Taylor & Francis
Taylor & Francis Group

Australasian Journal of Engineering Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/teen20

How multi-modal approaches support engineering and computing education research

Idalis Villanueva Alarcón, Saira Anwar & Zahra Atiq

To cite this article: Idalis Villanueva Alarcón, Saira Anwar & Zahra Atiq (12 Nov 2023): How multi-modal approaches support engineering and computing education research, Australasian Journal of Engineering Education, DOI: 10.1080/22054952.2023.2274513

To link to this article: https://doi.org/10.1080/22054952.2023.2274513

Published online: 12 Nov 2023.

AUSTRALASIAN JOURNAL OF ENGINEERING EDUCATION
https://doi.org/10.1080/22054952.2023.2274513

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS Check for updates

How multi-modal approaches support engineering and computing education research

Idalis Villanueva Alarcón [a], Saira Anwar [b] and Zahra Atiq [c]

[a]Department of Engineering Education, University of Florida, Gainesville, FL, USA; [b]Department of Multidisciplinary Engineering, Texas A&M University, College Station, TX, USA; [c]Department of Computer Science and Engineering, The Ohio State University, Columbus, OH, USA

ABSTRACT

Multi-modal approaches in engineering and computing education are still in its early stages. With the advent of new technologies and communication platforms, understanding the principles and elements of multi-modal work will help scholars to answer complex research questions related to engineering and computing education. Multi-modal approaches consist of research principles and practices that aim to explore the multi-sensory ways humans experience the complexity and multiplicity of their surrounding world as it happens. This manuscript will elaborate on the principles of multi-modal research, highlight examples in the engineering and computing education literature, and share considerations and strategies. The manuscript's purpose is to guide scholars who wish to capture participant experiences of phenomena naturalistically and authentically, and in near-real-time..

ARTICLE HISTORY
Received 31 August 2022
Accepted 19 October 2023

KEYWORDS
Multi-modal approaches; engineering education; computing education

1. Introduction

Multi-modality, which stems from sociolinguistics, represents multiple modes of communication, such as reading, writing, and oral communication, and how individuals make sense of their surrounding world (e.g. Ledin and Machin 2017). As an individual experiences their surrounding world, they acquire a 'wealth of information to support interaction with the world and with one another' (Turk 2014, 189). Since interactions between the individual and their surrounding world are multi-sensory and intertwined, a multi-modal approach is needed to research and capture these dynamics in near-real-time. Information is collected via a comprehensive understanding of the dynamics of experiences and senses, expressions, interactions, and meanings, as shown in Figure 1. Please note that the stages in this figure will be explained later in the manuscript.

Multi-modal approaches are defined as a set of practices and principles used to explore and understand multiple constructions of people's realities and experiences, including phenomena, contexts, and factors in near-real-time to said event. Although the concept of multi-modality is not new (e.g. Bandura 1969; Bunt, Beun, and Borghuis 1998; Lachs 2017; Lazarus 2005, 2006; Ledin and Machin 2017; Palmer 2006), its application to understand education-related problems is relatively nascent. By combining inter- and multi-disciplinary techniques and methods, and broadening the data collection strategies, multi-modal scholars 'try to collect many simultaneous events to represent the messiness and immediacy of life' (Villanueva Alarcón and Anwar 2022, 277).

Multi-modal approaches are situated within the umbrella term defined by the National Science Foundation (NSF): *convergence research*. Convergence research, one of NSF's Big 10 Ideas (National Science Foundation 2016), combines knowledge, theories, methods, data analysis, and interpretation strategies across multiple disciplines. These combinations result in new research approaches that intertwine disciplinary practices, communications, and cultures (Goundar 2021; Stegmaier 2009). Although not all components of convergence research align with multi-modal research, they share a common thread: eclecticism.

For multi-modal approaches, eclecticism is central as it allows scholars to 'import and apply a broad range of potent strategies' (Lazarus 2005, 105) while respecting 'science and data-driven findings, and … empirically supported methods when possible' (Lazarus 2005, 106). Another place where multi-modal approaches found their place was in the Engineering Education Research (EER)-Taxonomy version 1.2 (Finelli 2020), where this term was added to section 12.d.iv. Both NSF- convergence research and the EER-Taxonomy highlight the increasing interest to holistically explore and study both complexity and context using various data modalities across and in several areas of expertise.

CONTACT Idalis Villanueva Alarcón i.villanueva@ufl.edu Department of Engineering Education, University of Florida, 1949, Stadium Road Wertheim Building 464, P. O. Box 116120, Gainesville, FL 32611, USA

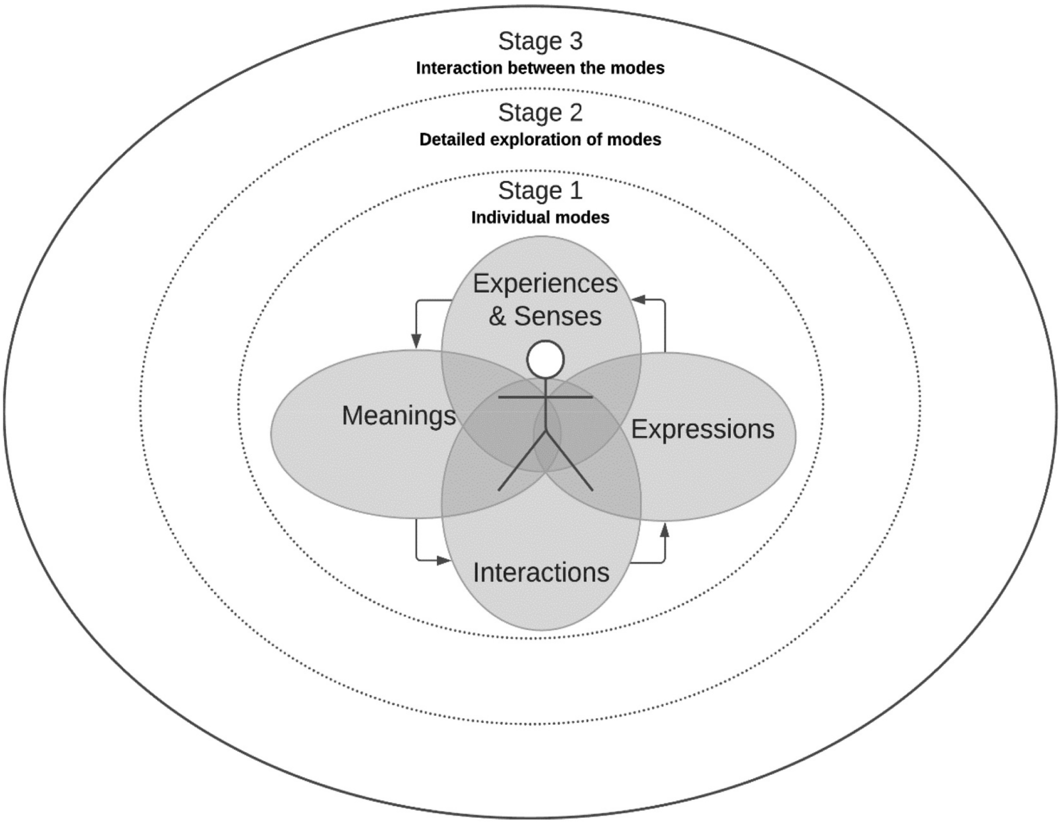

Figure 1. An overview of multi-modal approaches and its stages.

