Computers & Education: X Reality 4 (2024) 100063

Contents lists available at ScienceDirect

Computers & Education: X Reality

journal homepage: www.journals.elsevier.com/computers-and-education-x-reality

Applying multimodal data fusion to track autistic adolescents' representational flexibility development during virtual reality-based training

Jewoong Moon [a,*], Fengfeng Ke [b], Zlatko Sokolikj [b], Shayok Chakraborty [b]

[a] The University of Alabama, Tuscaloosa, AL, USA
[b] Florida State University, Tallahassee, FL, USA

A R T I C L E I N F O

Keywords:
Multimodal data fusion
Virtual reality
Representational flexibility
Autism spectrum disorder
Educational data mining

A B S T R A C T

In our study, we harnessed multimodal data to develop a predictive model aimed at assessing the development of representational flexibility (RF) in autistic adolescents engaged in virtual reality (VR)-based cognitive skills training. Recognizing VR's potential to enhance RF through immersive 3D simulation tasks, we addressed the research gap in analyzing learners' digital interactions within this environment. This data mining study integrated diverse data sources—including behavioral cues, physiological responses, and direct interaction logs—collected from 178 training sessions with eight autistic adolescents. This comprehensive dataset, encompassing both audio and screen recordings, was analyzed using advanced machine learning techniques. Through decision-level data fusion, particularly employing the random forest algorithm, our model demonstrated enhanced accuracy in predicting RF development, surpassing single-source data approaches. This research not only contributes to the effective use of VR in educational interventions for autistic adolescents but also showcases the potential of multimodal data fusion in understanding complex cognitive skills development.

The underrepresentation of autistic adolescents in the fields of science, technology, engineering, and mathematics (STEM) has been a critical issue that demands extensive research and practices (Ehsan et al., 2018). Autistic individuals often encounter developmental hurdles, such as different social and cognitive conditions, that hinder their engagement and success in STEM areas. One of the key skills that is typically less developed in autistic learners is representational flexibility (RF), which involves the ability to comprehend, adapt, and utilize various representations crucial in STEM learning. Abilities such as mental rotation and perspective-taking are integral to RF and have been linked to improved understanding of complex STEM tasks (Cartwright, 2008; Farrar & Ashwell, 2008).

This study has been propelled by the potential of virtual reality (VR) to significantly enhance autistic adolescents' RF development (Moon et al., 2020; Sokolikj et al., 2023). The immersive nature of VR allows for a unique, engaging learning environment that is highly adaptable to the individual needs of learners. Unlike other technologies, VR has provided a three-dimensional, visually-enriched, and highly-interactive digital space where autistic adolescents can practice and develop RF skills in a controlled, yet flexible setting. This functional capability is especially critical for autistic learners, as it supports personalized learning trajectories and encourages active engagement with STEM concepts through visual and sensory experiences (Dahlstrom-Hakki et al., 2021, 2024). However, a notable gap exists in previous studies to effectively monitor and evaluate these learners' progression in VR training (Drey et al., 2020). Limited empirical research on in-VR automatic assessment approaches reflects a broader challenge in capturing the nuanced, dynamic learning processes of autistic students within immersive learning environments (Lorenzo et al., 2019, 2023; Rizzo et al., 2001). Conventional learning assessment methods often appear limited in accounting for the complex interplay of cognitive, emotional, and behavioral factors that influence learning in these settings (Spector, 2006; Webb, 1980).

To further address these challenges, the exploration of multimodal data fusion becomes crucial in the design of assessments for immersive learning environments. Multimodal data fusion is a data mining technique that systematically integrate various types of data to achieve a more holistic analysis than what could be obtained through any single data source alone (Chango et al., 2022; Gaw et al., 2022). Multimodal data fusion involves the strategic amalgamation of various data types—ranging from behavioral and physiological signals to interaction

* Corresponding author.
E-mail addresses: jmoon19@ua.edu (J. Moon), fke@fsu.edu (F. Ke), zsokolikj@fsu.edu (Z. Sokolikj), shayok@cs.fsu.edu (S. Chakraborty).

https://doi.org/10.1016/j.cexr.2024.100063
Received 21 November 2023; Received in revised form 24 March 2024; Accepted 24 March 2024
Available online 4 April 2024

patterns and learning analytics—to construct a detailed, multidimensional understanding of students' learning experiences. This technique, by leveraging the strengths of diverse data sources, has revolutionized the assessment and enhancement of autistic students' learning processes in VR settings.

To bridge the existing gap in understanding and supporting autistic learners within VR-based learning environments, the present study proposes and evaluates the development of a predictive model that utilizes multimodal data fusion. This approach is suggested to highly enhance the comprehension of how autistic students engage and progress in immersive learning environmental settings. By harnessing multiple data sources that span behavioral, physiological, and interactional dimensions, this model is designed to offer a comprehensive view of the nuanced learning-related behaviors exhibited by these students (Akhuseyinoglu & Brusilovsky, 2021).Such an approach is not only novel but also highly relevant for creating personalized learning experiences that can address the unique challenges faced by autistic adolescents in developing RF. The utilization of multimodal data fusion in VR-based training environments opens up new avenues for empirical research (Jin, Liu, Yarosh, Han, & Qian, 2022), enriching learning experiences. This inquiry is particularly pertinent given the potential of VR technology to create immersive learning experiences that can be tailored to autistic learners' individual needs (Zhang et al., 2022). The predictive model's ability to synthesize diverse data types promises to provide deeper insights into how these students engage with and benefit from VR-based STEM education. The overarching research question of this study is as follows: How do data mining techniques perform in tracking and assessing autistic adolescents' representational flexibility development in VR-based training, comparing unimodal versus multimodal data sources?

1. Literature review

1.1. Representational flexibility in integrated STEM education

Representational Flexibility (RF) is an important characteristic that helps humans pursue complex tasks (Ionescu, 2012). RF refers to the cognitive skill of shifting focus and selecting different forms of representations while processing information (Acevedo Nistal et al., 2009; Herbert & Hayne, 2000). Individuals with strong RF are adept at understanding, replicating, and altering knowledge to tailor their problem-solving approaches to various situations (Ionescu, 2012). RF requires individuals to understand and reproduce mental representations following spatial description reading, including perspectives and depictions. It is particularly essential to consider a variety of factors to understand the multidimensional nature of spatial mental representations (Brunyé et al., 2008). This skill is closely linked to an individual's capacity to devise solutions and strategies in problem-solving scenarios (Acevedo Nistal et al., 2009; 2021; Gagatsis, Elia, & Mousoulides, 2006). There is an increasing attention to nurturing RF within integrated STEM education, with research indicating its close relationship to STEM-related skills (Pande & Chandrasekharan, 2017). Integrated STEM is a holistic approach to teaching and learning that exceeds the traditional boundaries of these individual disciplines (Heil et al., 2013). Instead of treating separate subjects, integrated STEM education combines them into a cohesive learning paradigm based on real-world applications. This approach is intended to promote a deeper understanding of each subject through the context of the others, promoting students to develop a and synthesize multidisciplinary perspective, as well as enhance problem-solving skills (Heba et al., 2017; Thibaut et al., 2018).

RF's role is particularly pronounced in integrated STEM education, a pedagogical approach that emphasizes the interconnectedness of science, technology, engineering, and mathematics (STEM) subjects. During these activities, they often participate in multiple rounds of problem-solving. For example, engineering tasks in STEM education generally require students to select, adjust, and reinvent representations suited to the given context (Hadgraft & Kolmos, 2020; Mentzer, 2011). These tasks typically involve scenarios that require them to synthesize and apply knowledge, where students should collect and analyze ill-structured data to grasp the nuances of a design challenge and its potential solution. Through repeated testing and prototype refinement, students become adept at selecting and evaluating various representations. Integrated STEM education requires learners to perform information retrieval and processing in multifaceted learning tasks.

Autistic adolescents often encounter significant challenges in integrated STEM education, primarily due to differences in their RF development. This skill enables learners to navigate and solve complex problems, which are fundamental in STEM disciplines. For instance, in mathematics, RF allows students to switch between different representations of a function—graphical, numerical, and algebraic—enhancing their understanding and problem-solving abilities. In science, it enables the interpretation of data through different modalities, such as charts or simulations, fostering a deeper comprehension of scientific concepts. Research has consistently shown that autistic adolescents with lower RF face difficulties in engaging with the dynamic problem-solving tasks inherent in STEM education (Ke et al., 2020; Moon, Ke, Sokolikj, & Dahlstrom-Hakki, 2022). The cognitive-affective states of autistic individuals, where learning-related stress may reduce inhibition and hinder attention-switching, can lead to quick disengagement when faced with the intricacies of STEM tasks (Messinger et al., 2015; De Vries & Geurts, 2015). For example, an autistic learner might find a multi-step physics problem overwhelming if they struggle to conceptualize the problem through different representational forms, such as diagrams and equations. The burgeoning interest in nurturing RF within integrated STEM education reflects a broader recognition of its critical role in equipping students with the skills necessary to thrive in an increasingly complex and technology-driven world.

1.2. Multimodal representations in virtual reality

VR environments inherently provide a multi-sensory experience (Sveistrup, 2004), incorporating visual, auditory, and at times, tactile feedback, catering to a wide range of preferences (Melo et al., 2020). This multi-sensory engagement allows learners to interact with content in various ways, including through visual-spatial, auditory, and physical interactions (Anastopoulou et al., 2011; Massaro, 2012). VR immerses users more fully than conventional educational settings, fostering a deeper connection to the virtual space and making the learning journey more engaging and memorable (Makransky & Mayer, 2022; Radianti et al., 2020). Such direct and immersive digital experiences aid in grasping and retaining complex concepts and processes, thus improving a learner's capacity for mental visualization and object manipulation (Molina-Carmona et al., 2018; Salzman et al., 1999). VR is especially advantageous for disciplines that depend on understanding spatial relations and visual representations (Paes et al., 2017). Through direct interaction with 3D objects and immediate observation in a virtual world, learners gain a richer comprehension of spatial relationships and concepts (Fogarty et al., 2018), enhancing essential spatial reasoning abilities—an integral aspect of representational flexibility. For autistic learners, who may exhibit unique learning profiles and sensitivities, the opportunity to engage with educational content through various modalities can offer a more inclusive and comfortable learning environment. VR supports the development of highly personalized learning experiences.

