 3 boxplots show that the mean values for both groups are similar, with overall high variance. Specifically, the mean values for the heterogeneous groups are higher in time 1, while the homogeneous group has higher mean values in time 2.

Fig. 4 boxplots show higher values for the homogeneous condition as well as more variance, suggesting that the homogeneous group formation may be more effective in promoting performance outcomes over time.

Table 6
Individual-level and group-level predictors of performance.

	Dependent variable: Assignment			
	(1)	(2)	(3)	(4)
Constant	57.051***	62.821***	63.036***	61.727***
	(3.432)	(1.800)	(2.443)	(3.053)
Experimental Condition	−5.096	−7.872***	−8.825***	−7.255***
	(4.794)	(2.630)	(2.688)	(2.657)
Extraversion			−2.088	−0.754
			(1.860)	(1.954)
Consciousness			1.913	2.457
			(2.019)	(2.047)
Neuroticism			1.521	1.107
			(1.753)	(1.876)
Agreeableness			−0.494	−6.126**
			(1.639)	(2.601)
Openness			0.467	0.085
			(1.517)	(1.625)
Team orientation			−1.198	−1.637
			(1.154)	(1.215)
Extraversion2				1.641
				(1.078)
Conscientiousness2				1.384
				(2.053)
Agreeability2				−0.755
				(1.246)
Neuroticism2				−2.992***
				(1.090)
Openess2				0.528
				(1.286)
Teamorientation2				−0.562
				(0.662)
Observations	235	235	233	233
Log Likelihood	−1037.717	−1069.331	−1058.326	−1053.269
Akaike Inf. Crit.	2087.434	2154.661	2144.652	2146.539
Bayesian Inf. Crit.	2108.191	2182.338	2192.967	2215.560

Note. Treatment Effect: Extraversion homogeneous = 0, heterogeneous = 1. Unstandardized coefficients are reported. Robust standard errors are in parentheses. Missing data handled with case deletion. *p**p***$p < 0.01$.

4. Discussion

The purpose of our research was to evaluate the impact of group formation while experimentally manipulating the distribution of extraversion as a group formation criterion. In interpreting the results, it is evident that group-level effects significantly influenced certain outcome variables, supporting the study's theoretical framework. Contrary to the common belief favoring heterogeneity, results revealed a homogeneous distribution of extraversion to significantly contribute to enhance group performance and participation. Although literature associates extraverted leadership with improved group performance in passive settings [65], findings diverge from this, suggesting that a similar distribution of extraversion in groups could be favorable. Given that group homogeneity in extraversion yielded better outcomes, not only challenges traditional notions of the superiority of group hierarchy [107], but also calls into question the dominance complementarity theory, whereby both authority and dominance are balanced by compliance and submissiveness [67]. In the context of extraversion, a homogeneous distribution may be more advantageous, despite the extolled benefits of diversity in perspectives and skills [108]. This may elucidate why a homogeneous distribution of extraversion in groups led to superior performance and participation in our study. Align with findings is the concept of homophily, denoting the tendency for individuals to associate with others similar to themselves, which has been linked to improved group cohesion and performance [95]. Furthermore, research indicates that groups with a shared understanding of goals and a strong sense of collective identity tend to perform better [142], which could be rather apparent in homogeneously extroverted groups.

Interestingly, team orientation, exhibited a significant elevation in homogeneously extraverted groups, indicating an interplay between distributions. This phenomenon further elucidated by French & Kottke [85], who observed that extraversion dispersion (high group variance akin heterogeneous distribution) influenced the relationship between teamwork interest (akin to team orientation) and satisfaction, acting as a suppressor effect. Consequently, the satisfaction level of individuals more inclined towards group work depended on a homogeneous distribution of extraversion. In our study, the heightened team orientation observed in homogeneously extraverted groups may, to some extent, be attributed to shared effects of these characteristics. However, there were no significant effects of groups on outcome satisfaction, again contrary to literature on extraverts' greater subjective well–being [68], higher leisure satisfaction and happiness [69]. This underscores how varying degrees of extraversion within groups, rather than absolute values or levels, may influence various outcomes.

4.1. The premise of analyzing the level of outcomes with extraversion

This discussion section highlights the respective levels adjustment of group formation outcomes, with emphasis on the relationship between the experimental groups, manipulated through group-level extraversion distributions.

The variance in outcome variables postulated to be primarily influenced by group-level effects rather than individual-level effects could be confirmed for the dependent variables of performance and participation. For satisfaction the variability could be better explained at the individual level, as the value of the variable rather than varying between groups, varied according to individual preconditions. We attribute this finding mainly to a ceiling effect, as all participants reported being very satisfied with their groups, leading to high overall satisfaction scores. The use of algorithmic group formation in the study was rather unusual, as in higher education students typically self-select or are allocated to groups. As a result, the study's approach drew attention to the composition of the groups, which may have increased satisfaction among the students.

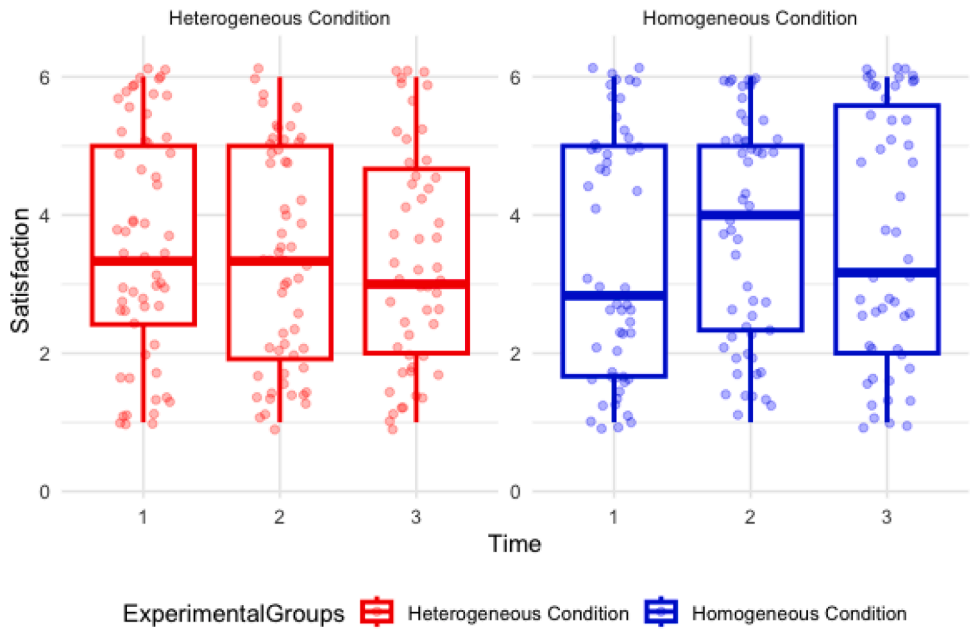

Fig. 2. *Boxplots of Satisfaction over Time by Experimental Group Conditions*
Note. box = interquartile range (IQR), whiskers = min/ max range, middle line = median.