2. Multi-modal approaches: its principles and traits

In disciplinary fields like education and psychotherapy, human beings were originally viewed from the perspective of three modes: cognition, affect, and behaviour (Lazarus 2005).

Lazarus, a psychotherapist, challenged this trimodal notion and argued that at the core, human beings possess at least seven modalities or dimensions: Behavior (action and reaction), Affect (responses of the heart and mind), Sensation (response to the sensory stimuli), Imagery (conjuring of signs, sounds, and other events in our mind), Cognition (entertainment of beliefs, opinions, values, and attitudes), Interpersonal (relationships), and Drugs and biological functions (these seven modalities spell BASIC ID; Lazarus 2005). Much of the multi-modal therapies we see today originated from Lazarus's BASIC ID model (Lazarus 2005; Palmer 2006).

Today, modern multi-modality is not limited to these seven dimensions and include additional dimensions such as spiritual, physical, neurological, genetic, communicative, expressive, among others (e.g. Anna 2022; Bagga-Gupta 2012; Ćosić et al. 2013; Lachs 2017; Ledin and Machin 2017; Leppänen and Tapionkaski 2021; Mills and Unsworth 2018). As a result, today's multi-modal approaches seek to explore beyond traditional unimodal, bimodal, or even trimodal domains of human dimensions. Multi-modal approaches aim to provide ample research evidence and multiple perspectives to understand the complexity that human beings carry and the evolving contexts in which related phenomena prevail.

When defining multi-modal approaches, within or as its own research study, it is essential to situate what traits define multi-modality in order to contextualise what is considered an experience of phenomena. Adapted from Lazarus (2005), Villanueva Alarcón and Anwar (2022) describe multi-modality based on four core principles:

- Principle 1: Humans respond, act, and interact with their environment via several modalities in contextually-informed ways
- Principle 2: Multi-modalities may be intertwined through behavioural and psycho-physiological mechanisms that reciprocate and inform each other. Some or all modalities can co-exist (converge or integrate) or act divergently (split) from a common point of inflection or function independently
- Principle 3: To explore multi-modality, tools and techniques must be systematically applied to study independent and co-dependent modes of meaning-making and their dynamics
- Principle 4: Multi-modal approaches must consider how individuals communicate, represent, and contextualise realities in multiple forms

Furthermore, multi-modal approaches, as summarised in a recent editorial by Villanueva Alarcón and Anwar (2022), contain the following traits:

- Trait 1: Includes many representations (or layers) by which an individual can experience phenomena; these experiences are usually intertwined and complex
- Trait 2: The experience of phenomena should be observed from a wide view (meta-understanding) as well as a focused view (nuanced) of how individual(s) experienced these phenomena
- Trait 3: Individual(s') experiences of phenomena could be focused on internalised (e.g. cognition, affect) and externalised (e.g. behaviour, verbalisation) elements
- Trait 4: Exploring how an individual experiences the phenomena from a multi-modal lens will require multiple tools, techniques, disciplinary expertise, frameworks, and approaches to handling large amounts of data
- Trait 5: Multi-modal experiences may be composed of multiple phenomena occurring almost simultaneously
- Trait 6: Based on the integrated tools, disciplinary expertise, frameworks, and approaches to large data, the research team may need to develop their own methods and strategies to compare, contrast, and validate the collected data to meet their research goals

It is important to note that the authors have previously differentiated between multi-method and mixed-methods research (Villanueva Alarcón and Anwar 2022) and readers are encouraged to refer to this publication for additional details. Here, we summarise the difference between multi-modal and mixed-method briefly: 'Both multi-methods and mixed-methods approaches do not necessarily require internalization and/or externalization of simultaneous and multiple layers of phenomena, which is the premise of multi-modal research' (p.279). In other words, multi-modal synchronization of events by time and space is essential in capturing external (e.g., behavioral observations) or internal (e.g., cognition) responses to an experience. Furthermore, multi-modal approaches require capturing data in near-real-time, which would necessitate 'several tools and/or expertise (e.g. psychology, physiology, neuroscience), multiple frameworks (e.g. intersectionality, literacy, community cultural wealth), and potentially comprehensive approaches to handle big data (e.g. machine learning, statistics, artificial intelligence)' (Villanueva Alarcón and Anwar 2022, 279). As such, techniques and approaches may need to be customised to the research needs.

3. Utility and uses of multi-modal approaches

Multi-modal approaches have been used by scholars who seek to capture comprehensive meaning-making in their research in response to important and complex questions. These questions may address recognition, resources, access, participation, equity, social relations, social justice, mental health, informal learning, governance, and policy-making, among others (e.g. Archer and Newfield 2014; Jewitt et al. 2021).

Within the scope of the practice of education, Stein (2007) suggests using different modes to define communication features in and out of class for curriculum, pedagogy, and assessment practices. According to Stein (2007), instructors use multiple semiotic resources and select combinations of modes to construct ensembles of course materials and strategies to elicit meaning in their students. The instructors provide students access to the material using multiple modes and means (e.g. computers, books, hands-on activities); use symbolic notions (epistemological access to knowledge via tutorials, videos, music, imagery, etc.); promote interactive discourses (e.g. discussions, reflections, debates); design and make changes to curriculum and instruction based on students' facial expressions and cues to cognitive struggles to a concept or topic; and conduct formal assessments (e.g. exams, quizzes, projects). In turn, the interpretation of students to these materials is also multi-modal (Martin 2016).

While these data sources could be treated independently, most of these modes are observed and assessed by the instructor at the moment (Blicblau and Pocknee 2003). For example, collecting all these data sources from students in the class could inform the instructor of the changes that need to be made in to make on the next lesson or course offering. Similarly, in fields like discipline-based educational and social science research, near-real-time events help paint the picture of all the factors that inform a participant's decisions, actions, behaviours, and other representations of the experiences of a given phenomenon.

4. Multi-modality in engineering and computing education

There have been studies already conducted in engineering and computing education using multi-modal approaches, even if they may not have been explicitly described as such (e.g. Atiq 2018; Husman et al. 2019; Villanueva Alarcón et al. 2021; Villanueva et al. 2019; Wert et al. 2021). Explanations of these studies have been covered in Villanueva Alarcón and Anwar (2022). These scholars argue for a need to 'weave together different data sources and expertise to understand how engineering [and computing] students experience the context of their learning or research environments' (p. 280).

Within the context of engineering and computing education, meaning-making and the capturing of multi-modal layers of meaning-making could expand beyond traditional qualitative (e.g. interviews, focus groups, observation), quantitative (e.g. surveys, performance scores), multi- or mixed-methods (e.g. concurrent, sequential) designs to include a new array of digital technologies. With the advent of augmented reality, virtual reality, online education, gaming technologies, social media, sensor technologies, conversational platforms (e.g. Zoom, Teams), and other technological advancements, multi-modal methods could be included in new and exciting ways. The possibilities are endless. However, capturing all these possibilities would be a challenge. In this manuscript, we provide some ways in which engineering and computing education researchers could adopt multi-modal methods in their research areas although these expositions are not meant to be prescriptive.