1.3. Virtual reality-based training for autistic learners

Innovations in virtual reality (VR)-based training programs have been increasingly employed as a highly-interactive tool to aid autistic learners in developing their social and cognitive abilities (Parsons & Cobb, 2011). VR is unique to be particularly effective in enhancing autistic learners' RF development (Babu et al., 2018; Zhang et al., 2022).

First, VR engages autistic learners through interactive 3D environments that include 3D objects and their simulations to allow them to select and adapt various representations. immersive and open-ended virtual environments provide autistic learners with the tools to conduct 3D simulations and experiments, allowing them to test hypotheses related to specific design challenges. The application of VR in activities that emphasize inquiry, design, and problem-solving significantly enhances their grasp of fundamental STEM principles (Makransky & Petersen, 2021). Furthermore, research indicates that representational fluency, which involves a deep understanding of scientific concepts and reasoning through various means such as verbal explanations, graphs, and drawings, benefits greatly from VR. By offering a visually enriched and augmented 3D space, VR supports autistic learners in developing more concrete and tangible concepts and visualizations, thereby enhancing their ability to solve design problems effectively.

Prominent examples of VR-based training have recently emerged, demonstrating design-based problem solving approaches that engage autistic adolescents in 3D engineering design challenges (Ke et al., 2020; Moon et al., 2020). This study involves participants in the creation of elevation bridges, designed to meet the needs of various environmental and societal contexts. During this training, participants engage in collaborative and iterative design conversations with facilitators and non-playable characters (NPCs), working together to develop and test 3D prototype solutions. Additionally, another study by Ke and Lee (2016) explores the impact of VR-based design problem-solving on enhancing perspective-taking skills and flexibility in autistic children. This research particularly emphasizes the support of design-oriented social interactions, highlighting the potential of VR to foster significant skill development in this domain.

Despite the growing interest on VR interventions and their application for autistic youth, there is a significant gap in empirical studies regarding learning assessment of VR-based training. Research has emphasized that existing attempts show limited success in monitoring the cognitive-affective states of learners during VR-based learning (Moon et al., 2020, 2023). Understanding individuals' cognitive-affective states through real-time and multimodal behavioral data is critical in its depth, immediacy, and context-specific insights. This approach captures the nuanced ways in which these states influence the development of RF in dynamic learning environments like VR-based training. Real-time monitoring offers immediate feedback on learners' engagement and emotional responses, crucial for understanding the fluid nature of cognitive-affective states during complex tasks. Multimodal data, encompassing verbal or behavioral, provides a rich view of these states beyond what self-reported measures can achieve. This is particularly important for capturing the subtle, moment-to-moment changes in learner behavior, emotion, and their dynamics that impacts learning processes. Moreover, the reliance on questionnaires for estimating cognitive-affective states presents particular challenges for autistic individuals, who may encounter difficulties in accurately expressing their emotions, experiences, and perceptions through conventional verbal or written means (Hill et al., 2004; Molnar-Szakacs & Heaton, 2012). The inherent abstract nature of questionnaires, demanding interpretation and articulation of responses (Block, 1998) can pose substantial obstacles. Autistic individuals' tendency to interpret questions very literally further complicates this issue, potentially leading to misunderstandings if the questionnaire design does not account for such literal interpretation.

Given these considerations, tracking their real-time and multimodal behaviors is thus crucial for a comprehensive understanding of their learning-related states, emphasizing the need for the development, implementation, and refinement of multimodal, data-driven assessments. Such assessments aim to drive timely and adaptive learning support, aligning with the unique developmental trajectories of autistic learners.

1.4. Data fusion in education research

While there is a burgeoning interest in VR-based training for autistic adolescents, research on VR-based learning assessment remains limited. Current research predominantly focuses on assessing learning enhancements by contrasting outcomes before and after VR interventions. This approach overlooks the necessity for in-process evaluations that can accurately trace the evolution of students' learning trajectories. Recent trends, however, indicate an increasing adoption of data analytics methods aimed at providing a more nuanced understanding of the learning process, as exemplified by the work of and Tlili and Chang (2019). Researchers have largely explored a range of computational methods to develop automated learning assessments that can identify learning-related behavioral evidence (Moon, Ke, Sokolikj, & Dahlstrom-Hakki, 2022; Yu & Denham, 2023). For instance, some studies have applied machine learning techniques to discern the cognitive-affective states of autistic learners. A notable example includes Moon et al. (2020), who developed a natural language processing (NLP)—based prediction model to track the growth of RF in autistic adolescents using supervised machine learning. This study aimed at analyzing their speech related to various aspects of RF. In addition, Chen, Cui, and Chu (2020) have utilized computer log data to create prediction models for assessing learning progress, employing methods like Bayesian knowledge tracing in the automatic game-based assessment system "Raging Skies". This game enables students to participate in various learning-related actions in Weather science: observing, describing, and interpreting weather phenomena and relating it to Earth's heating and cooling. It specifically targets six weather-related knowledge outcomes.

However, prediction model development approaches in existing studies face limitations in accurately tracking learners' states, often relying on unimodal data. This approach is likely to be less effective in VR environments that aim at multimodal interactions, making it challenging to fully capture and predict learning behaviors using solely unimodal data. In response to this challenge, multimodal learning analytics (MmLA) has emerged as a promising approach for observing learner behaviors in highly interactive digital learning environments (Worsley & Martinez-Maldonado, 2018). MmLA focuses on predicting learners' skill and knowledge development by aggregating and interpreting diverse data inputs with computational methodologies (Worsley & Martinez-Maldonado, 2018). It operates on the premise that learners manifest learning evidence through multiple channels, such as speech, gestures, and other observable actions. MmLA researchers have sought a way to capture and synthesize multiple data inputs effectively. Research increasingly suggests the feasibility of multimodal data fusion. Multimodal data fusion can be effectively applied to immersive learning environments. Immersive learning environments—such as VR—allow learners to interact with 3D stimuli, and such VR settings support multimodal behaviors. Integrating MmLA with data fusion in VR interventions could effectively collect and synthesize learner behaviors to drive adaptive learning support. For example, Sharma et al. (2019) successfully employed multimodal data fusion to merge physiological data from learners (e.g., eye-tracking, facial expression, and arousal). Henderson et al. (2020) developed a multimodal affect detection model by amassing and merging several forms of learner trace data. However, there has been a minimal investigation into the viability of building, synthesizing, and executing multimodal data fusion in VR-based training. Chango et al. (2022)'s comprehensive analysis discussed various data fusion techniques in MmLA studies across different educational settings: traditional in-person classroom, online environment, and hybrid/blend learning setting. They state that the adaptation of data fusion techniques to educational settings is still growing, with its potential for more sophisticated and accurate models of learning assessment. In particular, the findings echo those fewer studies focused on immersive learning environment contexts for underrepresented learner populations.

2. Methods

2.1. Comparison of data mining approaches

Aligned with the research question, our study includes a comprehensive comparison of the data mining techniques applied to different conditions of data sources. This study uniquely compares the efficacy of these techniques in accurately tracking and assessing the development of representational flexibility (RF) among autistic adolescents engaged in VR training. We systematically collected and processed multimodal data, including behavioral, physiological, and interactional signals, from VR training sessions. We implemented several data mining algorithms, including decision trees, random forests, and support vector machines, under various data conditions—namely, unimodal (single type of data source) and multimodal (fusion of multiple data sources) approaches. The comparison focuses on the algorithms' performance in identifying and predicting RF development indicators, considering the complexity and variability of autistic learners' interactions with the VR environment.

2.2. Data

We gathered two distinct categories of data: (1) recordings of both audio and video capturing the behaviors of autistic adolescents during their VR-based training sessions, and (2) computer-generated logs that documented the actions of learners throughout these VR training sessions. We considered systematic collection and processing of behavioral, physiological, and interactional data from VR training sessions. In total, we accumulated 178 recordings, each spanning 60–90 min of VR sessions involving eight autistic adolescents. These participants engaged in regular, weekly VR sessions focused on design and problem-solving tasks. As a key data source for our classifier, we captured computer-generated logs aligned with learner actions within the VR-based training sessions. These logs capture behaviors such as navigation movements through the virtual environment, interactions with virtual objects (e.g., selection, movement, modification), decision points and the choices made, time allocation to specific tasks, and instances of errors and subsequent corrections. For the purpose of processing and analyzing this data, we employed RapidMiner, a software specialized in data mining (Hofmann & Klinkenberg, 2016).

The determination of our sample size was guided by a careful consideration of several factors: the nature of the study and feasibility of recruiting participants within the specific population of autistic adolescents. Given the exploratory nature of our research on the effectiveness of various data mining techniques in a VR-based educational setting, we aimed to ensure that our sample size was sufficient to detect meaningful differences and trends in representational flexibility (RF) development. Considering the challenges associated with recruiting autistic adolescents for such specialized interventions, we aimed for a pragmatic yet scientifically justifiable sample size.

2.3. Selection process

We carefully designed the procedure of selecting the eight autistic adolescents for our VR-based training sessions. These criteria focused on ensuring that participants had a formal ASD diagnosis, fell within a targeted developmental and age range (12–18 years), possessed a minimum level of cognitive functioning suitable for VR interaction. In addition, in order to control for potential confounding factors, such as prior experience with digital platforms and subject matter familiarity, we conducted preliminary baseline assessments of all study participants. These assessments evaluated each participant's initial representational flexibility (RF) and cognitive abilities related to STEM learning, establishing a pre-intervention benchmark. This approach allowed us to isolate the VR-based training's impact on RF development. Qualitative data from reflective journals and feedback sessions provided further context on how prior experiences might influence engagement and learning outcomes within the VR environment. Through these steps, we ensured that our findings on the effectiveness of VR-based training in enhancing RF among autistic adolescents were accurately reflective of the intervention's impact, discounting pre-existing disparities in digital or STEM proficiency.