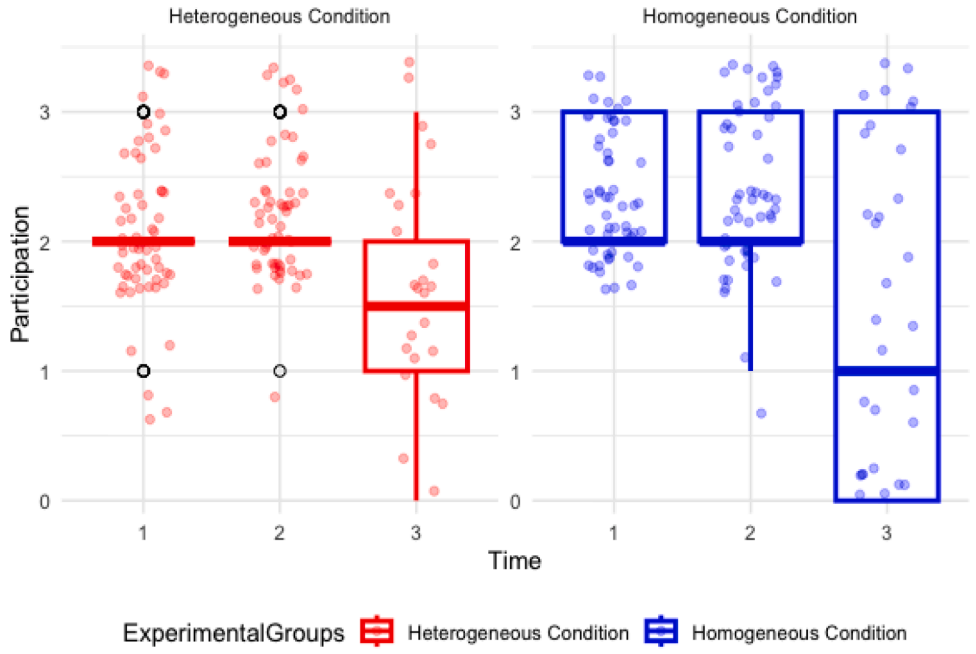

Fig. 3. *Boxplots of Participation over Time by Experimental Group Conditions*
Note. box = interquartile range (IQR), whiskers = min/ max range, middle line = median.

Although participants were unaware of the criteria used to form groups and the underlying hypotheses, we believe that they expected the algorithm to choose the perfect group members for them, resulting in greater satisfaction, even in the control group. Therefore, we assume that algorithmic group formation and participation in the research project led to higher levels of satisfaction among participants, consistent with the research participation effect [109].

To contextualize our findings, the literature on social capital, provides a broader theoretical and transformative perspective for understanding the mechanisms underlying group dynamics [88,89]. By focusing on the role of an important sociostructurally trait like extraversion for group formation, the research could to some extend contribute to existing discussions on micro-level determinants. In summary, research findings contribute to the dialogue on the intricate interplay of individual traits within the context of group work social dynamics.

4.2. Impact of extraversion distribution on group work outcomes

In the following section, we focus on the effect of extraversion distribution on outcome variables. Here, the second hypothesis posited that groups with a heterogeneous distribution of extraversion will report greater satisfaction with group formation and group work, show higher group member participation, and achieve better performance results

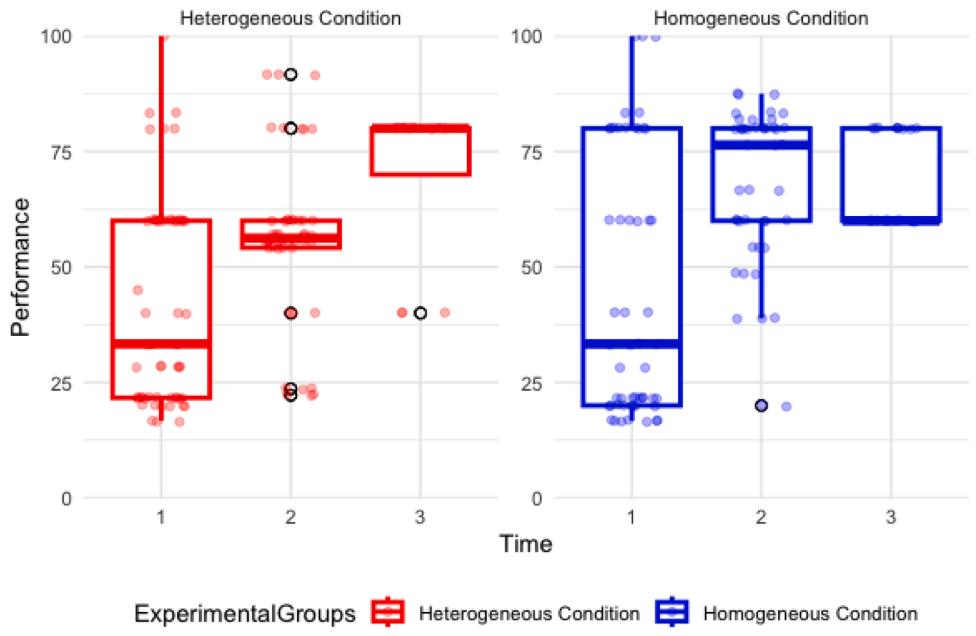

Fig. 4. *Boxplots of Performance over Time by Experimental Group Conditions*
Note. box = interquartile range (IQR), whiskers = min/ max range, middle line = median.

than groups with a homogeneous distribution of extraversion. In contrast to expectations, a homogeneous distribution of the trait extraversion was significantly superior for the outcome variables of performance and participation. Moreover, it is essential to acknowledge the inability to confirm our second hypothesis, which posited the benefits of trait heterogeneity in groups based on existing literature suggesting its advantages mostly for outcome of group work success [74,81,110]. This finding of higher participation resonates with social capital theory [89], emphasizing the value of social cohesion and homophily [111] making members more likely to evolve in group work process. Research in older adults found that social capital provides a viable explanation for the association between extraversion and volunteering, such as participation in organizations, and contact with friends [112]. This further aligns with the recognized dimensions of social capital, such as shared trust, rules, norms, and network resources that could be expected evolving similarity, e.g. homogeneity, as results of the underlying group work dynamics [113].

However, it is important to note that our results challenge previous literature that suggested the benefits of a heterogeneous distribution of extraversion to be beneficial to performance results [70,85,99]. In our sample, team orientation was significantly higher in homogeneously groups. As previously argued, we maintain that team orientation did not affect the nature of either outcome since both relate more to other characteristics, based on theoretical assumptions [85].

4.2.1. Exploring the dynamics of group formation outcomes across time and levels

The following theoretical conceptualization of groups as complex adaptive systems [114], offers a robust framework for interpreting the outcomes of group formation based on extraversion distribution. This approach contributes to a nuanced understanding of how individual differences dynamically interact and evolve within groups, shaping outcomes over time.

Extraversion known as a correlate of leadership [62], by dynamics and communication patterns being linked to the constellation of individuals' personality traits While previously assumed to respective outcome of group performance [42,43], questions about its enduring impact need to be raised based on our contrary research results. As such, extraverted individuals, perceived as leaders, appeared to decline in status within groups over time [115]. The evidence further suggests that the positive social influence of extraversion might diminish over time, with peers initially valuing extraverted behaviors like warmth and assertiveness but subsequently placing importance on other qualities as the socialization process unfolds [62,116]. This evolving dynamic highlights that the expected positive impact of hierarchical group formations may not be evident in long-term group work.