5. Components of multi-modal approaches

5.1. Pragmatism

Similar to mixed-methods, multi-modality is best situated in a pragmatic worldview, which allows for the versatile use of problem-oriented, innovative, dynamic, and boundary-spanning approaches to answer the research questions (Creswell 2013). The principles of pragmatism, such as recognising the interconnectedness of components, emphasising constructive knowledge, and viewing inquiry as an experiential process (Friedrichs and Kratochwil 2009), align with multi-modality (Bezemer and Jewitt 2010). Furthermore, multi-modality considers creating multi-layered ensembles and interpretations from multiple dimensions of an individual's experience of phenomena (Villanueva Alarcón and Anwar 2022), requiring a blending of multiple techniques and strategies (Creswell 2013). It is noteworthy that experiences of phenomena are socially, contextually, and culturally situated; as such, other paradigms may be present.

5.2. Versatility

Versatility in multi-modal research design allows for its design to become a component of other methodologies (e.g. qualitative, quantitative, mixed- or multi-method studies) or be their own type of study (multi-modal study). When including multi-modal approaches in another methodology, it is essential to include at least some core principles and traits (refer to section 1). Depending on the level of nuance desired, the combination of modalities could be utilised to complement existing methodologies or derive new methods to explore the data from multiple perspectives. These versatile ways to use and analyse the multi-modal data could lead to new understandings of the experience(s) of the phenomena.

For example, in a study by Villanueva et al. (2019b), the researchers explored how undergraduate students performed on engineering statics practice problems. They collected data using multiple modalities, including electrodermal activity sensors, self-efficacy surveys, and salivary stress biomarkers. The authors found that students, independent of the task's difficulty level, experienced mid-to-high levels of self-efficacy, which corresponded with bi-modal levels of emotional arousal and stress. However, they also found that introducing stress biomarkers to the study decreased emotional arousal levels by about five to six times. The versatility of multi-modal approaches and tools is an essential element of a research design using these strategies. Similar examples can be found in Atiq (2018), Wert et al. (2021), and Villanueva et al. (2014, 2019).

5.3. Research questions and/or hypotheses

Depending on the nature of the study and research design, the framing of the experience of phenomena could be represented as a research question, a hypothesis, or a blend of both. Multi-modal research questions and/or hypotheses could be part of the study's primary goal or sub-goal. As a rule of thumb, the multi-modal components for a research question and hypothesis should encapsulate the notion of multiple modalities and their interactions as described in the next sections: data collection, analytical framework, and interpretation.

The following guiding questions can assist the researcher in understanding the types of reflections they should make when considering multi-modal questions. The provided questions are separated into two categories: 1) as part of a research design or 2) as their own study. It is noteworthy that these guiding questions are not meant to be prescriptive and should serve as a starting point for researchers to decide on the approach to take.

As part of a research design:

- How could [different modes of reality] be captured to provide individual meaning to [the phenomena] experienced?
- How could [different modes of reality] be captured to provide comprehensive meaning to [the phenomena] experienced?
- How could [modes of reality] shed new light into multi-faceted paradigms [epistemological,

ontological, axiological] present in [the said phenomena]?
- How could [modes of reality] better contextualise the [selected problem, phenomena, or participant(s)] in both nuanced and comprehensive ways?

As its own multi-modal study:
- How could the researchers in a given study help re-envision [validity, reliability] considerations to my [selected problem, phenomena, or participant(s)] contexts?
- How could a new method be derived from integrating [modes of reality]? How could this new method support or serve to introduce further innovation into the field of [discipline of study]?
- How could [multiple modes of reality] explain the connection between the participant(s') [discourse, practices, perspectives, etc.] and dimensions of meaning-making [representational, social, organisational, contextual, ideological]?

In introducing a multi-modal research question and hypothesis, the researcher will need to align with the goal of the study. As Villanueva Alarcón and Anwar, 2022) indicated, it will be necessary for the researcher to create and justify their choices within the research design and phenomena. Also, an important aspect to remember for multi-modal work is the essence of time. In other words, while multi-modal work could be used in retrospective-type studies, the researcher should aim to capture the event as close to the occurrence of said event as possible (Villanueva Alarcón and Anwar 2022). This immediacy is important because in order to capture the multiple layers of the experience, multi-modal scholars need to consider the authenticity and trueness of the experience. Since recollection of events is subject to recall bias, capturing the event as soon as it happens will ensure more attuned authenticity and validity of the captured phenomena (Villanueva Alarcón and Anwar 2022; Villanueva et al. 2019). However, sometimes capturing the event in near real-time may not be possible. In that case, the researchers need to incorporate elements in their research design that would minimise the recall bias. For instance, researchers have added additional data modes (video of screen capture and eye-gaze scan path) to help participants recall their experience during a retrospective think-aloud interview (Atiq 2018; Wert et al. 2021). These sources of additional data, along with electrodermal activity, were used to provide an additional layer of meaning to the retrospective think-aloud interviews (Atiq 2018; Wert et al. 2021).

5.4. Data collection, analytical framework and its interpretations, and triangulation

In addition to the previous elements, researchers need to consider the following multi-modal components: data collection, analytical frameworks, and interpretation.

5.4.1. Data collection

For data collection, multi-modal approaches seek nuanced, near-real-time data capturing. The goal of collecting data with various modalities is to attain a wide, comprehensive view of the formation and dynamics of the meaning-making process (Villanueva Alarcón and Anwar 2022). The 'near-real-time' requirement is a distinguishing factor of multi-modal work as it requires that the participant has as close of a recollection of the experience as possible (e.g. minutes or days of an event).

While data collected in multi-modal approaches could consist of multiple data types and sources, primary forms of collected evidence rely on self-reports (e.g. surveys, interviews, focus groups), artefacts (e.g. documents, entries), sensory data (e.g. eye trackers, sensors), computational data (e.g. models, algorithms, analytics), and monitoring data (e.g. video, audio, fieldnotes, observational protocols). However, because multi-modal approaches are based on disciplinary expertise and capability, including other types of data modalities is also possible.

5.4.2. Analytical framework

The analytical framework within multi-modal approaches involves three stages, as shown in Table 1 (Bezemer and Jewitt 2010). The three stages include: 1) using individual modes as separate entities and analysing them as independent components of inquiry, 2) rigorous stages of looking at modes together in multiple layers, and 3) examining the interaction between the modes. Since data are typically collected simultaneously using different tools and techniques, it may be necessary to analyse and take notes of the different instances in which these data were collected. The exemplar papers (Section 6) synthesise two studies and provide examples of the three-stage analytical framework described in this section.

5.4.3. Interpretation using multi-modal dimensions

For interpretation, the multi-modal meaning-making process could consider five dimensions (Cope and Kalantzis 2009): 1) ***Representational*** (What do the meanings refer to with respect to modalities and phenomena?); 2) ***Social*** (How does meaning connect the persons they involve?); 3) ***Organizational*** (How do the meanings converge or diverge together?); 4) ***Contextual*** (How do the meanings fit the context of

Table 1. Analytical framework stages of multi-modal approaches.