When selecting and including study participants of the training we prioritized ethical considerations and rigorous consent/assent processes, adhering to their respect and beneficence. Approval from the Florida State University IRB was obtained to ensure compliance with ethical standards. A detailed informed consent form was provided to study participants and their caregivers, explaining the study goal, procedures, potential risks, and benefits. A week before the training, an age-appropriate assent form was presented to the participants in the presence of their caregivers, accompanied by a demo of the VR training to offer a clear, tangible overview of the activities involved. This approach not only facilitated informed decision-making by both caregivers and participants but also set a foundation of trust and transparency. By allowing participants to experience the VR environment beforehand, we aimed to alleviate any concerns, adjust expectations, and ensure a comfortable and engaging training experience. The privacy rights of individuals involved in this study were maintained by securely storing all related data in a suitable protected repository.

2.4. Virtual reality-based training

This VR-based training program (see Fig. 1) includes a variety of design challenges, such as constructing elevation bridges and creating non-playable characters (NPCs), in order to foster RF development of autistic adolescents. The training program was systematically developed and repeatedly refined withing the framework of a long-term, design-based research initiative (Moon et al., 2020). During this training program, study participants were required to closely observe and analyze both natural and societal elements within a virtual world, craft 3D design solutions, and engage in repeated testing and refinement of the solutions as part of their design quests. To enhance understanding of the VR-based training program's design challenges and foster a deeper engagement with the developmental objectives for autistic adolescents, we offer detailed examples and scenarios that were integral to our study's tasks.

In the first training module, study participants were tasked with designing an 3D elevation bridge within a virtual environment that simulated a realistic island ecosystem. The scenario was set in a virtual world where the bridge was crucial for the island's inhabitants, enabling them to transport goods across a river while allowing for unobstructed ship passage beneath. Study participants roleplayed an engineer engaging in a complex design process that required them to consider various design factors, such as the bridge's height, structural integrity, and its functions to integrate seamlessly into the island's existing transportation network. The task involved using VR tools to construct the bridge piece by piece, starting from laying down the foundation to assembling the roadway and support columns. Study participants were then required to write and modify block-based computer scripts (LSL) to simulate the bridge's functionality. This hands-on approach allowed participants to experiment with Newtonian physics engineering principles and spatial reasoning in a supportive, immersive environment. The second module focused on the creation and development of interactive non-playable characters (NPCs) that could inhabit the virtual world. Study participants encountered challenge of designing NPCs with distinct personalities, roles, and behaviors that would interact with users and other elements within the VR landscape. This task involved drafting pseudocode or utilizing block-based scripting to operate the NPCs' actions, dialogue, and reactions to player inputs. For example, one scenario required participants to create an NPC shopkeeper, requiring the participants to code the NPC's ability to receive orders.

Fig. 1. Training module examples (Left: Elevation bridge design, Right: NPC design.

2.5. Data wrangling

We conducted four stages of data wrangling. First, we gathered recorded speech and logged behavior data from study participants during VR-based training. We analyzed study participants' verbal utterances and text chats to track their RF development during 3D design activities in a VR world. We gained learner speech data by recording audio and screen data sources. In terms of behavior tracking, we used a tracker application on the VR server, monitored and logged study participants' actions when they performed building or scripting actions. Fig. 2 demonstrates the structure of the log data.

Second, we organized the metadata from the gathered data sources based on 5-min segments of the VR-based training activities. This categorization is visually represented in Fig. 3. For instance, in a typical 60-min training session, this method resulted in the creation of 12 separate data instances. Overall, from the training sessions involving all study participants, we successfully retrieved a total of 1441 data instances. These instances were then systematically utilized in the training and testing of our prediction model. This structured approach to data segmentation allowed for more granular and detailed data analytics of the training session.

Third, we conducted manual ground-truth labeling to build the benchmarking data set on the enactment of RF. There are four facets of RF enactment: (1) attention switching/cognitive shifting (ASC), (2) alternative representation (AR), (3) pattern development (PD), and (4) pattern contextualization (PC). After conducting an extensive review of existing literature, we created a coding scheme which underwent multiple rounds of testing and refinement. Table 1 outlines four key aspects of RF. Expert coders, well-versed in both VR-based training and RF-related studies, were responsible for labeling and coding the video recordings from the training sessions. These coders initially worked independently to identify data instances of the various aspects of RF, then collaboratively reviewed and reconciled their findings to achieve complete consensus. The outcomes were categorized into three distinct labels: neutral, positive, and negative, for each RF aspect. Table 2 provides examples of both positive and negative occurrences pertaining to the RF aspects, with negative examples highlighting unsuccessful attempts when applying RF. The development of the coding criteria for this precise labeling was continuously honed through successive discussions among peer coders.

Fourth, we trained the prediction model based on the ground-truth labeling results. We converted the raw speech and log data sets to the ones ready for machine learning model training. We employed speech-to-text conversion software to transform the spoken words of the participants into textual data. Text vectorization was then performed using the NLP-driven machine learning tool LightSIDE (Mayfield & Rosé,

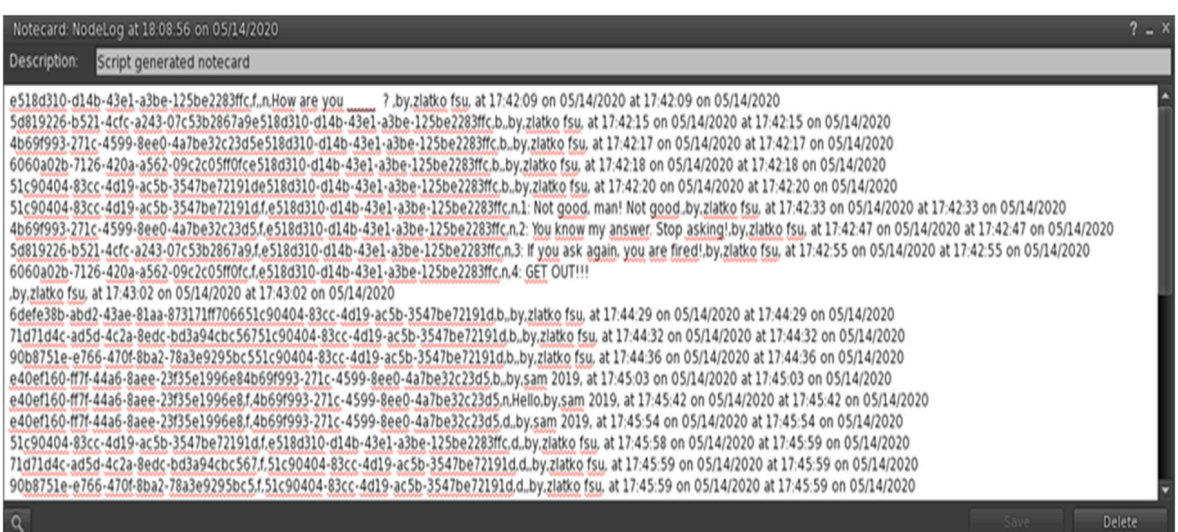

Fig. 2. Behavior logs in VR-based training.

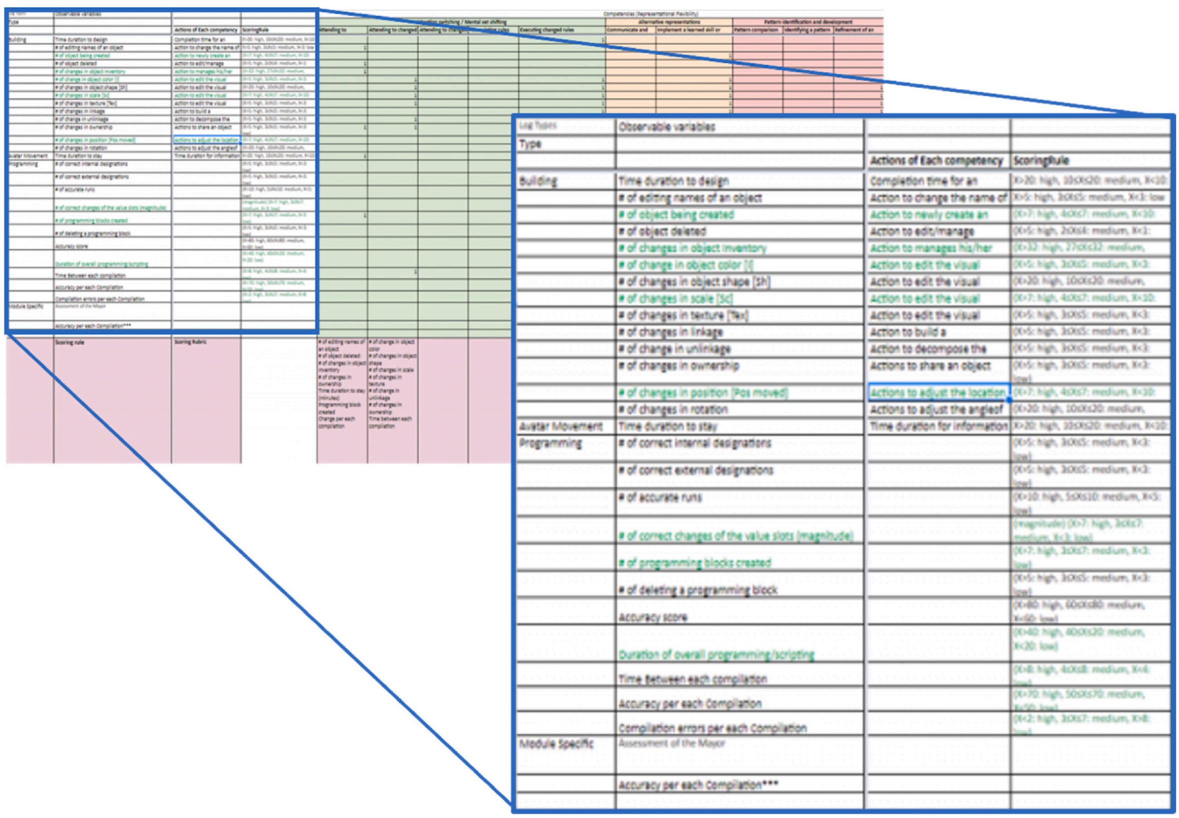

Fig. 3. A Q-matrix for log data mining.

Table 1
RF facets (Moon et al., 2020).