To discern the independent influence of group formation on outcomes over time, distinct from individual characteristics, it is essential to evaluate the impact of group formation outcomes rooted in the distribution of extraversion, considering the structural interplay among individual differences within the group. Social capital provides a viable explanation for the association between extraversion and outcomes of social behavior [112]. Existing literature emphasizes the importance of distinguishing between group- and individual-level factors to gain a comprehensive understanding of the intricacies of group behavior [38]. Moreover, previous studies have consistently shown that group formation can distinctly affect group-level processes, like decision-making, problem-solving, and creativity [39–41].

Moreover, the discussion of group-level variance within the context of social capital aligns with Lin [117], emphasizing the significance of identification and belonging within social groups as central to social manifestation across various dimensions, from individual to societal levels. This perspective aligns with the idea that group dynamics, including the emergence of hierarchies, are implicitly influenced by the configuration of trait expressions and thus affect outcomes over time, which is consistent with our results and other studies [37]. The anticipated effects of group formation resonate with the concept that 'the whole is more than the sum of its parts' [90]. Instead of focusing solely on isolated traits of individual group members, our research reinforces the importance of these trait configurations at the group level, influenced by the collective trait expressions of group members.

4.3. Implications

The findings of our study lead to several theoretical, methodological, and practical implications. In order to determine the factors that contribute to successful group work, it is important to consider the level of groups' similarity or dissimilarity in personality traits. In an

algorithmic approach, experimental groups were created by adjusting the level of extraversion while keeping predictors such as prior knowledge and motivation evenly distributed. Thus, a focused analysis of how the levels of extraversion affect outcomes could be conducted. According to our findings, extraversion appeared to be relevant in the groups process [70], consequently its role should be further investigated within the framework of experimental group design.

Methodologically, our study highlights the fact that effectively enhancing group formation remains a challenge that can only be tackled through interdisciplinary research [72,21]. Conducting randomized field experiments with manipulation of group distributions is only possible by combining the diagnostic expertise from fields like psychology together with the algorithmic expertise of computer scientists. Also, we want to emphasize the methodological decision to analyze curvilinear patterns [118] in our data. Even though in our case, we did not find such non-linear effects, the idea of "too-much-of-a-good-thing" is a statistical method that can be applied to many contexts.

From a practical perspective, algorithmic group formation might be a promising option for many educational settings as soon as feasible grouping criteria will be empirically established through more research. Especially in distance education, when learners have very little information about their potential group members, it is difficult for them to form groups on their own. Also, limited social interactions between group members make group work in distance education particularly difficult. The algorithmic selection of group members holds the promise of providing the necessary skills for this kind of group work in every single group—not only in one group in which the strongest students gathered based on their homophily. Further, implementing algorithmic group formation saves educators time and effort by eliminating the need for manual group formation, allowing them to focus on other critical aspects of teaching.

The results of a homogeneous distribution of extraversion being beneficial for student group performance align with the new concept of agile group work. Agile teams favor a flat, non-hierarchical structure where people are given the autonomy to work independently and organize themselves [119]. In the context of extraversion, a homogeneous distribution means that all members of the group have similar levels in this trait. This can lead to a more balanced and harmonious group dynamic with less conflict or imbalance.

However, ethical considerations such as informed consent, proper authorization for psychometric tests, and privacy must be considered when implementing such strategies. In line with previous findings [120], the results stress active collaborative learning while integrating new technologies to improve students' learning performance. Embedded in an authentic setting, the randomized field experiment contributes to the academic discourse relevant to higher education. By involving students from different majors and universities for an extended period, employing both subjective and objective outcomes, we enhance the relevance and generalizability of our findings, offering a comprehensive understanding of group dynamics while enriching the overall research findings.

4.4. Limitations of the study

Despite its strengths, our study faces several limitations that warrant consideration in interpreting the findings. The relatively small sample size, especially concerning the group-level focus, is a common challenge in empirical group research. The formation of small groups, while increasing the number of groups, introduces heightened variance at the group level, potentially elevating the risk of type I errors.

The reliance on self-report measures for assessing collaborative participation, while common in comparable studies [121,122], presents a limitation. The absence of behavioral data or teacher ratings may limit the objectivity of our outcome measures. While our measurement instruments underwent rigorous testing for reliability and validity, the use of self-reports introduces a subjective element that should be acknowledged. Furthermore, while transparency in communicating the experiment's overview adhered to ethical research practices, it may have influenced participant behavior and introduced potential biases. We provided participants with a general overview of the experiment without specific information on the hypotheses posed to ensure transparency. This approach aimed to maintain natural interactions within the groups and avoid inducing artificial behavior or perceptions. While this transparency may have influenced participants' awareness of the experiment, it aligns with ethical research practices. Still, this may have influenced participant behavior and could thereby introduce limitations to the outcomes of the study. While algorithmic methods offer objectivity and scalability, they are not inherently superior to teacher-based methods and may present their own set of challenges, including privacy concerns and the potential for bias.

Although our study aimed to experimentally manipulate groups based on the distribution of extraversion variance, it is important to acknowledge potential limitations to internal validity. The algorithmic process used may have introduced challenges in ensuring that extraversion was the sole determinant of group membership, as factors beyond extraversion could have inadvertently influenced the grouping process. As personality traits are partly correlated [123], our algorithm's experimental manipulation of the distribution of extraversion in groups may have simultaneously led to a more homogeneous or more heterogeneous distribution of other traits. This makes it difficult to isolate the impact of extraversion from other variables that contribute to group outcomes. Assuming that a homogeneous distribution of extraversion leads to more cohesive collaboration may oversimplify the interplay of individual differences. To further address this concern, it's crucial to consider the explanation and transparency of the algorithm's processing logic. As detailed in the methodology, the algorithm employed both homogeneous and heterogeneous criteria simultaneously, aiming to equalize or differentiate specific characteristics within the groups. which could complicate the understanding of extraversion's impact on group outcomes.

By acknowledging these limitations, future research can address methodological gaps and refine our understanding of group processes. As educational practices evolve, the integration of algorithmic group formation requires careful consideration of its implications for privacy, fairness, and effectiveness in fostering collaborative learning environments. Through continued interdisciplinary research, we can advance our understanding of group dynamics and enhance the quality of collaborative learning experiences in diverse educational contexts.

4.5. Future directions

Future research on group formation based on personality traits should address the identified limitations and strive for improvements. This involves increasing the sample size, controlling for additional variables (e.g., self-regulation, emotional intelligence) that may impact group outcomes, and exploring the impact of personality traits as group formation criteria in a variety of settings. Apart from extraversion, there are several other possible conditioning factors to consider, such as the optimal group size, the type of task to be performed, or the group's goal. While this study focused on experimentally manipulating extraversion, it represents only one facet of the broader picture. Further research should focus on exploring the interaction effects of extraversion with other personality traits. Understanding these interactions could lead to more nuanced algorithmic groupings This line of inquiry has the potential to provide insights into factors that either inhibit or moderate group dynamics [85,124].

To advance our understanding further, it is recommended to include control groups, for example, with randomized group formation, and employ other objective outcome measures or group process data, such as video analysis and logfile data. Utilizing well-established outcome measures, like the observational rubric for assessing collaborative disciplinary engagement in groups [125], can improve the validity of

results and reduce the influence of extraneous factors.