Description	Data Collection	Analysis	Interpretation	Triangulation
Stage 1: Individual modes				
Treat individual modes as separate entities to analyze them as independent components of inquiry.	• Self-report • Video recording • Observational notes • Sensory data*Important Consideration:* If data is being collected concurrently, researchers need to ensure that individual modes are synchronized spatially and temporally.	• Descriptive self-reports (e.g. qualitative approach) • Themes from video recording (e.g. qualitative approach) • Key aspects from observations (e.g. qualitative or quantitative approach) • Trends derived from sensor data (e.g. qualitative or mixed methods approaches)	What is the meaning derived from each mode of inquiry?	Triangulation may not be necessary at this stage since each mode is analyzed individually. However, individual data modes may have many dimensions that could be explored together. For example, video data may be recorded using different cameras. Each camera may collect different perspectives of the same phenomena.
Stage 2: Detailed exploration of modes				
Revisiting the individual modes from the perspective of adjacent mode(s) allows for the layering of the data with each other. This may not necessarily entail integration or confirmation of each data source.	• If the data modes are not synchronised at stage 1, researchers could identify the points of convergence between different modes. • Creating multi-layers based on events	• Diagnostic assessment of each revisited mode to identify new layers or knowledge of said phenomena. (OR) • Exploration of multiple possibilities of the nature of the data source and finding	Which layers collectively provide the same perspective, and which add new dimensions to meaning-making?	*Layered Convergence:* Different modes tell researchers the same story and perspective of the phenomena at a certain point in time. *Layered Divergence:* Different modes tell researchers a different story and perspective of the phenomena, even if it is by a single mode.
Stage 3: Interaction between the modes				
Examine the interaction between the modes to collectively capture an 'eyewitness' account of the phenomena and experience.	• Creation of multi-modal wholes'	• Finding the commonality and variations between modes and layers • Side-by-side comparison of convergent and divergent modes and layers	Convergences, divergences, integrated instances, and independence of modes in the meaning-making process and its corresponding dimensions.	*Contextualized Convergence:* Different layers formed at stage 2 are placed side-by-side to identify the similarities between the stories and perspectives of the phenomena (eyewitness accounts). *Contextualized Divergence:* Different layers formed at stage 2 are placed side-by-side to identify the differences between the stories and perspectives of the phenomena.

the study for meaning making?); 5) ***Ideological*** (Whose interest(s) do the meaning(s) serve?).

Since the meaning-making process is closely tied to the representative modalities, these modes could provide a plethora of representations of a given participant's experience within a single layer of meaning, for instance, writing, voices, images, gestures, or spatial understandings. The nuance in interpreting these data, based on the collected source, is up to the discretion and resource capabilities of the researcher (Cope and Kalantzis 2009).

For example, meaning may refer to nouns and verbs in written descriptions, shapes, and colours in an image (representational). These words and images may have physical or symbolic expressions discussed in groups (social). These modes could be further contextualised to provide cohesion of meaning, such as focusing on the sequence of words or the placement of objects in an image (organisational) to provide a larger picture of how this meaning fits into the larger world (contextual) and explore the motivation behind participants' uncovering of its meaning (ideological). Added to these meanings, there should be a conscious approach by the researcher to consider secondary, tertiary, or additional layers of meanings based on additional data sources or angles (layers) of the same collected data source. As a reminder, in multi-modal work, including layers of the experience is essential (Villanueva Alarcón and Anwar 2022). The exemplary papers section describes meaning-making dimensions with existing studies (refer to section 6).

Multi-modal meaning-making could involve more than one dimension. It is up to the researcher's discretion to identify how many layers to include in their analysis. The decision of when to stop layering meanings will come when no new meanings are identified (e.g. like points of saturation in a grounded theory study). Once these data have been put together, approaches such as but not limited to learning analytics and machine learning could be used to present a 360-degree picture of the experienced phenomena and summarised into 'multi-modal wholes'. One important consideration when using learning analytics is that while capturing the data could occur in real-time, these may not fully represent the authenticity of the experience, especially if variables are being synthetically controlled during experimentation. The researchers are encouraged to carefully think of all the stages presented in Table 1 to ensure all elements of multi-modal approaches are being addressed.

As indicated in stage 3 (interaction between modes) in Table 1, it is important to note that based on the goal of the researcher, interpretation could serve to divergently explain phenomena (separate accounts or instances of phenomena to represent diversity in the experience), could be converged to explain phenomena (to explain how individual instances intersect, be both convergent and divergent (where the integrated instances and divergent experiences serve to represent the flux and flow of a complex and multi-faceted phenomena), or could be treated as an independent event.

5.4.4. Triangulation

A noteworthy aspect of creating 'multi-modal wholes' is an extension to the triangulation process (Boyd 2007, 2008; Boyd and Crawford 2011). Typically, triangulation uses data from different times, spaces (immediate, retrospective, future-oriented), theoretical perspectives, methodologies, and people in which multiple investigators may collect or analyse the data for a phenomenon (Denzin 2012; Hammersley 2008). In such cases, triangulation aims for coherence between data sources to reduce bias and establish credibility and viability (Hammersley 2008). In other words, triangulation seeks confirmation.

Triangulation uses different data sources and modes and assumes that these sources will align to paint a holistic picture of the studied phenomena (Denzin 2012). However, when data or methods diverge, the triangulation and trustworthiness of the phenomena are put into question (Hammersley 2008). Multi-modal approaches posit that even if data or methods diverge, there is value in understanding other aspects of the phenomenon that may be interacting differently with the individual(s). Multi-modal approaches are typically situated in the dynamic and raw ways individuals make sense of their experiences. It doesn't mean that participants can't think retrospectively or prospectively about an experience, but rather that data collection tools and techniques should have the central aim of ensuring that they can capture multi-layered phenomena via different data collection sources (e.g. interviews, videos, physiological sensors, AI). This pairing of data collection strategies helps provide a richer and more nuanced understanding of experience. This also means that the researcher is intentionally introducing more ways to analyse data by considering multiple angles, aspects, and views of the phenomena with the aim of presenting multiple layers of interpretation of the same event or experience.

Multi-modal approaches 'must, perforce, move around and adopt different perspectives to capture these different aspects' of phenomena" (Hine 2015, 29). This may mean that the research object is not viewed in the same way throughout the triangulation process, even if it is the case of a single context study. This type of extension to triangulation is useful for internet research (Boyd 2007, 2008; Boyd and Crawford 2011), for example, where participants fluctuate over the course of a study, or do not follow a distinct sample from one site to another, or where it is less straightforward to identify a single research object. If more than one mode provides additional views of the phenomenon and these don't coalesce, then most likely, a multi-modal approach extension to triangulation approaches may be required. The definitions of triangulation at different multi-modal stages are explained in Table 1.

6. Exemplar papers

6.1. Positionality statement

For this work, the authors are candidly disclosing how their professional experiences over a number of years and collaborations have informed the current work. The first author has over 10 years of experience in developing and creating multi-modal research approaches, and the second and third authors have approximately 5–6 years of experience developing multi-modal research studies. All agree that there is a scholarship that uses multi-modal methods within quantitative, qualitative, multi-method, or mixed-methods studies but seldom is considered as its own set of studies, or do they get explicitly categorised as multi-modal. Each author carries multidisciplinary expertise in engineering and computing education research and has professional experiences in science, engineering, and computer science. For this reason, the authors are uniquely positioned to identify and categorise multi-modal work used in the examples explained below.