RF Facets	Definition	Literature
Attention Switching/ Cognitive Shifting (ACS)	A capability to redirect attention or alter actions in response to changing rules and contextual requirements. It involves adaptability in thought and behavior in varying situations.	Reed and McCarthy (2012)
Alternative Representation (AR)	An ability to notice and utilize multiple forms of representation. It highlights the skill to perceive and apply different approaches to represent information or solve problems.	Linebarger and Norton-Meier (2016); Takeuchi and Dadkhahfard (2019)
Pattern Development (PD)	An ability to identify and create a pattern or rule during the in-depth analysis of a design problem. It is indicative of proficiency in discerning underlying structures or regularities within complex information.	MacDonald, Westenskow, Moyer-Packenham, & Child (2018)
Pattern Contextualization (PC)	A RF facet to pinpoint the practical application to adapt and tailor the pattern to fit a new and unique implementation context. It represents the ability to apply established patterns in innovative and context-specific ways.	Yee and Bostic (2014)

Table 2
Positive examples of RF enactment.

RF Facets	Examples
ASC	This ability is indicative of effectively recognizing and understanding design cues specific to a given design challenge. In other words, it is an ability to discern and interpret key elements or indicators.
AR	This facet refers to the capability to express and convey an individual's design concept via various representational formats. It includes students' illustration of design ideas using multiple mediums, such as 3D object manipulations or 2D drawings.
PD	This skill includes conducting a thorough analysis of the interrelationships among different features of a design prototype or artifact. It is indicative of the ability to comprehend and articulate how various elements of a prototype design interact and contribute to the overall system structure.
PC	This facet involves the adaptation and customization of a design pattern or prototype for different design contexts, taking into account various environmental or social factors. It highlights the ability to modify and apply design principles flexibly across diverse scenarios.

*Note. ASC refers to Attention Switching/Cognitive Shifting; AR refers to Alternative Representation, PD refers to Pattern Development, and PC is equal to Pattern Contextualization.

2013). LightSIDE enables researchers to build an NLP-driven prediction model by extracting features from lexical data. We extracted a bag-of-word representation from the study participants' verbal utterance data during the training. With the logged behavior data, we computed the frequency of each behavior feature. Specifically, we built and used a Q-matrix that maps behavior data features by each RF facet as the target competency. A Q-matrix locates VR training tasks and data features on the row, as well as the competencies on the columns. Each cell in the Q-matrix contains 0/1 data, indicating how each task and data features are connected to which competencies. We then determined how to score raw feature data into discrete categories (low, medium, and high), which result in a scoring rule (Shute et al., 2019). Fig. 3 demonstrates how Q-matrix is designed for machine learning model training. Based on the ground truth-labeling results, we developed an independent machine-learning classifier with unimodal data input (i.e., speech and behavior data) to generate the estimated probability scores of RF facet development.

2.6. Data fusion

We employed a multimodal data fusion approach to monitor autistic adolescents' RF development by integrating the two distinct data sources previously mentioned. Fig. 4 illustrates how the data fusion approach was designed and applied in this study. We applied decision-level fusion (Meng et al., 2020) to merge classification results from independent data sources. By decision-level fusion, we conducted independent classification with each data source's estimated probability and then computes final decisions after fusing the probability scores of the individual classification results, which concatenates features from different modalities, on which the classification models are trained. In our study, we employed classifiers based on confidence-based fusion (Alam et al., 2015). This approach involved averaging confidence scores derived from each respective data source at a sub-decision level. We established a predefined threshold condition, set at above 50% to ascertain the final result of the classification in the machine learning model.

2.7. Feature engineering

We explored and tested multiple methodological characteristics to fine-tune the classifier settings, including those for data imputation and splitting in model training and testing. We used the k-nearest neighbor algorithm (kNN) for data imputation. Data imputation was used to computationally determine and replace the missing data in the data set for the model prediction.

For the training of our model, we selected a data distribution ratio, allocating 80% for training purposes and reserving 20% for testing. We also used tree-based pruning for the random forest (RF) classifier. We performed pruning to reduce the volume of the decision trees and reduce the risk of overfitting (i.e., a problem in which a model fits the training data very well but does not generalize to unseen test data). During the data fusion process, we identified a potential issue of class imbalance within the dataset used for model training. Data imbalance is a typical problem when the class results are not equally distributed during the model training. We used a computational technique called synthetic minority over-sampling technique (SMOTE) (Cenggoro et al., 2017; Chawla et al., 2002) to alleviate a potential concern about the class-imbalanced data and carry out purposeful data resampling by generating sets of data of minority class.

2.8. Classification algorithm selection

The present study used multi-label classification, a method where multiple outcomes are predicted from a single dataset, as discussed by Tsoumakas and Katakis (2007). This approach aims to assign several

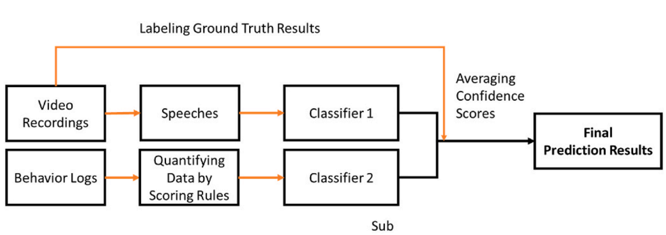

Fig. 4. Multimodal data fusion implementation.

potentially overlapping class labels to data. Multi-label classification allows for a range of classifications, from none to several, within individual data instances. Our evaluation included various candidate algorithms, including the support vector machine (SVM), decision tree (DT), random forest (RForest), and Naïve Bayes (NB). The SVM is a type of supervised learning model that creates a hyperplane in a multidimensional space to segregate classes. It classifies samples based on which side of the hyperplane they are located, aiming to maximize the distance between the division of classes. Despite SVM's performance in high-dimensional spaces on its precision in binary classification tasks, its performance in multi-label contexts and with large datasets can sometimes be less efficient. The DT is a visual method where each internal node denotes a test on a feature, divides symbolize test outcomes, and leaf nodes represent class labels, with the path from root to leaf signifying classification rules. The RForest, an ensemble learning method, generates multiple decision trees from the training data using bootstrapping and classifies samples based on the majority vote from these trees. NB features fast and simple, often underperforms in comparison to more complex models like RForest when dealing with dependency among variables. After exploring these algorithms, we observed that the RForest classifier provided the most accurate predictions. In particular, the robustness of the RForest algorithm was notable. This is largely due to its mechanism of constructing multiple decision trees and using their aggregate predictions, which naturally mitigates the risk of overfitting associated with any single tree. In addition, the RForest's adeptness at handling multi-label datasets was advantageous for our research. Autistic adolescents' RF development is a multifaceted phenomenon that is not clearly captured by mere binary classifications. The RForest algorithm's inherent design allows it to effectively manage multiple labels by assessing the output of individual trees, making it well-suited to our study's needs Consequently, we chose to use a decision-level fusion classification approach utilizing the RF algorithm.

During the data analysis phase, several challenges and limitations emerged, notably the issue of class imbalance within the dataset and the risk of overfitting with the RF classifier. As we noted above, the class imbalance problem, where some classes were underrepresented in the training dataset, posed a significant risk of biased model predictions. We addressed this challenge by implementing the Synthetic Minority Oversampling Technique (SMOTE), which helped to equalize the class distribution by generating synthetic examples of the minority class, thus providing a more balanced dataset for model training. To mitigate the risk of overfitting—a common concern with complex models like RForest that might perform exceedingly well on training data but poorly on unseen data—we employed tree-based pruning techniques. Pruning helped in reducing the size of the decision trees within the RForest, simplifying the model without sacrificing its predictive power, and enhancing its generalizability to new, unseen data. These methodological adjustments ensured that our data fusion approach remained robust and effective, capable of accurately reflecting the RF development in autistic adolescents through the integrated analysis of multimodal data sources.

2.9. Performance evaluation metrics

To determine whether the integration of multimodal data enhanced the accuracy of predictions, we conducted a comparative analysis between the prediction model outcomes derived from the individual data sources and those from the combined, or fused, data source. We employed several key metrics to evaluate the performance of the classifier: Area Under the Curve (AUC), precision, and the F1 Score. The AUC metric is insightful because it evaluates the balance between the true positive rate and the false-positive rate, proving especially beneficial when dealing with imbalanced data distributions. Precision is used to assess the proportion of correctly predicted positive observations to the total predicted positives. The F1 Score, on the other hand, represents the harmonic mean of precision and recall, providing a balanced measure of the classifier's accuracy.

3. Results

3.1. Classifier parameter setting

We configured multiple parameter settings to yield the best prediction performance of the classifiers. We used shuffled random sampling, which divides the data set randomly, and then ran tests on different parts of the data set. This sampling approach is generally conducted to avoid bias caused by an imbalanced data set (Gupta, Varshney, Sharma, Pachauri, & Verma, 2022). In addition, we optimized data imputation and *SMOTE*. During model training, we applied *kNN* for data imputation to replace missing data, whereas *SMOTE* was used to control data imbalance issues. These methodological choices underscore our attempt to ensuring the accuracy and generalizability of our findings across diverse learning scenarios within the VR environment. Table 3 presents the parameters and settings used to render classifiers in this study.

3.2. Classification performance results

Table 4 displays the overall prediction performance results of unimodal classifiers (i.e., only speech or behavior data-based classification) and the fused one. The performance metric scores in bold indicate the best scores for positive and negative labels across training modules, reflecting our approach's nuanced understanding of different aspects of representational flexibility. In general, the fused classifier yielded the best performance results on most performance metrics, illustrating the advantages of multimodal data fusion in accurately assessing and tracking the development of representational flexibility in autistic adolescents. The fused classifier had the best AUC, precision, and F1 scores and in tracking most RF facets. Specifically, the fused classifier had higher overall prediction performance scores (overall AUC = 0.780, Precision = 0.730, F1 Score = 0.675) on ASC, AR, and PD.

There are a few different patterns of prediction performance across the two training modules. The detailed analysis of these patterns reveals the complementary strengths of unimodal and multimodal approaches in understanding the nuanced dynamics of RF development in VR-based training. The fused classifier yielded the best AUC performance on the facet of pattern development in the elevation bridge module, whereas the classifier with the speech data yielded the best performance on pattern development during the NPC design module. Such differential performance highlights the contextual sensitivity of our assessment tools, illustrating our methodology's precision in identifying RF development nuances. In addition, although the classifier with the behavior data earned the best prediction performance on pattern contextualization during the NPC design module (AUC = 0.916, precision = 0.929, and F1 score = 0.865), this classifier appeared to show poor prediction results on other RF facets in the same module. Interestingly, the fused classifier, after synthesizing the two different data inputs, failed to yield the best prediction on pattern contextualization during the NPC design module.