While contributing to the current understanding of group dynamics, this research also points towards a future path for optimizing collaborative learning experiences, particularly in higher education settings. The development of specialized training programs and tools, including apps, can aid individuals in mastering group work and conflict resolution based on their unique characteristics, thereby enhancing the overall group learning experience. Better knowledge of group formation driven by research could enable learners to make better decisions in situations where they are in charge of group formation by themselves [126].

Looking ahead, we recognize the evolving landscape of human-AI collaboration, where AI-based systems become integral group members in various scenarios [127–129]. This transformation necessitates a reevaluation of established theories on group phenomena and processes, considering how human-AI collaboration impacts the way groups work together [64,127]). While there are similarities between human-human and human-machine interactions, many aspects of human-AI collaboration still require further investigation. Besides well-explored topics like trust in AI, reciprocity in human-AI collaboration, and anthropomorphism, we propose a focus on group formation and, specifically, the potential roles of AI-based systems within a group.

4.6. Conclusion

In conclusion, the proposed solution to student group formation using support algorithms and extraversion level as criteria is a useful contribution to collaborative learning and automates the group formation process efficiently and effectively. Although our experiment on the distribution of extraversion within groups did not provide a full understanding of its impact, we found that groups with a more similar distribution of extraversion perform better.

The study utilized a unique approach to determining the efficacy of grouping students by incorporating individual traits, such as extraversion, into the group formation process with the help of an algorithm. However, it is essential to consider ethical considerations when implementing algorithmic approaches, such as informed consent, proper authorization for psychometric testing, and privacy acts. Overall, our findings highlight the importance of considering personality traits in group formation and call for more attention to this aspect in future research.

CRediT authorship contribution statement

Adrienne Mueller: Data curation, Formal analysis, Project administration, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. **Anna Goeddeke:** Conceptualization, Data curation, Supervision. **Petra Kneip:** Supervision. **Johannes Konert:** Resources, Software. **René Röpke:** Resources, Software. **Henrik Bellhäuser:** Methodology, Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Statements and Declarations

The authors did not receive support from any organization for the submitted work.

The authors have no relevant financial or non-financial interests to disclose.

References

[1] Ertmer PA, Newby TJ. Behaviorism, cognitivism, constructivism: Comparing critical features from an instructional design perspective. Performance Improvement Quarterly 2013;26(2):43–71. https://doi.org/10.1002/piq.21143.
[2] Johnson, D.W., Johnson, R.T., & Stanne, M.B. Cooperative learning methods: a meta-analysis. Cooperative Learning Center at the University of Minnesota Press; 2000. Retrieved November 6, 2021, from http://www.tablelearning.com/uploads/File/EXHIBIT_B.pdf.
[3] Kyndt E, Raes E, Lismont B, Timmers F, Cascallar E, Dochy F. A meta-analysis of the effects of face-to-face cooperative learning. Do recent studies falsify or verify earlier findings? Educ Res Rev 2013;10:133–49. https://doi.org/10.1016/j.edurev.2013.02.002.
[4] Chang Y, Brickman P. When group work doesn't work: insights from students. CBE Life Sci Educ 2018;17(3):ar42. https://doi.org/10.1187/cbe.17-09-0199.
[5] Chiriac E. Group work as an incentive for learning - students' experiences of group work. Front Psychol 2014;5:558. https://doi.org/10.3389/fpsyg.2014.00558.
[6] Mathieu JE, Hollenbeck JR, van Knippenberg D, Ilgen DR. A century of work teams in the journal of applied psychology. J Appl Psychol 2017;102(3):452–67. https://doi.org/10.1037/apl0000128.
[7] Graesser AC, Fiore SM, Greiff S, Andrews-Todd J, Foltz PW, Hesse FW. Advancing the science of collaborative problem solving. Psychol Sci Public Int 2018;19(2):59–92. https://doi.org/10.1177/1529100618808244.
[8] LePine JA, Buckman BR, Crawford ER, Methot JR. A review of research on personality in teams: accounting for pathways spanning levels of theory and analysis. Hum Res Manage Rev 2011;21(4):311–30. https://doi.org/10.1016/j.hrmr.2010.10.004.
[9] Borges S, Mizoguchi R, Bittencourt II, Isotani S. Group Formation in CSCL: A Review of the State of the Art. In: Cristea A, Bittencourt I, Lima F, editors. Higher education for all. From challenges to novel technology-enhanced solutions. HEFA 2017. Communications in Computer and Information Science. 832. Springer; 2017. p. 71–88. https://doi.org/10.1007/978-3-319-97934-2_5.
[10] Bellhäuser H, Konert J, Müller A, Röpke R. Who is the perfect match? effects of algorithmic learning group formation using personality traits. Journal of Interactive Media (i-Com) 2018;17(1):65–77. https://doi.org/10.1515/icom-2018-0004.
[11] Müller A, Bellhäuser H, Konert J, Röpke R. Effects of Group Formation on Student Satisfaction and Performance: A Field Experiment. Small Group Research 2021;53(2):244–73. https://doi.org/10.1177/1046496420988592.
[12] Müller A, Röpke R, Konert J, Bellhäuser H. Investigating group formation: An experiment on the distribution of extraversion in educational settings. Acta psychologica 2023;242:104111. https://doi.org/10.1016/j.actpsy.2023.104111.
[13] Mohammed S, Angell LC. Personality heterogeneity in teams: which differences make a difference for team performance? Small Group Res 2003;34(6):651–77. https://doi.org/10.1177/1046496403257228.
[14] Taggar S, Hackett R, Saha S. Leadership emergence in autonomous work teams: antecedents and outcomes. Pers Psychol 2006;52(4):899–926. https://doi.org/10.1111/j.1744-6570.1999.tb00184.x.
[15] Wilmot MP, Wanberg CR, Kammeyer-Mueller JD, Ones DS. Extraversion advantages at work: a quantitative review and synthesis of the meta-analytic evidence. J Appl Psychol 2019;104(12):1447–70. https://doi.org/10.1037/apl0000415.
[16] Maqtary N, Mohsen A, Bechkoum K. Group formation techniques in computer-supported collaborative learning: a systematic literature review. Technol Knowl Learn 2019;24(2):169–90. https://doi.org/10.1007/s10758-017-9332-1.
[17] Hwang G-J, Yin P-Y, Hwang C-W, Tsai C-C. An enhanced genetic approach to composing cooperative learning groups for multiple grouping criteria. Educ Technol Soc 2008;11(1):148–67. https://www.jstor.org/stable/jeductechsoci.11.1.148.
[18] Maina EM, Oboko RO, Waiganjo PW. Using machine learning techniques to support group formation in an online collaborative learning environment. Int J Intell Syst Appl 2017;9(3):26–33. https://doi.org/10.5815/ijisa.2017.03.04.
[19] Sun G, Shen J. Teamwork as a service: a cloud-based system for enhancing teamwork performance in mobile learning. In: 2013 IEEE 13th International Conference on Advanced Learning Technologies (ICALT; 2013. p. 376–8. https://doi.org/10.1109/ICALT.2013.115.
[20] Wilkinson, D.J., & Golightly, A. (2010). Markov chain Monte Carlo algorithms for SDE parameter estimation. *Learn Inf Comput Syst Biol*.
[21] Odo C, Masthoff J, Beacham N. Group formation for collaborative learning. In: Isotani S, Millán E, Ogan A, Hastings P, McLaren B, Luckin R, editors. Artificial Intelligence in Education. AIED 2019 (Vol. 2, pp. 206–12). Springer; 2019. https://doi.org/10.1007/978-3-030-23207-8_39.
[22] Konert J, Burlak D, Steinmetz R. The group formation problem: an algorithmic approach to learning group formation. Lecture Notes in Computer Science. 2014. p. 221–34. https://doi.org/10.1007/978-3-319-11200-8_17.
[23] Konert J, Bellhäuser H, Röpke R, Gallwas E, Zucik A. MoodlePeers: Factors relevant in learning group formation for improved learning outcomes, satisfaction and commitment in ELearning scenarios using GroupAL. In: Verbert K, Sharples M, Klobucar T, editors. Adaptive and Adaptable Learning: Proceedings of the 11th European Conference on Technology Enhanced Learning, EC-TEL 2016, Lecture Notes in Computer Science; 2016. p. 390–6. https://doi.org/10.1007/978-3-319-45153-4_32.
[24] Odo C, Masthoff J, Beacham N. Group formation for collaborative learning. In: Isotani S, Millán E, Ogan A, Hastings P, McLaren B, Luckin R, editors. Artificial