6.2. Selecting the exemplary papers

It is noteworthy to mention that the authors did a methods search of premier journals and conferences of engineering and computer science education with the terms 'modal', 'multi-method', 'mixed-method', and 'multiple'. These journals and conferences include, but not limited to: the International Journal of Engineering Education (IJEE), the Journal of Engineering Education (JEE), the Australasian Journal of Engineering Education (AJEE), the European Journal of Engineering Education (EJEE), Studies in Engineering Education (SEE), Advances in Engineering Education (AEE), ACM Special Interest Group in Computer Science Education (SIGCSE), IEEE Frontiers in Education (FIE), ACM International Computing Education Research (ICER), ACM Transactions on Computing Education (TOCE), IEEE Transactions on Education (TOE). Most of the search yielded examples of work that did not typically fit the definition of multi-modal approaches posed in this manuscript. We did identify a handful of examples that were close to multi-modal approaches. We present two exemplary papers, one from engineering and computer science education and one from learning and instruction, that adopt multi-modal approaches to answer their research questions. The selection of exemplary papers was guided by criteria of using multi-modal approaches, a relatively recent publication in national and/or international scholarly venues, and a clear explanation of the collective perspective of various modes.

6.3. Discussing the exemplary papers

This section summarises each paper with its overview, research questions and hypotheses, data collection modalities, three-stage analytical framework, and dimensions of multi-modal meaning-making. Since multi-modal research is still in its infancy in engineering and computing education research, these studies may not have explicitly applied to the three-stage framework or the interpretation dimensions. As such, we opted to map the analysis and results of the papers to these elements to help readers understand the possibilities of better communicating the multi-modal approaches to their work. We also included a schematic representation of each paper to help the reader better situate the study. It is the hope of the authors that future papers will begin to situate their multi-modal work to the recommended suggestions provided in this manuscript.

Paper 1: Villanueva et al. (2018). A multi-modal exploration of Engineering students' emotions and electrodermal activity in design activities. Journal of Engineering Education, 107(3), 414–441.

6.3.1. Overview

The paper uses multi-modal approaches to explore first-year mechanical engineering students' classroom emotional experiences during an engineering graphics and design course. For the study, students were

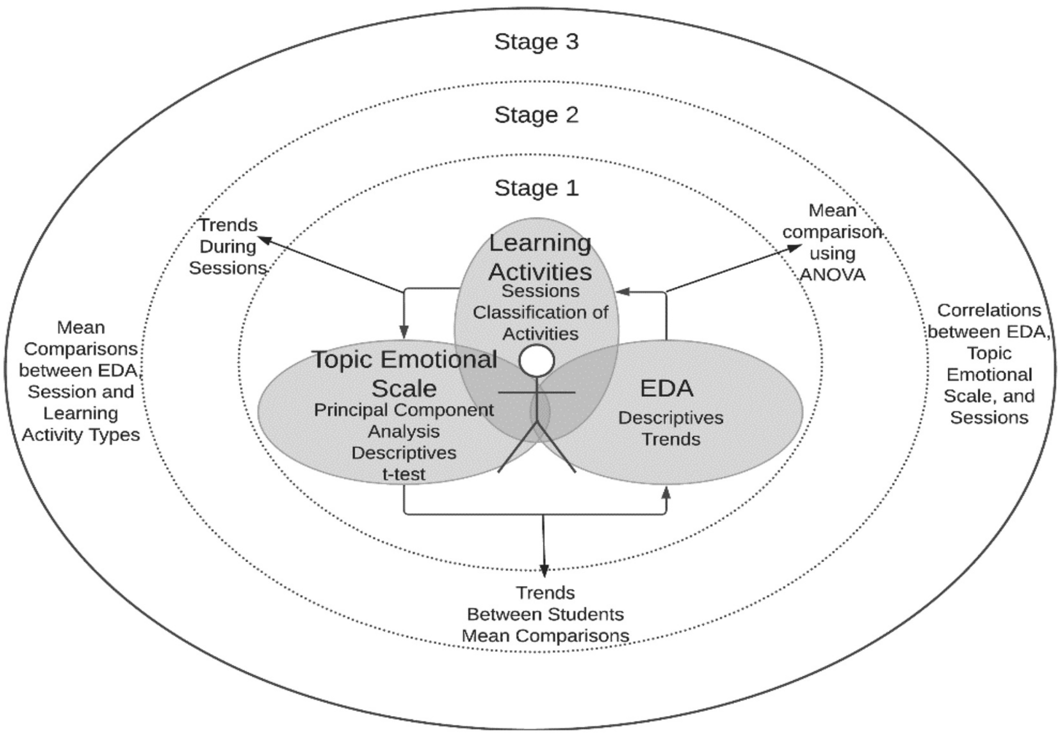

Figure 2. Visual representation of paper 1.

engaged in five sessions specifically created to enhance students' design skills while working on an engineering design problem. These sessions collectively were referred to as the Engineering Design Method Workshop. Each session was taught using a mix of passive and active learning activities. The researchers emphasised the need to use multi-modal approaches for a deep and holistic examination of engineering students' internalisation of the socialised learning process. Also, they associated emotional experiences for each session and throughout the semester. A visual representation of this paper is summarised in Figure 2.

6.3.2. Research questions and/or hypotheses

The paper's hypothesis is based on Piagetian constructivist tasks and activities, informing those cognitive processes that may suppress emotional responses (Kleinginna & Kleinginna, 1981) independent of the topic or instructional format. The researchers hypothesised that regardless of the type of learning activities, there would be no change between students' self-reported topic emotions from pre- to post- for each session. Also, there will be no association between electrodermal activity (EDA, a measure of emotional arousal) and self-reported topic emotions throughout the design activities.

6.3.3. Modalities of data collection

The data were collected using two sources of students' emotional experiences. These sources are: 1) self-reported students' emotions using a modified and validated Topic Emotions Scale (Broughton et al. 2013) from 88 students in a pre-post manner. The survey comprised 13 items for both positive and negative emotions and 2) EDA data from selected 14–18 students who wore a non-invasive wrist sensor during each design session. The paper described the delivery and content of the five sessions, where three types of learning activities existed for each session, i.e. passive, individually active, and collaborative active. Additionally, the timeline for each type of learning activity was established within each session.