On average, the fused classifiers had fair performance results (AUC >0.70) when tracking all RF facets, demonstrating the efficacy of multimodal data fusion in providing a balanced and comprehensive assessment of RF development. This balanced performance across different facets and modules directly responds to our research question, confirming the effectiveness of data mining techniques, particularly multimodal data fusion, in tracking and assessing RF development in autistic adolescents. In comparison, the unimodal classifier with the behavior data tends to show lower prediction performance (lowest AUC scores) on most RF facets. Among the performance metric results, interestingly, the fused classifier yielded the best performance in terms of the AUC and precision scores. Given that the negative occurrences of RF facets were minority classes in the current dataset, high precision scores on the fused classifier suggest that the proposed fused classifier is satisfactory in detecting learners' minority class results. These findings not only support the efficacy of our methodological approach but also values future approaches to deploying personalized learning interventions within VR environments.

Table 3
Key parameters and their settings for classifier development.

Purpose	Parameter Type	Details
Data sampling	Shuffled random sampling	• Training (80%) and testing data (20%)
Data imputation algorithm	k-nearest neighbors (kNN)	• Number of k = 5 • Mixed measures = Mixed Euclidean distance
Classification algorithm	Random forest	• Number of trees = 100 • Criterion = Gain ratio (Effective in choosing the most informative features for splitting data at each node) • Maximal depth = 10 (Prevent overfitting) • Pruning (confidence = 30%, simplifying the model and potentially improving its generalizability) • Voting = majority voting
Data resampling technique for imbalanced data	Synthetic minority oversampling (SMOTE)	• Normalization • Number of neighbors = 5 • Nominal change rate = 50%

4. Discussion

4.1. Enhancing predictive accuracy via multimodal data fusion

This study's findings not only reaffirm but also extend the understanding of the superior capability of classifiers that incorporate multimodal data fusion, compared to those with single modality (either speech or computer-generated logs), emphasizing the nuanced benefits of multimodal approaches in various contexts. Our detailed analysis of the methodological integration of various data sources at the decision level enhances the accuracy of tracking specific RF facets (i.e., ASC and PD). Moreover, while the multimodal approach did not surpass the effectiveness of speech data in detecting AR, it showed a marked improvement over classifiers based only on behavioral data. The fused classifiers achieved commendable AUC scores (higher than 0.70) especially in tracking most RF facets (ASC, PD, and PC). This comprehensive performance underscores the robustness of our multimodal approach in diverse learning scenarios within VR environments, offering a nuanced perspective on the adaptive challenges and opportunities presented by RF facets. The precision scores, which reflect the accuracy of minority class predictions—were notable for the fused classifiers particularly in identifying cases of PC. This result is indicative of the fact that the model's robustness in providing timely and adaptive learning support for the specific challenges by certain facets during VR-based training. These outcomes strongly support the integration of multiple data streams to improve the accuracy of predictive models in evaluating student competencies, especially within VR-based and immersive learning environments (Chango et al., 2022; Mu et al., 2020; Zhang et al., 2017).

Our proposed data mining approach addressed the limitations of prior approaches by incorporating continuous, in-process learning assessments (Zahabi & Abdul Razak, 2020). By monitoring students' interactions, choices, and problem-solving strategies within the VR environment in real time, we aim to map learning trajectories (Camilleri et al., 2013; Moon et al., 2020; Prakash & Rajendran, 2022). Such detailed insights into the learning process enable the adaptation of VR content and instructional approaches to effectively meet learners' needs. Furthermore, by employing a more comprehensive data analytics approach, we gain deeper insights into learning behaviors beyond

Table 4
Prediction performance results.

RF	Module	P/N	Speech Data Only			Log Data only			Fused		
			AUC	Precision	F1 Score	AUC	Precision	F1 Score	AUC	Precision	F1 Score
ASC	Bridge	P	0.500	0.501	0.660	0.517	0.250	0.003	**0.741**	**0.698**	**0.612**
		N	0.539	0.463	0.633	0.500	UNK	UNK	**0.931**	**0.886**	**0.876**
		Avg	0.520	0.482	0.647	0.508	0.250	0.003	**0.836**	**0.792**	**0.744**
	NPC	P	0.580	0.850	0.227	0.535	0.714	0.152	**0.675**	**0.722**	**0.467**
		N	0.683	0.518	0.670	0.566	0.833	0.588	**0.770**	**0.612**	**0.744**
		Avg	0.632	0.684	0.449	0.551	0.774	0.370	**0.723**	**0.667**	**0.606**
		Overall	0.576	0.583	0.548	0.530	0.512	0.187	**0.780**	**0.730**	**0.675**
AR	Bridge	P	0.544	0.500	0.002	0.554	**0.962**	0.195	**0.789**	0.704	0.664
		N	0.843	0.497	0.664	0.514	0.545	0.705	**0.800**	**0.624**	**0.655**
		Avg	0.694	0.499	0.333	0.534	0.754	0.450	**0.794**	**0.664**	**0.659**
	NPC	P	**0.676**	0.964	0.319	0.506	**0.966**	**0.979**	0.621	0.961	0.289
		N	0.663	0.657	0.632	0.548	0.669	0.801	**0.665**	**0.711**	**0.816**
		Avg	**0.669**	0.810	0.475	0.527	0.818	0.890	0.643	**0.836**	**0.553**
		Overall	0.681	0.654	0.404	0.531	0.786	0.670	**0.718**	**0.755**	**0.606**
PD	Bridge	P	0.615	0.559	0.674	0.538	**0.846**	0.167	**0.690**	0.598	**0.667**
		N	0.632	UNK	UNK	0.553	**0.971**	0.205	**0.886**	0.608	**0.725**
		Avg	0.624	0.559	0.674	0.546	0.909	0.186	**0.788**	**0.603**	**0.696**
	NPC	P	**0.862**	0.784	0.719	0.605	0.585	0.565	0.834	**0.810**	**0.797**
		N	**0.763**	0.510	0.671	0.590	**0.531**	0.610	0.525	0.522	**0.682**
		Avg	**0.813**	0.647	0.695	0.598	0.558	0.588	0.680	0.666	0.740
		Overall	0.718	0.603	0.684	0.572	0.734	0.387	**0.734**	**0.635**	**0.718**
PC	Bridge	P	0.770	**0.781**	0.575	0.516	0.494	0.660	**0.817**	0.718	**0.782**
		N	0.853	0.815	**0.851**	UNK	UNK	UNK	**0.878**	**0.838**	0.842
		Avg	**0.832**	0.798	0.713	0.516	0.494	0.660	**0.848**	0.778	**0.812**
	NPC	P	0.666	**0.916**	0.144	**0.861**	0.906	**0.765**	0.600	0.869	0.217
		N	0.786	0.775	0.873	**0.971**	**0.951**	**0.964**	0.853	0.705	0.760
		Avg	0.726	0.846	0.509	**0.916**	**0.929**	**0.865**	0.727	0.787	0.489
		Overall	**0.794**	0.822	0.611	0.716	0.712	0.763	0.787	0.783	0.651

Notes. RF = Representational flexibility, ASC = Attention switching/Cognitive shifting, AR = Alternative representation, PD = Pattern development, PC = Pattern contextualization, Bridge = Elevation bridge design module, NPC=NPC design module, P = Positive occurrence of RF, N = Negative occurrence of RF, UNK = Unknown, Avg = Average prediction performance score between positive and negative occurrences. Overall = average prediction performance score from both training modules.

conventional assessment methods (Bienkowski et al., 2012; Blikstein & Worsley, 2016). This includes the ability to identify common misconceptions, to assess the effectiveness of specific VR instructional design, and to predict learning outcomes based on observed in-process behaviors (Fernández-Gallego et al., 2013; Kazanidis et al., 2021).

4.2. Effect of virtual reality-based training on representational flexibility development

Recent research has demonstrated the beneficial impact of Virtual Reality (VR)-based training on enhancing cognitive abilities in individuals with autism. Notably, the work of Zhao et al. (2021; 2022) revealed that VR-based cognitive training significantly boosts developmental capabilities in this demographic. Further supporting these findings, studies by various authors (Ke et al., 2020; Moon et al., 2020; Sokolikj et al., 2023), have demonstrated that VR training not only augments cognitive skills but also fosters social skills development among autistic learners. A series of investigations (de Moraes et al., 2020; Drigas, Mitsea, & Skianis, 2022) have underscored the efficacy of VR to enhance both cognitive and metacognitive abilities.

This current research documents the effectiveness of VR-based training on RF development through behavior observation records, providing new insights into the unique contributions of VR to fostering representational flexibility. The findings of this study support prior research on cognitive skill enhancement and broaden the conversation about how VR-based tasks contribute to RF development. The tasks involved require participants to tackle complex, sequential problems, such as designing elevated bridges and creating interactive NPCs, challenging them to think abstractly and concretely, thereby enhancing their comprehension and manipulation of real-world and theoretical concepts. For example, the bridge design task requires understanding physical principles and considering environmental factors, prompting participants to mentally simulate and refine their designs. Furthermore, scripting tasks that control bridges and NPCs offer participants the opportunity to convert intricate, real-life problems into executable, computer-interpretable commands. While the beneficial outcomes of VR-based training have been noted, it is important to acknowledge that this study was not conducted within a randomized controlled trial framework. Therefore, future research should aim to investigate the impact of VR-based training under more controlled experimental conditions.

4.3. Multimodal data mining in virtual reality-based training

This methodological research sought to apply multimodal data fusion in VR-based learning environments, an area still nascent and under resolved in current MmLA research, which has predominantly focused on physical learning settings (Chan et al., 2020; Giannakos & Cukurova, 2023). It highlights the effective integration of multimodal data fusion techniques in dynamic and highly interactive VR and Metaverse environments, suggesting that such approaches can effectively improve MmLA performance in VR and immersive learning environment systems.