Intelligence in Education. AIED 2019; 2019. p. 206–12. https://doi.org/10.1007/978-3-030-23207-8_39.
[25] Karau SJ, Williams KD. Social loafing: A meta-analytic review and theoretical integration. Journal of Personality and Social Psychology 1993;65(4):681–706. https://doi.org/10.1037/0022-3514.65.4.681.
[26] Moreno J, Ovalle DA, Vicari RM. A genetic algorithm approach for group formation in collaborative learning considering multiple student characteristics. Computers and Education 2012;58(1):560–9. https://doi.org/10.1016/j.compedu.2011.09.011.
[27] Dwork C, Hardt M, Pitassi T, Reingold O, Zemel R. Fairness through awareness. In: Proceedings of the 3rd innovations in theoretical computer science conference; 2012. p. 214–26. https://arxiv.org/abs/1104.3913.
[28] Revelo-Sánchez O, Collazos CA, Redondo MA, Pinto II. Homogeneous group formation in collaborative learning scenarios: an approach based on personality traits and genetic algorithms. IEEE Trans Learn Technol 2021;4(14):486–99. https://doi.org/10.1109/TLT.2021.3105008.
[29] Liang C, Majumdar R, Ogata H. Learning log-based automatic group formation: system design and classroom implementation study. Res Pract Technol Enhanc Learn 2021;(14):16. https://doi.org/10.1186/s41039-021-00156-w.
[30] Mujkanovic A, Bollin A. Empowering learners for life in the digital age (OCCE 2018). In: Passey D, Bottino R, Lewin C, Sanchez E, editors. IFIP Advances in Information and Communication Technology Vol. 524, pp. 207–17. Springer; 2019. https://doi.org/10.1007/978-3-030-23513-0_21.
[31] Patrício L, Franco M. A systematic literature review about team diversity and team performance: future lines of investigation. Adm Sci 2022;12(1):31. https://doi.org/10.3390/admsci12010031.
[32] Hübscher R. Assigning students to groups using general and context-specific criteria. IEEE Trans Learn Technol 2010;3(3):178–89. https://doi.org/10.1109/TLT.2010.17.
[33] Revelo-Sánchez O, Collazos CA, Redondo MA. Collaborative learning group formation based on personality traits: an empirical study in initial programming courses. In: Agredo-Delgado V, Ruiz PH, Villalba-Condori KO, editors. Human-computer interaction. communications in computer and information science (Vol. 1334, pp 73–84). Springer; 2020. https://doi.org/10.1007/978-3-030-66919-5_8.
[34] Bell ST, Brown SG, Colaneri A, Outland N. Team composition and the ABCs of teamwork. Am Psychol 2018;73(4):349–62. https://doi.org/10.1037/amp0000305.
[35] Bowers CA, Pharmer JA, Salas E. When member homogeneity is needed in work teams: a meta-analysis. Small Group Res 2000;31(3):305–27. https://doi.org/10.1177/104649640003100303.
[36] Hitt MA, Beamish PW, Jackson SE, Mathieu JE. Building theoretical and empirical bridges across levels: multilevel research in management. Acad Manage J 2007;50(6):1385–99. https://doi.org/10.5465/amj.2007.28166219.
[37] Blanco-Fernández D, Leitner S, Rausch A. Interactions between the individual and the group level in organizations: the case of learning and group turnover. Cent Eur J Oper Res 2023;31:1087–128. https://doi.org/10.1007/s10100-023-00843-7.
[38] Kozlowski SWJ, Bell BS. Work groups and groups in organizations: review update. In: Schmitt N, Highhouse S, editors. Handbook of psychology. Industrial and organizational psychology (2nd editor, Vol. 12, pp. 412–69). Wiley; 2013. https://doi.org/10.1002/0471264385.wei1214.
[39] Mannix E, Neale MA. What differences make a difference? The promise and reality of diverse teams in organizations. Psychol Sci Pub Int 2016;6(2):31–55. https://doi.org/10.1111/j.1529-1006.2005.00022.x.
[40] van Knippenberg D, Schippers MC. Work group diversity. Annu Rev Psychol 2007;58(1):515–41. https://doi.org/10.1146/annurev.psych.58.110405.085546.
[41] Voltmer JB, Reich-Stiebert N, Raimann J, Stürmer S. The role of multi-attributional student diversity in computer-supported collaborative learning. Internet High Educ 2022;55:100868. https://doi.org/10.1016/j.iheduc.2022.100868.
[42] Gawande AA, Zinner MJ, Studdert DM, Brennan TA. Analysis of errors reported by surgeons at three teaching hospitals. Surgery 2003;133(6):614–21. https://doi.org/10.1067/msy.2003.169.
[43] Zennouche M, Zhang J, Wang B. Factors influencing innovation at individual, group and organisational levels: a content analysis. Int J Inf Syst Change Manag 2014;7(1):23–42. https://doi.org/10.1504/IJISCM.2014.065052.
[44] Hough LM, Oswald FL. Personnel selection: looking toward the future - remembering the past. Annu Rev Psychol 2000;51(1):631–64. https://doi.org/10.1146/annurev.psych.51.1.631.
[45] Hough LM, Oswald FL. They're right, well ... mostly right: research evidence and an agenda to rescue personality testing from 1960s insights. Hum Perform 2005;18(4):373–87. https://doi.org/10.1207/s15327043hup1804_4.
[46] Saucier G, Ostendorf F. Hierarchical subcomponents of the Big Five personality factors: a cross-language replication. J Pers Soc Psychol 1999;76(4):613–27. https://doi.org/10.1037/0022-3514.76.4.613.
[47] Stewart GL. Trait bandwidth and stages of job performance: assessing differential effects for conscientiousness and its subtraits. J Appl Psychol 1999;84(6):959–68.
[48] Costa P, McCrae R. Revised NEO personality (NEOPI-R) and NEO five factor inventory (NEO-FFI) professional manual. Psychol Assess Resour 1992. http://www.oalib.com/references/7491296.
[49] Costa P, McCrae R, Boyle GJ, Matthews G, Saklofske DH. The revised NEO personality inventory (NEO-PI-R). 2008. The SAGE handbook of personality theory and assessment: Volume 2 — Personality measurement and testing. SAGE Publications Ltd; 2008. https://doi.org/10.4135/9781849200479.