6.3.4. Three-stage analytical framework

- In Stage 1 (Individual Modes), the survey data was validated using principal component analysis. The validation process identified one dimension and two factors (positive and negative topic emotions) within the survey. The mean and standard deviation for survey data were calculated for positive and negative topic emotions. Researchers also administered t-tests for mean comparison between pre-and post-survey. The results indicated a non-significant mean difference between the pre-and post-survey. The EDA was initially mean-corrected to reduce error variance. Also, the EDA data were analysed to examine the trends for each student. Each participant who wore the EDA sensor showed a different, fluctuating response indicating a nonmonotonic (nonlinear) trend.
- In Stage 2 (Detailed exploration of modes), the data were analysed in different layers of two adjacent modes. In the first layer, survey data was analysed while accounting for the context of the sessions. The trends were examined for each session individually for both positive and negative topic emotions. The results indicated that students experienced higher positive topic emotions for each session. However, only the second session showed significant results. In another layer, survey data was examined alongside EDA, where researchers compared students' topic emotions between those who wore the sensors against those who only completed the survey and found no difference. The next layer compared the EDA data within each session for trends and mean comparisons. The trends indicated a linear trend for four sessions, while a cubic trend was observed for the fifth session. Further, the mean comparison analysis was conducted using within contrast Analysis of Variance (ANOVA), indicating a significant change in EDA data for all sessions.
- In Stage 3 (Interaction between the modes), the 'multi-modal wholes' were created in two combinations. For the first combination, the data was analysed with collective modes of EDA, sessions, and learning activity types. Using a t-test, the results indicated that for the fifth session, there was a significant difference between learning activity types, where EDA mean values were significantly higher for collaborative learning activities than passive activities. The results of repeated measures ANOVA on session 4, which included all three types of learning activities, indicated significant differences between passive and individual active learning activities. Also, the same significant results were observed between passive and collaborative learning activities. For the second combination, the collective examination of EDA data with the survey responses and sessions was conducted. The Pearson correlation between EDA data, positive and negative topic emotion per session, indicated moderate correlations for sessions 1, 2, and 4 with negative emotions and sessions 3 and 5 with positive emotions. Also, EDA and negative emotions showed a strong significant correlation for session 1 only.

6.3.5. Dimensions of multi-modal meaning-making

The paper used all five dimensions of multi-modal meaning-making.

- *Representational*: The representational dimension allowed the exploration of physiological and behavioural indicators for understanding students' experiences while engaged in problem-solving during workshop sessions.
- *Organisational*: The meaning-making process was organisational as it used two primary modes (i.e. survey and EDA) for convergence and divergence with other modes, such as learning activities and workshop sessions.
- *Contextual*: The contextual aspect of the workshop session and activities helped in identifying participants' emotions during the learning activities and within each session.
- *Social*: For the social dimension, the study examined data from collaborative activities. The inclusion and analysis of collaborative active participation allowed individual participants to work with other students, which may have caused emotional arousal.
- *Ideological*: The paper describes the ideological dimension by considering the interest of students' engagement in classroom settings. Also, the study provided future researchers and practitioners the initial steps to use a multi-modal interdisciplinary approach in engineering and computing education.

Paper 2: Järvelä, S., Malmberg, J., Haataja, E., Sobocinski, M., & Kirschner, P. A. (2021). What multi-modal data can tell us about the students' regulation of their learning process. Learning and Instruction, 72(7).

6.3.6. Overview

The paper discusses and demonstrates how multi-modal data is collected and triangulated to generate evidence about the self-regulated learning (SRL) process in collaborative learning environments. The authors presented the evidence using five examples to explain three aspects: 1) interactions between different facets of regulation, including cognition, motivation, and emotion, 2) occurrence and temporality of different types of regulation; and 3) temporality and cyclical process of regulation involving planning, enacting strategies, reflecting, and adapting. The data for each example were collected from 48 high school students who were divided into groups of three students during a 75-minute-long collaborative science experiment. In this paper, the collaborative work of three group members was considered. A visual representation of this paper is summarised in Figure 3.

6.3.7. Research questions and/or hypotheses

The papers hypothesise that multi-modal data can comprehensively understand the self-regulatory process in collaborative activities. Furthermore, the triangulation of multiple data sources has the potential to advance the theoretical and conceptual progress of social aspects of SRL theory. Furthermore, the authors suggested using five case examples to demonstrate

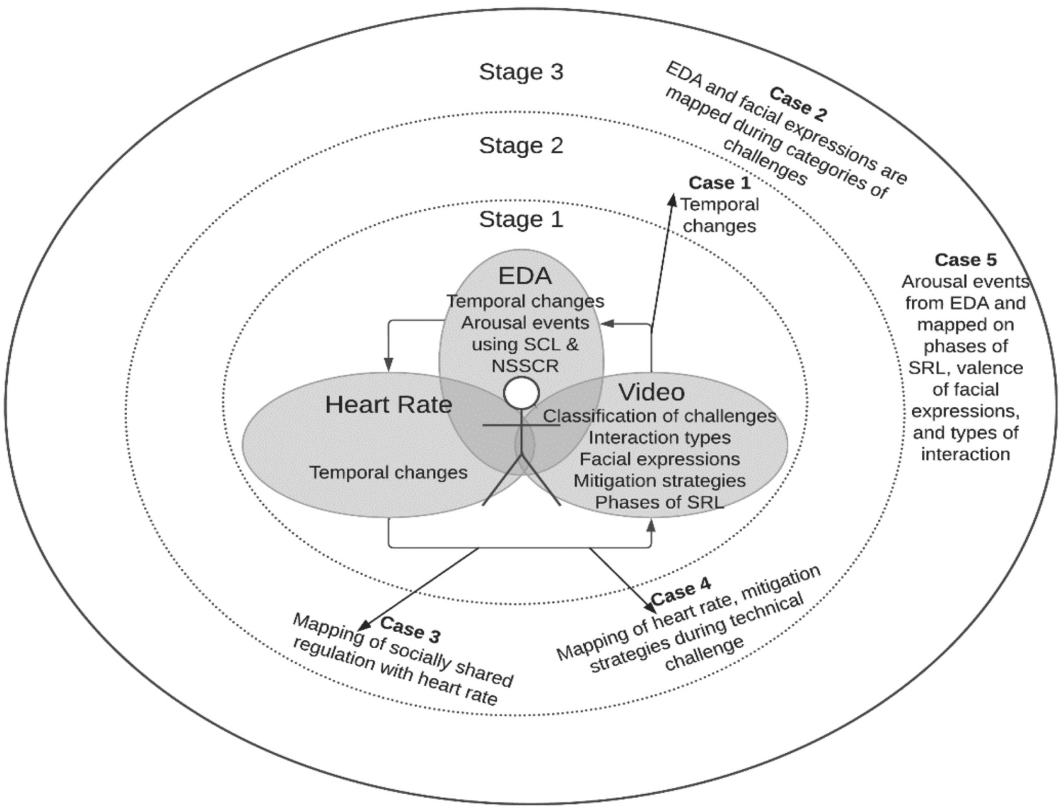

Figure 3. Visual representation of paper 2.

how multi-modal data is collected, used, contextualised, and analysed for future research.

6.3.8. Modalities of data collection

The data were collected from different data channels or sources, which authors categorised as subjective (e.g. repeated and contextualised self-reports of perceptions and behaviours), and objective (e.g. log data, physiological measures, behavioural and mental process information captured from video data). These data were collected using objective and subjective measures of regulation of learning. The data modalities included video and physiological data comprising electrodermal activity and heart rate.