The study navigates challenges such as missing data and class imbalance using strategies like data imputation and the Synthetic Minority Over-sampling Technique (SMOTE). These approaches are particularly relevant given the diverse nature of VR training modules and their impact on different RF facets. For instance, in design challenges like elevation bridge design, autistic adolescents demonstrated more instances of AR, as compared to other RF facets (ASC, PD, PC) observed in modules like NPC design. To counter the infrequency of negative RF facet cases and ensure balanced class representation, deliberate resampling was implemented, underscoring the strategic significance of such feature engineering methods in enhancing the fidelity and inclusivity of assessments within VR-based educational frameworks. Particularly, the application of these techniques proves beneficial for more accurate data tracking in neurodiverse student populations (Plunk et al., 2022; Sewell, 2022), which are generally smaller than the population of typically developing individuals. When data analytics and model development are applied to neurodiverse populations, issues of imbalanced data often arise, making it crucial to optimize prediction models for limited data scenarios (e.g., few-shot or zero-shot learning). In this regard, the findings of the study could contribute to data optimization in diverse data collection contexts.

Lastly, the results offer insights into refining MmLA systems for VR-based training. This nuanced understanding facilitates a deeper exploration into the alignment of data-driven insights with pedagogical practices, ideally shaping more responsive and effective VR learning environments. While the fused data approach did not yield the highest accuracy in tracking PC, it was found that classifiers using speech data were more effective in this respect. This observation indicates a need for further refinement in data collection and fusion techniques to track various competency facets more accurately in VR learning environments.

4.4. Inclusive design in multimodal learning analytics

The study's findings highlight the critical role of inclusive design principles in the development of multimodal learning analytics (MmLA) systems within immersive learning environments (Tamura et al., 2019; Martínez-Monés et al., 2019). Inclusive design of MmLA should aim at creating systems that are accessible and accommodating to a wide range of users, including those with diverse neurological profiles. This approach ensures that the technology is not just for the majority but is sensitive to the needs of all users, especially those who might interact with technology differently. This inclusive approach is essential, as it not only democratizes technology access but enriches learning experiences by recognizing and valuing the diverse ways in which individuals engage with and process information.

Using post-hoc behavior analysis of video recordings that captured VR training sessions, we found a significant variability observed in how quickly learners adapted to the VR interface and controls. Some study participants became proficient in 3D design and scripting rapidly, enhancing their task performance, while others remained slow, affecting their overall task efficiency. It echoes the need for adaptive interface design within a VR training system, where the complexity and delivery of instruction can be adjusted based on the learner's proficiency. Adaptive tutorials and practice sessions need to be offered to those struggling with the interface, ensuring that technical barriers do not impede learning. As another example found in behavior observations, in design problem-solving scenarios, the way learners responded to feedback varied widely. Some immediately applied feedback to refine their solutions, while others either ignored feedback or were unable to apply it effectively to improve their outcomes. Emphasizes the need for personalized feedback mechanisms in VR training that consider the learner's receptivity and ability to apply feedback. For learners who struggle to incorporate feedback, additional guidance or alternative explanations could be provided to help bridge the gap between understanding and application. Specifically, neurodiverse learners may have different cognitive processes that affect how they solve problems, understand concepts, and apply feedback. Multiple representation-supporting learning supports can cater to these differences by offering various ways to approach and solve problems. For instance, some learners might benefit from animated simulations, while others might prefer abstractive and symbolic information when identifying design-related clues and aligned knowledge.

The current findings also revealed the importance of various data representation. In our VR training module, we ensure the data we collected and analyzed represents a broad spectrum of the autistic experience. We accordingly presented a variety of design quests, each predominantly necessitating distinct types of learning-related actions (evidence-centered design), along with their respective data channels, which were then integrated into the assessment system. This approach allowed for a diverse and comprehensive evaluation of the learning process, ensuring that various aspects of student interaction and performance were captured and analyzed effectively. We believe that our approach can contribute to reducing data biases from the singular data channels and develop a more inclusive learning system.

5. Limitation and future work

The present study encounters several limitations that need improvement in future research. First, the prediction models employed exhibit limited effectiveness in consistently tracking the development of Representational Flexibility (RF) among autistic adolescents. Despite employing data resampling techniques to mitigate biases, difficulties persisted in accurately assessing the classifier's ability to detect less frequent RF occurrences. This issue was particularly pronounced with behavior data during the bridge design module, suggesting that these data were not comprehensively captured. The unique characteristics of this module might have contributed to ambiguous performance metrics for specific RF dimensions. To address these limitations, future research should focus on collecting a broader array of speech and behavioral data from autistic adolescents. Such efforts would not only improve the model's accuracy but also extend its applicability across different contexts, thereby enhancing our understanding of RF development in this demographic.

Second, our study utilized a decision-level fusion approach for synthesizing data features, which, while accurate, tends to be less efficient due to the need for processing multiple data characteristics, resulting in slower performance. Moreover, decision-level fusion might not be effective if one type of data is missing. Future research should explore feature-level fusion, which integrates multiple data characteristics before classifier construction, to compare the efficacy of both fusion methods in tracking RF in VR-based training. An additional limitation emerged from the observation that the fused classifier was not consistently superior in tracking all aspects of RF, particularly less effective than speech data alone in identifying Alternative Representation (AR). This result suggests inherent limitations in the fused classifier's ability to leverage multimodal data for specific RF facets. The discrepancy highlights the complexity of effectively integrating diverse data types and the potential for certain modalities to dilute or obscure relevant signals. Future research should delve into optimizing data fusion techniques, possibly through advanced algorithms that can more accurately maintain the integrity of each data type's contribution to understanding RF.

Third, the scope of our study was constrained to speech data, including lexical features and behavior logs, which may not comprehensively capture the RF developmental state of autistic adolescents. To improve the multimodal data fusion approach within the practical confines of VR environments, future research should consider integrating VR-feasible equipment that is less intrusive and more aligned with the realities of deploying such extended reality technologies in educational settings. Specifically, incorporating eye-tracking technology integrated within VR headsets could provide insightful data on participants' focus, attention switching, and engagement without the need for additional cumbersome equipment. This approach would allow for a more nuanced understanding of RF facets, particularly those related to visual attention and cognitive processing in a way that both feasible and minimally disruptive to the immersive experience of VR-based training can be applied. By leveraging such VR-compatible technologies, future studies could significantly enhance the richness of the dataset, offering a comprehensive picture of autistic adolescents' learning processes and their interaction with virtual learning environments.

Last, our analysis overlooked the temporal dynamics inherent in the data, a critical aspect considering the evolving nature of RF development during VR-based training. The data's sequential structure may include valuable insights into how autistic adolescents' engagement and cognitive strategies dynamically change over time. Future studies should explore the application of deep learning models, like recurrent neural networks (RNNs) or Long Short-Term Memory network (LSTM), renowned for their proficiency in handling with timeseries data. We believe that these models can complicate the temporal progression of learners' responses, providing a granular view of the changes in RF across various stages of the VR experience. By integrating these advanced computational techniques, future research could achieve a dynamic and nuanced portrayal of the learning trajectories among autistic adolescents, facilitating the creation of more adaptive and responsive VR-based educational interventions.

Funding

This research was supported by National Science Foundation (Award number: 1837917)

CRediT authorship contribution statement

Jewoong Moon: Writing – review & editing, Writing – original draft, Supervision, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Fengfeng Ke:** Writing – review & editing, Writing – original draft, Funding acquisition, Conceptualization. **Zlatko Sokolikj:** Validation, Software, Methodology, Investigation. **Shayok Chakraborty:** Writing – review & editing, Supervision, Methodology, Conceptualization.

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.

References

Acevedo Nistal, A., Van Dooren, W., Clarebout, G., Elen, J., & Verschaffel, L. (2009). Conceptualising, investigating and stimulating representational flexibility in mathematical problem solving and learning: A critical review. *ZDM, 41*(5), 627–636. https://doi.org/10.1007/s11858-009-0189-1

Acevedo Nistal, A., Van Dooren, W., & Verschaffel, L. (2012). What counts as a flexible representational choice? An evaluation of students' representational choices to solve linear function problems. *Instructional Science, 40*, 999–1019. https://doi.org/10.1007/s11251-011-9199-9

Akhuseyinoglu, K., & Brusilovsky, P. (2021). Data-driven modeling of learners' individual differences for predicting engagement and success in online learning. June. In *Proceedings of the 29th ACM conference on user modeling, adaptation and personalization* (pp. 201–212).

Alam, M. R., Bennamoun, M., Togneri, R., & Sohel, F. (2015). A confidence-based late fusion framework for audio-visual biometric identification. *Pattern Recognition Letters, 52*, 65–71. https://doi.org/10.1016/j.patrec.2014.10.006

Anastopoulou, S., Sharples, M., & Baber, C. (2011). An evaluation of multimodal interactions with technology while learning science concepts. *British Journal of Educational Technology, 42*(2), 266–290. https://doi.org/10.1111/j.1467-8535.2009.01017.x

Babu, S. K., Krishna, S., Unnikrishnan, R., & Bhavani, R. R. (2018). Virtual reality learning environments for vocational education: A comparison study with conventional instructional media on knowledge retention. July. In *2018 IEEE 18th international conference on advanced learning technologies (ICALT)* (pp. 385–389). IEEE.

Bienkowski, M., Feng, M., & Means, B. (2012). *Enhancing teaching and learning through educational data mining and learning analytics: An issue brief. Office of educational technology*. US Department of Education.

Blikstein, P., & Worsley, M. (2016). Multimodal learning analytics and education data mining: Using computational technologies to measure complex learning tasks. *Journal of Learning Analytics, 3*(2), 220–238. https://doi.org/10.18608/jla.2016.32.11

Block, D. (1998). Exploring interpretations of questionnaire items. *System, 26*(3), 403–425. https://doi.org/10.1016/S0346-251X(98)00022-0

Brunyé, T. T., Rapp, D. N., & Taylor, H. A. (2008). Representational flexibility and specificity following spatial descriptions of real-world environments. *Cognition, 108*(2), 418–443.

Camilleri, V., de Freitas, S., Montebello, M., & McDonagh-Smith, P. (2013). A case study inside virtual worlds: Use of analytics for immersive spaces. April. In *Proceedings of the third international conference on learning analytics and knowledge* (pp. 230–234).