[50] Bell ST. Deep-level composition variables as predictors of team performance: a meta-analysis. J Appl Psychol 2007;92(3):595–615. https://doi.org/10.1037/0021-9010.92.3.595.
[51] MacDonnell R, O'Neill TA, Kline TJ, Hambley LA. Bringing group-level personality to the electronic realm: a comparison of face-to-face and virtual contexts. Psychol-Manag J 2009;12(1):1–24. https://doi.org/10.1080/10887150802371773.
[52] Peeters MAG, Van Tuijl HFJM, Rutte CG, Reymen IMMJ. Personality and group performance: a meta-analysis. Eur J Pers 2006;20(5):377–96. https://doi.org/10.1002/per.588.
[53] Wang ML, Hsu BF. A study to explore the team virtualization level and team effectiveness from the team personality composition. J Knowl Econ 2012;3(2):199–216. https://doi.org/10.1007/s13132-011-0079-y.
[54] Fleeson W, Gallagher P. The implications of big five standing for the distribution of trait manifestation in behavior: fifteen experience-sampling studies and a meta-analysis. J Pers Soc Psychol 2009;97(6):1097–114. https://doi.org/10.1037/a0016786.
[55] Marks MA, Mathieu JE, Zaccaro SJ. A temporally based framework and taxonomy of team processes. Acad Manage Rev 2001;26(3):356–76. https://doi.org/10.2307/259182.
[56] Perry JL. Work team diversity: refocusing through the lens of team power and status. Soc Personal Psychol Compass 2021;15(12):e12646. https://doi.org/10.1111/spc3.12646.
[57] Prewett MS, Brown MI, Goswami A, Christiansen ND. Effects of team personality composition on member performance: a multilevel perspective. Group Organ Manag 2018;43(2):316–48. https://doi.org/10.1177/1059601116668633.
[58] Driskell JE, Goodwin GF, Salas E, O'Shea PG. What makes a good team player? Personality and team effectiveness. Group Dyn: Theory Res Prac 2006;10(4):249–71. https://doi.org/10.1037/1089-2699.10.4.249.
[59] Lykourentzou I, Antoniou A, Naudet Y, Dow SP. Personality matters: balancing for personality types leads to better outcomes for crowd groups. In: 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing (CSCW '16); 2016. p. 260–73. https://doi.org/10.1145/2818048.2819979.
[60] Barrick MR, Stewart GL, Neubert MJ, Mount MK. Relating member ability and personality to work-team processes and team effectiveness. J Appl Psychol 1998;83(3):377–91. https://doi.org/10.1037/0021-9010.83.3.377.
[61] Hough LM. The "big five" personality variables–construct confusion: description versus prediction. Hum Perform 1992;5(1–2):139–55. https://doi.org/10.1080/08959285.1992.9667929.
[62] Judge TA, Bono JE, Ilies R, Gerhardt MW. Personality and leadership: a qualitative and quantitative review. J Appl Psychol 2002;87(4):765–80. https://doi.org/10.1037/0021-9010.87.4.765.
[63] Nonaka I, Hirose A, Takeda Y, Friedrich TL, Vessey WB, Schuelke MJ, Ruark GA, Mumford MD, D'innocenzo L, Mathieu JE, Kukenberger MR, Wang D, Waldman D, Zhang Z, Banks GC, McCauley KD, Gardner WL, Guler CE, Brower HH, Sawyer KB. What type of leadership behaviors are functional in teams? A meta-analysis. Leadership Q 2016;17(3):288–307. https://doi.org/10.1016/j.leaqua.2006.02.007.
[64] Krämer A, Bhave DP, Johnson TD. Personality and group performance: the importance of personality composition and work tasks. Pers Individ Dif 2014;58(24):132–7. https://doi.org/10.1016/j.paid.2013.10.019.
[65] Tiedens LZ, Fragale AR. Power moves: complementarity in dominant and submissive nonverbal behavior. J Pers Soc Psychol 2003;84(3):558–68. https://doi.org/10.1037/0022-3514.84.3.558.
[66] Yellen R, Winniford M, Sanford C. Extraversion and introversion in electronically-supported meetings. Inf Manag 1995;1(28):63–74. https://doi.org/10.1016/0378-7206(94)00023-C.
[67] Kiesler DJ. The 1982 interpersonal circle: a taxonomy for complementarity in human transactions. Psychol Rev 1983;90(3):185–214. https://doi.org/10.1037/0033-295X.90.3.185.
[68] Harris K, English T, Harms P, Gross J, Jackson J, Back M. Why are extraverts more satisfied? Personality, social experiences, and subjective well–being in college. Eur J Pers 2017;31:170–86. https://doi.org/10.1002/per.2101.
[69] Lu L, Hu C. Personality, leisure experiences and happiness. J Happiness Stud 2005;6:325–42. https://doi.org/10.1007/S10902-005-8628-3.
[70] Humphrey SE, Hollenbeck JR, Meyer CJ, Ilgen DR. Trait configurations in self-managed groups: a conceptual examination of the use of seeding for maximizing and minimizing trait variance in groups. J Appl Psychol 2007;92(3):885–92. https://doi.org/10.1037/0021-9010.92.3.885.
[71] Tekleap AG, Quigley NR. Team deep-level diversity, relationship conflict, and team members' affective reactions: a cross-level investigation. J Bus Res 2014;67(3):394–402. https://doi.org/10.1016/j.jbusres.2012.12.022.
[72] den Hartog SC, Runge JM, Reindl G, Lang JWB. Linking personality trait variance in self-managed teams to team innovation. Small Group Res 2019;51(2):265–95. https://doi.org/10.1177/1046496419865325.
[73] Roney R, Greenaway K, Anicich EM, Galinsky AD. The path to glory is paved with hierarchy: when hierarchical differentiation increases group effectiveness. Psychol Sci 2012;23(6):669–77. https://doi.org/10.1177/0956797611433876.
[74] Thanh PTH, Gillies R. Group composition of cooperative learning: does heterogeneous grouping work in Asian classrooms? Int Educ Stud 2010;3(3):12–9. https://doi.org/10.5539/ies.v3n3p12.
[75] Woolley AW, Chow RM, Mayo AT, Riedl C, Chang JW. Collective attention and collective intelligence: the role of hierarchy and team gender composition. Organization Science 2022;34(3):1315–31. https://doi.org/10.1287/orsc.2022.1602.