6.3.9. Three-stage analytical framework

- In Stage 1 (Individual Modes), the video data were analysed using qualitative content analysis to recognise student challenges and types of student interaction to develop mitigation strategies. The challenges were classified into three categories: task understanding, technical, and motivational challenges. The authors argued that identifying challenges during the experiment is important as they might trigger the learning regulation process and noted the observations of potentially triggering events. The interaction types were classified as low interaction (group members were processing information together with no visible regulation and less verbal communication), high interaction (group members were co-constructing meaning and regulated learning), and confusion (involves "markers of metacognitive monitoring and prompting others to regulate", pg.6). Also, the video data revealed instances of socially shared regulation processes (if occurred) after any challenge identification. Also, the regulation process was divided into SRL phases: task interpretation, planning, and enactment.
- Additionally, the video data was analysed for the detection of visible faces. The facial expression was recognised for each face in positive, neutral, or negative valence classifications. The EDA data were standardised and compared for temporal changes. The data were processed to measure the skin conductance level (SCL) and the frequency of non-specific skin conductance responses (NSSCR) changes. The data points were observed for the rapid occurrence of NSSCRs and increased SCL. Similarly, the heart rate data were analysed for temporal changes.
- In Stage 2 (Detailed exploration of modes), example cases 1 and 3 formed a layer of data for exploration. For example, in case 1, the video challenges observed in the collaboration from video data were layered with the EDA data. The results were analysed by comparing temporal changes with video segments. The results indicated that SCL and NSSCR were at the beginning of the learning session. Also, the higher NSSCRs were observed during the collaborative learning sessions. The authors concluded that the elevated levels indicated the potential to capture interactions between facets of regulations. For example, in case 3, the instance of socially shared regulation was analysed with the heart rate data. During socially shared regulation, the group received the task, and all three students were involved in the co-construction of knowledge. One student was reading the task out loud, and others were wondering what they needed to do. The heart rate data was mapped in 30-second windows over the total period of this socially shared regulation. A different heart rate was observed during different aspects of the tasks, with an increased heart rate found during the challenging situation and after regulation. However, there was a decrease in heart levels during socially shared regulation. These changes in heart rate were synchronised between the three group members. Similar to the example of case 3, In the example of case 4, the mitigation strategies data were analysed with the heart rate in a technical challenge. It was observed that students showed different heart rates based on their behaviours during the challenge. The behaviours could be off task, related to the challenge, and co-regulation. Also, the authors noted that students with the most cognitive activation through verbal participation showed the most fluctuation in heart rate. Collectively, example case 3 and example case 4 explained the occurrence and temporality of different types of regulation.
- In Stage 3 (Interaction between the modes), the collective aspects of multi-modal data were analysed and explained in the example case 2 and example case 5. For example, in case 2, the layers in case 1 were explored with a closer look at categories of challenges and were synchronised with EDA signal data and facial expressions during a chunk of elevated NSSCRs. These data were analysed by considering temporal EDA changes

and mapping them on the categories of challenges and students' expressions. It was observed that SCL and NSSCR increased simultaneously and synchronised during technical challenges. Also, the group evidently observed frustration, providing evidence of physiological synchrony. Collectively, example case 1 and example case 2 provided the evidence for the first aspect and explained: 'what multi-modal data can reveal about interactions between different facets of regulation' (pg.4). For example case 5, the arousal events from EDA, the different phases of SRL, the valence of facial expressions, and the type of interaction were analysed and mapped together for collective perspective. The results indicated that low interaction was observed during increased EDA signals. However, high EDA signals were not always associated with the regulation of learning. Also, low and high types of interaction were found to be associated with SRL phases, such as planning and task enactment, and confusion, as the type of interaction was not associated with any SRL phase. However, during confusing interactions, the students expressed negative expressions and increased EDA rates. In the 'multi-modal wholes', the authors argued that the collective perspective of the subjective and objective data channels could help capture the cyclical SRL process. However, due to data processing limitations, it is not possible to directly map each arousal on the regulation of learning.

6.3.10. Dimensions of multi-modal meaning-making

The paper used all five dimensions of multi-modal meaning-making.

- *Representational*: The representational dimension allowed the exploration of subjective and objective data channels for understanding regulatory processes in collaboration.
- *Organisational*: The meaning-making process was organisational as it set five examples to triangulate multiple data sources using EDA signals, heart rate, facial expressions, and observations made from video data. The authors extended the example case 1 and provided a zoomed-in version in case 2 for convergence of findings. Similarly, example case 4 provided an enhanced perspective on example case 3. Also, case 5 provided the collective perspective for convergence and divergence.
- *Contextual*: The contextual aspect of the collaborative learning task performed by a group of three students helped identify and map the regulation phase, quality of interaction, expression valence, and EDA signals in a synchronised manner.
- *Social*: The paper captured the social dimension by focusing on students' collaborative activities and types of interaction between the participants during the experiment. Notably, the social dimension was also achieved by considering the group as a unit of analysis, not individual students.
- *Ideological*: The paper describes the ideological dimension by considering the SRL processes in collaborative learning tasks. Also, the study provides a mechanism to use multi-modal data for unveiling the interaction between affective responses and strategic cognitive interactions.

7. Considerations and challenges of multi-modal approaches

Since multi-modal approaches are not limited to a single disciplinary practice, many considerations and challenges could be present across the multiple facets of the study. Some of these are introduced with respect to research design, resources, and techniques for data collection, analysis, and interpretation, among others.

In research design, the overarching consideration is based on the context of the study and its research questions and hypotheses. In a study that includes multi-modal approaches, every research question, hypothesis, and research context would lend themselves to using a different design. In turn, the research design could be influenced by the setting (e.g. classroom, workshop, lab, informal learning) where the research will be conducted. This will require that the researcher controls for confounding factors, a challenge of multi-modal studies. For this reason, many existing studies have opted for more lab-controlled environments rather than naturalistic settings (e.g. Atiq and Loui 2022; D'Mello and Graesser 2010; Harley, Bouchet, and Azevedo 2013; Harley et al. 2015; Villanueva, Valladares, and Goodridge 2016; Villanueva et al. 2018; Wert et al. 2021). However, naturalistic studies could also be conducted (Villanueva et al. 2018), although contrary to lab studies, would require large number of resources (people, equipment, time, money, expertise) to handle the data.

All the studies, whether conducted in lab-controlled environments or naturalistic settings, will require a keen understanding of the study's experimental setup, which will inform how to address unanticipated challenges in consideration of available resources. Also, each experimental setup will require specific expertise and training of research personnel. These trainings may include: 1) training on the use of specialised software and hardware equipment (e.g. eye tracker, facial recognition software), 2) understanding of modalities, and 3) how data should be analysed. In addition, such studies may incur costs related to the number of people required to conduct the study, collect data, and analyse data. If the researchers have budgetary constraints to purchase expensive software and hardware, they may spend additional time developing customised protocols before conducting the experiments (e.g. Villanueva et al. 2019). It is important to note that multi-modality may not require quantitative approaches, and a purely qualitative multi-modal study may allow for studies that are limited in terms of resources or tools.

Additionally, multi-modal approaches require several considerations regarding tools and techniques for collecting, analysing, and interpreting data. For example, consider the use of sensory data to explore emotions and cognition in engineering and computer education tasks (Villanueva et al. 2018; Wert et al. 2021), where each tool used to collect the sensory data comes with its own set of confounding factors (e.g. temperature, signal noise, movement restrictions). The instruments used to measure the sensory data also have manufacturer limitations that, if unknown or inadequately handled, could introduce skewness in the data. As such, these factors need to be considered from the onset of the study and not as an afterthought.