Cenggoro, T. W., Isa, S. M., Kusuma, G. P., & Pardamean, B. (2017). Classification of imbalanced land-use/land-cover data using variational semi-supervised learning. November. In *2017 international conference on innovative and creative information technology (ICITech)* (pp. 1–6). IEEE. https://doi.org/10.1109/INNOCIT.2017.8319149.

Chan, M. C. E., Ochoa, X., & Clarke, D. (2020). Multimodal learning analytics in a laboratory classroom. *Machine LearningParadigms: Advances in Learning Analytics*, 131–156.

Chango, W., Lara, J. A., Cerezo, R., & Romero, C. (2022). A review on data fusion in multimodal learning analytics and educational data mining. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 12*(4), e1458. https://doi.org/10.1002/widm.1458

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). Smote: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research, 16*, 321–357. https://doi.org/10.1613/jair.953

Chen, F., Cui, Y., & Chu, M. W. (2020). Utilizing game analytics to inform and validate digital game-based assessment with evidence-centered game design: A case study. *International Journal of Artificial Intelligence in Education, 30*(3), 481–503. https://doi.org/10.1007/s40593-020-00202-6

Dahlstrom-Hakki, I., Alstad, Z., Asbell-Clarke, J., & Edwards, T. (2024). The impact of visual and auditory distractions on the performance of neurodiverse students in virtual reality (VR) environments. *Virtual Reality, 28*(1), 29. https://doi.org/10.1007/s10055-023-00933-6

Dahlstrom-Hakki, I., Edwards, T., Larsen, J., Alstad, Z., Belton, G., Lougen, D., & Santana, D. (2021). Inclusive VR through inclusive co-design with neurodiverse learners. May. In *2021 7th international conference of the immersive learning research network (iLRN)* (pp. 1–5). Eureka, CA, USA: IEEE. https://doi.org/10.23919/iLRN52045.2021.9459322.

de Moraes, Í. A. P., Monteiro, C. B. D. M., Silva, T. D. D., Massetti, T., Crocetta, T. B., de Menezes, L. D. C., … Magalhaes, F. H. (2020). Motor learning and transfer between real and virtual environments in young people with autism spectrum disorder: A prospective randomized cross over controlled trial. *Autism Research, 13*(2), 307–319. https://doi.org/10.1002/aur.2208

De Vries, M., & Geurts, H. (2015). Influence of autism traits and executive functioning on quality of life in children with an autism spectrum disorder. *Journal of Autism and Developmental Disorders, 45*(9), 2734–2743. https://doi.org/10.1007/s10803-015-2438-1

Drey, T., Jansen, P., Fischbach, F., Frommel, J., & Rukzio, E. (2020). Towards progress assessment for adaptive hints in educational virtual reality games. April. In *Extended abstracts of the 2020 CHI conference on human factors in computing systems* (pp. 1–9).

Drigas, A., Mitsea, E., & Skianis, C. (2022). Virtual reality and metacognition training techniques for learning disabilities. *Sustainability, 14*(16), Article 10170. https://doi.org/10.3390/su141610170

Ehsan, H., Rispoli, M., Lory, C., & Gregori, E. (2018). A systematic review of STEM instruction with students with autism spectrum disorders. *Review Journal of Autism and Developmental Disorders, 5*(4), 327–348. https://doi.org/10.1007/s40489-018-0142-8

Farrar, M. J., & Ashwell, S. (2008). The role of representational ability in the development of phonological awareness inpreschool children. In K. B. Cartwright (Ed.), *Literacy processes: Cognitive flexibility in learning and teaching* (pp. 71–89). New York, NY: Guilford Press.

Fernández-Gallego, B., Lama, M., Vidal, J. C., & Mucientes, M. (2013). Learning analytics framework for educational virtual worlds. *Procedia Computer Science, 25*, 443–447. https://doi.org/10.1016/j.procs.2013.11.056

Fogarty, J., McCormick, J., & El-Tawil, S. (2018). Improving student understanding of complex spatial arrangements with virtual reality. *Journal of Professional Issues in Engineering Education and Practice, 144*(2), Article 04017013. https://doi.org/10.1061/(ASCE)EI.1943-5541.0000349

Gagatsis, A., Elia, I., & Mousoulides, N. (2006). Are registers of representations and problem solving processes on functions compartmentalized in students' thinking? *Revista Latinoamericana de Investigación en Matemática Educativa, RELIME, (Esp)*, 197–224.

Gaw, N., Yousefi, S., & Gahrooei, M. R. (2022). Multimodal data fusion for systems improvement: A review. *IISE Transactions, 54*(11), 1098–1116. https://doi.org/10.1080/24725854.2021.1987593

Giannakos, M., & Cukurova, M. (2023). The role of learning theory in multimodal learning analytics. *British Journal of Educational Technology, 54*(5), 1246–1267.

Gupta, H., Varshney, H., Sharma, T. K., Pachauri, N., & Verma, O. P. (2022). Comparative performance analysis of quantum machine learning with deep learning for diabetes prediction. *Complex & Intelligent Systems, 8*(4), 3073–3087. https://doi.org/10.1007/s40747-021-00398-7

Hadgraft, R. G., & Kolmos, A. (2020). Emerging learning environments in engineering education. *Australasian Journal of Engineering Education, 25*(1), 3–16. https://doi.org/10.1080/22054952.2020.1713522

Heba, E. D., Mansour, N., Alzaghibi, M., & Alhammad, K. (2017). Context of STEM integration in schools: Views from in-service science teachers. *Eurasia Journal of Mathematics, Science and Technology Education, 13*(6), 2459–2484. https://doi.org/10.12973/eurasia.2017.01235a

Heil, D. R., Pearson, G., & Burger, S. E. (2013). Understanding integrated stem education: Report on a national study. June. In *2013 ASEE annual conference & exposition* (pp. 23–1279).

Henderson, N., Rowe, J., Paquette, L., Baker, R. S., & Lester, J. (2020). Improving affect detection in game-based learning with multimodal data fusion. July. In *International conference on artificial intelligence in education* (pp. 228–239). Cham: Springer. https://doi.org/10.1007/978-3-030-52237-7_19.

Herbert, J., & Hayne, H. (2000). Memory retrieval by 18–30-month-olds: Age-related changes in representational flexibility. *Developmental Psychology, 36*(4), 473–484. https://psycnet.apa.org/doi/10.1037/0012-1649.36.4.473.

Hill, E., Berthoz, S., & Frith, U. (2004). Brief report: Cognitive processing of own emotions in individuals with autistic spectrum disorder and in their relatives. *Journal of Autism and Developmental Disorders, 34*, 229–235. https://doi.org/10.1023/b:jadd.0000022613.41399.14

Hofmann, M., & Klinkenberg, R. (Eds.). (2016). *RapidMiner: Data mining use cases and business analytics applications.* CRC Press.

Ionescu, T. (2012). Exploring the nature of cognitive flexibility. *New Ideas in Psychology, 30*(2), 190–200. https://psycnet.apa.org/doi/10.1016/j.newideapsych.2011.11.001.

Jin, Q., Liu, Y., Yarosh, S., Han, B., & Qian, F. (2022). How will VR enter university classrooms? Multi-stakeholders investigation of VR in higher education. April. In *Proceedings of the 2022 CHI conference on human factors in computing systems* (pp. 1–17).

Kazanidis, I., Pellas, N., & Christopoulos, A. (2021). A learning analytics conceptual framework for augmented reality-supported educational case studies. *Multimodal Technologies and Interaction, 5*(3), 9. https://doi.org/10.3390/mti5030009

Ke, F., & Lee, S. (2016). Virtual reality based collaborative design by children with high-functioning autism: Design-based flexibility, identity, and norm construction. *Interactive Learning Environments, 24*(7), 1511–1533. https://doi.org/10.1080/10494820.2015.1040421

Ke, F., Moon, J., & Sokolikj, Z. (2020). Tracking representational flexibility development through speech data mining. October. In *2020 IEEE frontiers in education conference (FIE)* (pp. 1–4). IEEE. https://doi.org/10.1109/FIE44824.2020.9273818.

Linebarger, D. L., & Norton-Meier, L. (2016). Scientific concepts, multiple modalities, and young children. In *Using multimodal representations to support learning in the science classroom* (pp. 97–116). Cham: Springer.

Lorenzo, G., Lledó, A., Arráez-Vera, G., & Lorenzo-Lledó, A. (2019). The application of immersive virtual reality for students with ASD: A review between 1990–2017. *Education and Information Technologies, 24*, 127–151. https://doi.org/10.1007/s10639-018-9766-7

Lorenzo, G. G., Newbutt, N. N., & Lorenzo-Lledó, A. A. (2023). Designing virtual reality tools for students with autism spectrum disorder: A systematic review. *Education and Information Technologies*, 1–49. https://doi.org/10.1007/s10639-022-11545-z

MacDonald, B. L., Westenskow, A., Moyer-Packenham, P. S., & Child, B. (2018). Components of place value understanding: Targeting mathematical difficulties when providing interventions. *School Science and Mathematics, 118*(1–2), 17–29.

Makransky, G., & Mayer, R. E. (2022). Benefits of taking a virtual field trip in immersive virtual reality: Evidence for the immersion principle in multimedia learning. *Educational Psychology Review, 34*(3), 1771–1798. https://doi.org/10.1007/s10648-022-09675-4

Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (camil): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review, 33*(3), 937–958. https://doi.org/10.1007/s10648-020-09586-2

Martínez-Monés, A., Villagrá-Sobrino, S., Georgiou, Y., Ioannou, A., & Ruiz, M. J. (2019). The INTELed pedagogical framework: Applying embodied digital apps to support special education children in inclusive educational contexts. June. In *Proceedings of the international Conference on human computer interaction* (pp. 1–4). https://doi.org/10.1145/3335595.3335652. Donostia-San Sebastián (Basque Country, Spain).

Massaro, D. W. (2012). Multimodal learning. *Encyclopedia of the Sciences of Learning*, 2375–2378. https://doi.org/10.1109/TVCG.2020.3010088

Mayfield, E., & Rosé, C. P. (2013). LightSIDE: Open source machine learning for text. In *Handbook of automated essay evaluation* (pp. 146–157). Routledge.