[76] Kirschner PA. Stop propagating the learning styles myth. Comput Educ 2017;106:166–71. https://doi.org/10.1016/j.compedu.2016.12.006.

[77] Klein C, DiazGranados D, Salas E, Le H, Burke CS, Lyons R, Goodwin GF. Does team building work? Small Group Res 2009;40(2):181–222. https://doi.org/10.1177/1046496408328821.

[78] Eckhaus E, Davidovitch N. Technology-supported teaching: technological progress or a sham? Eur J Edu Res 2019;8(3):697–702. https://doi.org/10.12973/eu-jer.8.3.697.

[79] Eckhaus E, Klein G, Kantor J. Experiential learning in management education. Bus Manage Econ Eng 2017;15(1):42–56. https://doi.org/10.3846/bme.2017.345.

[80] Filade BA, Bello AA, Uwaoma C, Anwanane BB, Nwangburuka K. Peer group influence on academic performance of undergraduate students in Babcock University, Ogun State, Nigeria. Afr Educ Res J 2019;7(2):81–7. https://doi.org/10.30918/AERJ.72.19.010.

[81] Zheng Z, Pinkwart N. A discrete particle swarm optimization approach to compose heterogeneous learning groups. In: Proceedings of the 2014 IEEE 14th International Conference on Advanced Learning Technologies. IEEE Computer Society; 2014. p. 49–51. https://doi.org/10.1109/ICALT.2014.24.

[82] Mehar R, Kaur P. Effect of online collaborative learning strategy on achievement in economics in relation to self-efficacy. Int J Scientif Technol Res 2020;2(9):6544–54.

[83] Shaw RS. The relationships among group size, participation, and performance of programming language learning supported with online forums. Comput Educ 2013;62:196–207. https://doi.org/10.1016/j.compedu.2012.11.001.

[84] Putwain DW, Becker S, Symes W, Pekrun R. Reciprocal relations between students' academic enjoyment, boredom, and achievement over time. Learn Instr 2018;54:73–81. https://doi.org/10.1016/j.learninstruc.2017.08.004.

[85] French KA, Kottke JL. Teamwork satisfaction: exploring the multi-level interaction of groupwork interest and group extraversion. Active Learn Higher Educ 2013;14(3):189–200. https://doi.org/10.1177/1469787413498034.

[86] Lau K, Jin Q. Chinese students' group work performance: does team personality composition matter? Educ Training 2019;61(3):290–309. https://doi.org/10.1108/et-06-2018-0141.

[87] Stipelman BA, Rice EL, Vogel AL, Hall KL. The role of team personality in team effectiveness and performance. In: Hall K, Vogel A, Croyle R, editors. Strategies for Team Science Success. Springer; 2019. p. 189–96. https://doi.org/10.1007/978-3-030-20992-6_14.

[88] Häuberer, J. (2011). The founding concepts of social capital - bourdieu's theory of capital and coleman's rational-choice approach to social capital. in: social capital theory. VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-92646-9_2.

[89] Oh H, Chung MH, Labianca G. Group social capital and group effectiveness: the role of informal socializing ties. Acad Manage J 2004;47(6):860–75. https://doi.org/10.5465/20159627.

[90] Loignon AC, Woehr DJ, Loughry ML, Ohland MW. Elaborating on team-member disagreement: examining patterned dispersion in team-level constructs. Group Organ Manag 2018;44(1):165–210. https://doi.org/10.1177/1059601118776750.

[91] Bekele R, Menzel W. A Bayesian approach to predict performance of a student (BAPPS): a case with Ethiopian students. Artif Intell Appl 2005;22(23):189–94. http://www.actapress.com/Abstract.aspx?paperId=18917.

[92] Moore RL. The effect of group composition on individual student performance in an introductory economics course. J Econ Educ 2011;42(2):120–35. https://doi.org/10.1080/00220485.2011.555694.

[93] Seong JY, Hong D-S. Collective personality fit and diversity – how effective are they in predicting relationship conflict? Int J Conflict Manage 2020;31(2):201–19. https://doi.org/10.1108/IJCMA-03-2019-0055.

[94] De Dreu CKW, Weingart LR. A contingency theory of task conflict and performance in groups and organizational teams. In: Wes MA, Tjosvold D, Smith K, editors. International Handbook of Organizational Teamwork and Cooperative Working. Wiley; 2003. p. 151–66.

[95] McPherson M, Smith-Lovin L, Cook JM. Birds of a feather: homophily in social networks. Rev Sociol 2001;27:415–44. https://doi.org/10.1146/annurev.soc.27.1.415.

[96] Bradley JH, Hebert FJ. The effect of personality type on team performance. J Manage Develop 1997;16(5):337–53. https://doi.org/10.1108/02621719710174525.

[97] Martín, E., Paredes, P. (2004). Using learning styles for dynamic group formation in adaptive collaborative hypermedia systems. In M. Matera & S. Comai (editors), *Engineering Advanced Web Applications. Proceedings of Workshops in Connection with 4th International Conference on Web Engineering (ICWE 2004)*, Rinton Press. 188–97. https://repositorio.uam.es/handle/10486/3680.

[98] Byrne D. The attraction paradigm. Academic Press; 1971.

[99] Humphrey SE, Hollenbeck JR, Meyer CJ, Ilgen DR. Personality configurations in self-managed groups: a natural experiment on the effects of maximizing and minimizing variance in traits. J Appl Soc Psychol 2011;41(7):1701–32. https://doi.org/10.1111/j.1559-1816.2011.00778.x.

[100] Lakens, D. (2021). *Sample size justification*. https://doi.org/10.31234/osf.io/9d3yf.

[101] Rammstedt B, John OP. Short version of the big five inventory (BFI-K): development and validation of an economic inventory for assessment of the five factors of personality. Diagnostica 2005;51(4):195–206. https://doi.org/10.1026/0012-1924.51.4.195.

[102] R. Core Team (2014). R: a language and environment for statistical computing. R Foundation for Statistical Computing. http://www.R-project.org/.

[103] Pinheiro, J., Bates, D., DebRoy, S., Sarkar, D., & R. Core Team. (2021). nlme: linear and nonlinear mixed effects models. [Computer software]. https://CRAN.R-project.org/package=nlme.

[104] Burnham KP, Anderson DR. Multimodel inference. Sociol Methods Res 2016. https://doi.org/10.1177/0049124104268644.

[105] Van Landeghem G, De Fraine B, Van Damme J. The consequence of ignoring a level of nesting in multilevel analysis: a comment. Multivariate Behav Res 2005;40(4):423–34. https://doi.org/10.1207/s15327906mbr4004_2.

[106] Hox JJ, Moerbeek M, van de Schoot R. Multilevel Analysis: Techniques and Applications. 3rd editor. Routledge; 2017. https://doi.org/10.4324/9781315650982.

[107] Oh H, Chung MH, Labianca G. Group social capital and group effectiveness: The role of informal socializing ties. Academy of Management Journal 2004;47(6):860–75. https://doi.org/10.5465/20159627.

[108] Bell ST, Brown SG, Colaneri A, Outland N. Team composition and the ABCs of teamwork. American Psychologist 2018;73(4):349–62. https://doi.org/10.1037/amp0000305.

[109] Mccambridge J, Kypri K, Elbourne D. Research participation effects: a skeleton in the methodological cupboard. J Clin Epidemiol 2014;67:845–9. https://doi.org/10.1016/j.jclinepi.2014.03.002.

[110] Nijstad BA, De Dreu CKW. Creativity and group innovation. Appl Psychol 2002;51(3):400–6. https://doi.org/10.1111/1464-0597.00984.

[111] Rienties B, Héliot Y. Enhancing (in) formal learning ties in interdisciplinary management courses: A quasi-experimental social network study. Studies in Higher Education 2018;43(3):437–51. https://doi.org/10.1080/03075079.2016.1174986.

[112] Okun MA, Pugliese J, Rook KS. Unpacking the relation between extraversion and volunteering in later life: the role of social capital. Pers Individ Dif 2007;42(8):1467–77. https://doi.org/10.1016/j.paid.2006.10.020.

[113] Fukuyama F. Social capital, civil society and development. Third World Q 2001;22(1):7–20.