Melo, M., Gonçalves, G., Monteiro, P., Coelho, H., Vasconcelos-Raposo, J., & Bessa, M. (2020). Do multisensory stimuli benefit the virtual reality experience? A systematic review. *IEEE Transactions on Visualization and Computer Graphics, 28*(2), 1428–1442. https://doi.org/10.1109/tvcg.2020.3010088

Meng, T., Jing, X., Yan, Z., & Pedrycz, W. (2020). A survey on machine learning for data fusion. *Information Fusion, 57*, 115–129. https://doi.org/10.1016/j.inffus.2019.12.001

Mentzer, N. (2011). High school engineering and technology education integration through design challenges. *Journal of STEM Teacher Education, 48*(2), 7. https://files.eric.ed.gov/fulltext/EJ952048.pdf.

Messinger, D. S., Duvivier, L. L., Warren, Z. E., Mahoor, M., Baker, J., Warlaumont, A., & Ruvolo, P. (2015). Affective computing, emotional development, and autism. In R. A. Carlvo, S. K. D'Mello, J. Gratch, & A. Kappas (Eds.), *The Oxford handbook of affective computing* (pp. 516–536). Oxford University Press.

Molina-Carmona, R., Pertegal-Felices, M. L., Jimeno-Morenilla, A., & Mora-Mora, H. (2018). Virtual reality learning activities for multimedia students to enhance spatial ability. *Sustainability, 10*(4), 1074. https://doi.org/10.3390/su10041074

Molnar-Szakacs, I., & Heaton, P. (2012). Music: a unique window into the world of autism. *Annals of the New York Academy of Sciences, 1252*(1), 318–324.

Moon, J., Choi, G. W., & Seo, J. (2023). Revisiting multimedia learning design principles in virtual reality-based learning environments for autistic individuals. *Virtual Reality, 27*, 3101–3113. https://doi.org/10.1007/s10055-023-00856-2

Moon, J., Ke, F., & Sokolikj, Z. (2020). Automatic assessment of cognitive and emotional states in virtual reality-based flexibility training for adolescents with autism. *British Journal of Educational Technology, 51*(5), 1766–1784. https://doi.org/10.1111/bjet.13005

Moon, J., Ke, F., Sokolikj, Z., & Dahlstrom-Hakki, I. (2022). Multimodal data fusion to track students' distress during educational gameplay. *Journal of Learning Analytics*, 1–13. https://doi.org/10.18608/jla.2022.7631

Mu, S., Cui, M., & Huang, X. (2020). Multimodal data fusion in learning analytics: A systematic review. *Sensors, 20*(23), 6856. https://doi.org/10.3390/s20236856

Paes, D., Arantes, E., & Irizarry, J. (2017). Immersive environment for improving the understanding of architectural 3D models: Comparing user spatial perception between immersive and traditional virtual reality systems. *Automation in Construction, 84*, 292–303.

Pande, P., & Chandrasekharan, S. (2017). Representational competence: Towards a distributed and embodied cognition account. *Studies in Science Education, 53*(1), 1–43. https://doi.org/10.1080/03057267.2017.1248627

Parsons, S., & Cobb, S. (2011). State-of-the-art of virtual reality technologies for children on the autism spectrum. *European Journal of Special Needs Education, 26*(3), 355–366. https://doi.org/10.1080/08856257.2011.593831

Plunk, A., Amat, A. Z., Wilkes, D. M., & Sarkar, N. (2022). Automated behavior labeling during team-based activities involving neurodiverse and neurotypical partners using multimodal data. August. In *International conference on pattern recognition* (pp. 193–209). Cham: Springer Nature Switzerland.

Prakash, A., & Rajendran, R. (2022). Mechanism to capture learners' interactions in virtual reality learning environment. July. In *2022 international conference on advanced learning technologies (ICALT)* (pp. 335–337). IEEE.

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education, 147*, Article 103778. https://doi.org/10.1016/j.compedu.2019.103778

Reed, P., & McCarthy, J. (2012). Cross-modal attention-switching is impaired in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 42*(6), 947–953. https://doi.org/10.1007/s10803-011-1324-8

Rizzo, A. A., Buckwalter, J. G., McGee, J. S., Bowerly, T., Zaag, C. V. D., Neumann, U., … Chua, C. (2001). Virtual environments for assessing and rehabilitating cognitive/functional performance a review of projects at the USC integrated media systems center. *Presence: Teleoperators and Virtual Environments, 10*(4), 359–374.

Salzman, M. C., Dede, C., Loftin, R. B., & Chen, J. (1999). A model for understanding how virtual reality aids complex conceptual learning. *Presence: Teleoperators and Virtual Environments, 8*(3), 293–316. https://doi.org/10.1162/105474699566242

Sewell, A. (2022). Understanding and supporting learners with specific learning difficulties from a neurodiversity perspective: A narrative synthesis. *British Journal of Special Education, 49*(4), 539–560. https://doi.org/10.1111/1467-8578.12422

Sharma, K., Papamitsiou, Z., & Giannakos, M. (2019). Building pipelines for educational data using AI and multimodal analytics: A "grey-box" approach. *British Journal of Educational Technology, 50*(6), 3004–3031.

Shute, V., Rahimi, S., & Lu, X. (2019). Supporting learning in educational games: Promises and challenges. In *Learning in a digital world* (pp. 59–81). Singapore: Springer.

Sokolikj, Z., Ke, F., Chakraborty, S., & Moon, J. (2023). Using deep learning to track representational flexibility development of students with autism in a nurturing virtual reality environment. In *Proceedings of the 11th international conference on information and education technology (ICIET)* (pp. 51–55). Fujisawa, Japan: IEEE. https://doi.org/10.1109/ICIET56899.2023.10111218.

Spector, J. M. (2006). A methodology for assessing learning in complex and ill-structured task domains. *Innovations in Education & Teaching International, 43*(2), 109–120. https://doi.org/10.1080/14703290600650368

Sveistrup, H. (2004). Motor rehabilitation using virtual reality. *Journal of NeuroEngineering and Rehabilitation, 1*, 1–8. https://doi.org/10.1186/1743-0003-1-10

Takeuchi, M. A., & Dadkhahfard, S. (2019). Rethinking bodies of learners through STEM education. In P. Sengupta, M. C. Shanahan, & B. Kim (Eds.), *Critical, Transdisciplinary and embodied Approaches in STEM education. Advances in STEM education*. Cham: Springer. https://doi.org/10.1007/978-3-030-29489-2_11.

Thibaut, L., Ceuppens, S., De Loof, H., De Meester, J., Goovaerts, L., Struyf, A., … Depaepe, F. (2018). Integrated STEM education: A systematic review of instructional practices in secondary education. *European Journal of STEM Education, 3*(1), 2. https://doi.org/10.20897/ejsteme/85525

Tlili, A., & Chang, M. (2019). Data analytics approaches in educational games and gamification Systems: Summary, challenges, and future insights. In *Data analytics approaches in educational games and gamification systems* (pp. 249–255). Singapore: Springer.

Tamura, K., Lu, M., Konomi, S. I., Hatano, K., Inaba, M., Oi, M., … Yamada, Y. (2019). Integrating multimodal learning analytics and inclusive learning support systems for people of all ages. In *Cross-Cultural Design. Culture and Society: 11th International Conference, CCD 2019, Held as Part of the 21st HCI International Conference, HCII 2019, Orlando, FL, USA, July 26-31, 2019, Proceedings, Part II 21* (pp. 469–481). Springer International Publishing. https://doi.org/10.1007/978-3-030-22580-3_35.

Tsoumakas, G., & Katakis, I. (2007). Multi-label classification: An overview. *International Journal of Data Warehousing and Mining, 3*(3), 1–13. https://doi.org/10.4018/jdwm.2007070101

Webb, N. M. (1980). A process-outcome analysis of learning in group and individual settings. *Educational Psychologist, 15*(2), 69–83. https://psycnet.apa.org/doi/10.1080/00461528009529217.

Worsley, M., & Martinez-Maldonado, R. (2018). Multimodal learning analytics' past, present, and potential futures. March. In *CrossMMLA@ LAK*.

Yee, S. P., & Bostic, J. D. (2014). Developing a contextualization of students' mathematical problem solving. *The Journal of Mathematical Behavior, 36*, 1–19. https://doi.org/10.1016/j.jmathb.2014.08.002

Yu, J., & Denham, A. R. (2023). Utilizing augmented reality for embodied mental rotation training: A learning analytics study. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.13397

Zahabi, M., & Abdul Razak, A. M. (2020). Adaptive virtual reality-based training: A systematic literature review and framework. *Virtual Reality, 24*, 725–752. https://doi.org/10.1007/s10055-020-00434-w

Zhang, M., Ding, H., Naumceska, M., & Zhang, Y. (2022). Virtual reality technology as an educational and intervention tool for children with autism spectrum disorder: Current perspectives and future directions. *Behavioral Sciences, 12*(5), 138. https://doi.org/10.3390/bs12050138

Zhang, L., Wade, J., Bian, D., Fan, J., Swanson, A., Weitlauf, A., … Sarkar, N. (2017). Cognitive load measurement in a virtual reality-based driving system for autism intervention. *IEEE Transactions on Affective Computing, 8*(2), 176–189. https://doi.org/10.1109/TAFFC.2016.2582490

Zhao, J., Zhang, X., Lu, Y., Wu, X., Zhou, F., Yang, S., … Fei, F. (2022). Virtual reality technology enhances the cognitive and social communication of children with autism spectrum disorder. *Frontiers in Public Health, 10*, Article 1029392. https://doi.org/10.3389/fpubh.2022.1029392

Zhao, J. Q., Zhang, X. X., Wang, C. H., & Yang, J. (2021). Effect of cognitive training based on virtual reality on the children with autism spectrum disorder. *Current Research in Behavioral Sciences, 2*, Article 100013. https://doi.org/10.1016/j.crbeha.2020.100013

Jewoong Moon is assistant professor of Instructional Technology at the University of Alabama. His research interests include digital game-based learning and inclusive and immersive e-learning. He is the corresponding author of this manuscript (* contact: jmoon19@ua.edu)

Fengfeng Ke is professor of the Department of Educational Psychology and Learning Systems at Florida State University.

Zlatko Sokolikj is a doctoral candidate at the Department of Scientific Computing at Florida state University.

Shayok Chakraborty is assistant professor of the Department of Computer Science at Florida State University.