[114] Ramos-Villagrasa PJ, Marques-Quinteiro P, Navarro J, Rico R. Teams as complex adaptive systems: reviewing 17 years of research. Small Group Res 2018;49(2):135–76. https://doi.org/10.1177/1046496417713849.

[115] Bendersky C, Shah NP. The downfall of extraverts and rise of neurotics: the dynamic process of status allocation in task groups. Acad Manage J 2013;56(2):387–406. https://doi.org/10.5465/amj.2011.0316.

[116] Kalish Y, Luria G. Traits and time in leadership emergence: a longitudinal study. Leadersh Q 2021;32(2):101443. https://doi.org/10.1016/j.leaqua.2020.101443.

[117] Lin N. Social capital. A theory of social structure and action. Cambridge University Press; 2001.

[118] Curşeu, P.L., Ilies, R., Vîrgă, D., Maricuţoiu, L., & Sava, F.A. (2019). Personality characteristics that are valued in teams: not always "more is better"?.

[119] Junker TL, Bakker AB, Gorgievski MJ, Derks D. Agile work practices and employee proactivity: a multilevel study. Hum Relat 2022;75(12):2189–217. https://doi.org/10.1177/00187267211030101.

[120] Blasco-Arcas L, Buil I, Hernández-Ortega B, Sese FJ. Using clickers in class. The role of interactivity, active collaborative learning and engagement in learning performance. Comput Educ 2013;62:102–10. https://doi.org/10.1016/j.compedu.2012.10.019.

[121] Fryer LK, Dinsmore DL. The promise and pitfalls of self-report: development, research design and analysis issues, and multiple methods. Front Learn Res 2020;8(3):1–9. https://doi.org/10.14786/flr.v8i3.623.

[122] Virk A, Joshi A, Mahajan R, Singh T. The power of subjectivity in competency-based assessment. J Postgrad Med 2020;66(4):200–5. https://doi.org/10.4103/jpgm.JPGM_591_20.

[123] van der Linden D, te Nijenhuis J, Bakker AB. The general factor of personality: a meta-analysis of big five intercorrelations and a criterion-related validity study. J Res Pers 2010;44(3):315–27. https://doi.org/10.1016/j.jrp.2010.03.003.

[124] Kickul J, Neuman G. Emergent leadership behaviors: The function of personality and cognitive ability in determining teamwork performance and KSAs. Journal of Business and Psychology 2000;15:27–51. https://doi.org/10.1023/A:1007714801558.

[125] Rogat TK, Hmelo-Silver CE, Cheng BH, Traynor A, Adeoye TF, Gomoll A, Downing BK. A multidimensional framework of collaborative groups' disciplinary engagement. Frontline Learning Research 2022;10(2):1–21. https://doi.org/10.14786/flr.v10i2.863.

[126] Magpili NC, Pazos P. Self-managing team performance: a systematic review of multilevel input factors. Small Group Res 2018;49(1):3–33. https://doi.org/10.1177/1046496417710500.

[127] Seeber I, Bittner E, Briggs RO, DeVreede GJ, DeVreede T, Druckenmiller D, Maier R, Merz AB, Oeste-Reiß S, Randrup N, Schwabe G, Söllner M. Machines as teammates: a collaboration research agenda. In: Proceedings of the 51st Hawaii International Conference on System Sciences (HICSS); 2018. http://hdl.handle.net/10125/49942.

[128] Mirbabaie M, Stieglitz S, Brünker F, Hofeditz L, Ross B, Frick NR. Understanding collaboration with virtual assistants–the role of social identity and the extended self. Business & Information Systems Engineering 2021;63:21–37. https://doi.org/10.1007/s12599-020-00672-x.

[129] Siemon D. Elaborating team roles for artificial intelligence-based teammates in Human-AI collaboration. Group Decision and Negotiation 2022;31:871–912. https://doi.org/10.1007/s10726-022-09792-z.

[130] Isotani S, Mizoguchi R, Isotani S, Capeli OM, Isotani N, De Albuquerque AR, Jaques P. A Semantic Web-based authoring tool to facilitate the planning of collaborative learning scenarios compliant with learning theories. Computers & Education 2013;63:267–84.

[131] Lambić D, Lazović B, Djenić A, Marić M. A novel metaheuristic approach for collaborative learning group formation. Journal of Computer Assisted Learning 2018;34(6):907–16. https://doi.org/10.1111/jcal.12299.
[132] Thumlert K, McBride M, Tomin B, Nolan J, Lotherington H, Boreland T. Algorithmic literacies: Identifying educational models and heuristics for engaging the challenge of algorithmic culture. Digital Culture & Education 2022;14(4):19–35.
[133] Sukstrienwong Anon. Group Formation based on Students' Learning Styles by Circular Genetic Algorithm. TEM Journal 2021;10:1016–21. https://doi.org/10.18421/TEM103-02.
[134] Garcia-Garcia A, Orts S, Oprea S, Villena-Martinez V, Martinez-Gonzalez P, Rodríguez JG. A survey on deep learning techniques for image and video semantic segmentation. Appl. Soft Comput. 2018;70:41–65. https://doi.org/10.1016/j.asoc.2018.05.018.
[135] McCabe KO, Fleeson W. What is extraversion for? Integrating trait and motivational perspectives and identifying the purpose of extraversion. Psychological Science 2012;23(12):1498–505. https://doi.org/10.1177/0956797612444904.
[136] Porter COLH, Hollenbeck JR, Ilgen DR, Ellis APJ, West BJ, Moon H. Backing up behaviors in groups: The role of personality and legitimacy of need. Journal of Applied Psychology 2003;88(3):391–403. https://doi.org/10.1037/0021-9010.88.3.391.
[137] Müller A, Röpke R, Konert J, Bellhäuser H. Investigating group formation: An experiment on the distribution of extraversion in educational settings. Acta psychologica 2023;242:104111. https://doi.org/10.1016/j.actpsy.2023.104111.
[138] Van Dijk H, Meyer B, van Engen ML, Lewin Loyd D. Microdynamics in diverse groups: A review and integration of the diversity and stereotyping literatures. Academy of Management Annals 2017;11(1):517–57. https://doi.org/10.5465/annals.2014.0046.
[139] Jackson CM, Jackson JC, Bilkey D, Jong J, Halberstadt J. The dynamic emergence of minimal groups. Group Processes and Intergroup Relations 2019;22(7):921–9. https://doi.org/10.1177/1368430218802636.
[140] Konert J, Bellhäuser H, Röpke R, Gallwas E, Zucik A. MoodlePeers: Factors relevant in learning group formation for improved learning outcomes, satisfaction and commitment in E- Learning scenarios using GroupAL. In: Verbert K, Sharples M, Klobucar T, editors. Adaptive and Adaptable Learning: Proceedings of the 11th European Conference on Technology Enhanced Learning, EC-TEL 2016, Lecture Notes in Computer Science. 9891; 2016. p. 390–6. https://doi.org/10.1007/978-3-319-45153-4_32.
[141] Konert J, Burlak D, Steinmetz R. The group formation problem: an algorithmic approach to learning group formation. Lecture Notes in Computer Science. 2014. p. 221–34. https://doi.org/10.1007/978-3-319-11200-8_17.
[142] Salas E, Cooke NJ, Rosen MA. On Teams, Teamwork, and Team Performance: Discoveries and Developments. Human Factors 2008;50(3):540–7. https://doi.org/10.1518/001872008X288457.