At the same time, tools used to collect multi-modal data may capture more than one source of evidence. For example, non-invasive electrodermal sensors are used to collect a participant's skin electrical conductivity, which is an indirect measurement of arousal. However, electrodermal activity is a proxy to measure emotional arousal (Benedek and Kaernbach 2010; Boucsein and Backs 2009, Bouscein, 2012; Villanueva et al. 2018) and is an indicator of cognitive engagement (Bouscein, 2012; Tranel 2000). The same could apply to eye-movement data as it could help provide an understanding of cognitive load while students work on a learning task, but it cannot be differentiated from attention (Köles 2017; Lai et al. 2013; Obaidellah, Al Haek, and Cheng 2018; Van Gog and Scheiter 2010). From the scenarios provided, even within a single data collection stream, it is unclear if the physiological data collected is due to emotions, cognitive load, or attentional events. As such, researchers may need to anticipate these limitations in the collected data as they may compromise the validity of a study (e.g. Villanueva et al. 2019). Researchers may also need to consider introducing other techniques that would allow the data to be teased out more succinctly (e.g. introducing emotional surveys, facial expression technologies, and observational protocols). However, this may imply that the research team should be equipped to handle either larger sample sizes or data sets to tease out the meaning-making layers of the experience of phenomena (Harley et al. 2015; Villanueva et al. 2018).

Additionally, to analyse and understand data coming from multiple sources so that individual, convergent, divergent meaning-making could be extracted, it is crucial to initially synchronise different sources of data as much as possible (Harley et al. 2015; Husman et al. 2019; Villanueva & Anwar 2022; Villanueva, Valladares, and Goodridge 2016). During the time synchronisations, the researcher will need to identify how to measure the individual layers of the experience, how they converge, and how they diverge. These approaches may require data mining and machine learning algorithms, as well as a semi- or full automation of the data analysis process (Chejara et al. 2019).

8. Possibilities and conclusion

Due to the complexity of contemporary societal challenges, we can no longer afford to work in disciplinary silos to seek answers to critical questions surrounding complex and contextual problems. Multi-modal approaches offer the potential to attend to these challenges by leveraging many aspects of inquiry. Although multi-modal approaches are not new, engineering and computing education research communities are strategically positioned to use new technologies and tools in their disciplinary fields of expertise to move the needle of innovation and inquiry. Using new tools and technologies could greatly inform new avenues of engineering and computer science education, such as online learning, artificial intelligence projects, team-based assignments, and other activities that inform educators in a timelier fashion about how their students are doing and how they are applying corresponding actions. As such, this manuscript was written to equip further the larger national and international engineering and computing education community to embrace and incorporate new and innovative multi-modal approaches that directly impact the education and research of its learners.

Furthermore, multi-modal approaches afford new levels of theoretical and analytical evidence for impact. For theoretical impact, multi-modality is a field of application informed by theory but is not necessarily bounded by it. The theoretical impact suggests that multiplicity in meaning-making could be captured while acknowledging the individuality, convergence, and divergence between modes. These multiple modes of meaning-making question the strict division of considering only modes that represent the part of the world we are trying to account for and suggests that meaning-making doesn't happen in isolation of one mode over another; instead, they co-exist.

Finally, when multi-modal approaches are embedded as part of quantitative, qualitative, mixed- or multi-method studies or as their own studies, they can further our understanding of the multiple layers that may exist in an experience of the phenomena in near-real-time. This allows scholars to situate the authenticity of the experience more holistically. Of course, researchers must also consider challenges and considerations associated with designing and conducting multi-modal studies and learn to mitigate its challenges to the best extent possible.

Acknowledgments

This material is based upon work supported, in part, by the National Science Foundation (NSF) Numbers DUE 2120451, 1661117, 1661100 (Dr. Idalis Villanueva Alarcón) and NSF IIS 2104729 (Dr. Zahra Atiq). Any opinions, findings, conclusions, or recommendations expressed in this material do not necessarily reflect those of NSF. Dr. Anwar acknowledges the start-up funds provided by Texas A&M University.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

The work was supported by the National Science Foundation DUE [2120451, 1661117, 1661100] to Dr. Villanueva Alarcón; Start-Up (Texas A&M University; Dr. Anwar) National Science Foundation IIS [2104729] to Dr. Atiq. Any opinions, findings, conclusions, or recommendations expressed in this material do not necessarily reflect those of NSF.

Notes on contributors

Dr. Idalis Villanueva Alarcón is an Associate Chair for Research & Graduate Studies and tenured Associate Professor in the Department of Engineering Education at the University of Florida. She is an expert in hidden curriculum, mentoring, and multi-modal approaches in engineering education. Note that "Villanueva" and "Villanueva Alarcón" in the citations are the same author.

Dr. Zahra Atiq is an Assistant Professor of Practice of Computer Science and Engineering at The Ohio State University. She is a computing education researcher who focuses on using multimodal methods to understand non-cognitive factors in the context of CS1 courses. She is also working on curriculum development of an undergraduate quantum computing course.

ORCID

Idalis Villanueva Alarcón http://orcid.org/0000-0002-8767-2576
Saira Anwar http://orcid.org/0000-0001-6947-3226
Zahra Atiq http://orcid.org/0000-0002-7905-2553

References

Anna, N. 2022. Multi-Modality in Instagram Posts by Psychotherapists as a Meaning-Making Tool. *Collection of Scientific Papers «SCIENTIA», Helsinki, Finland*, June 24, 50–51. Retrieved from https://previous.scientia.report/index.php/archive/article/view/284.

Archer, A., and D. Newfield. 2014. *Multi-Modal Approaches to Research and Pedagogy: Recognition, Resources, and Access*. New York: Routledge.

Atiq, S. Z. 2018. "Work in Progress: A Multi-Modal Method for Assessing Student Emotions During Programming Tasks." Paper presented at the *ASEE Annual Conference & Exposition*. Salt Lake City, Utah. https://doi.org/10.18260/1-2–31264

Atiq, Z., and M. C. Loui. 2022. "A Qualitative Study of Emotions Experienced by First-Year Engineering Students During Programming Tasks." *ACM Transactions on Computing Education (TOCE)* 22 (3): 1–26. https://doi.org/10.1145/3507696.

Bagga-Gupta, S. 2012. "Privileging Identity Positions and Multi-Modal Communication in Textual Practices: Intersectionality and the (Re) Negotiation of Boundaries." In *Literacy Practices in Transition: Perspectives from the Nordic Countries*, edited by A. Pitkänen-Huhta and L. Holm, 75–100. Multilingual Matters. https://doi.org/10.21832/9781847698414-006.

Bandura, A. 1969. "Social-Learning Theory of Identificatory Processes." In *Handbook of Socialization Theory and Research*, edited by D. A. Goslin, 213–262. Chicago, Il: Rand McNally.

Benedek, M., and C. Kaernbach. 2010. "A Continuous Measure of Phasic Electrodermal Activity." *Journal of Neuroscience Methods* 190 (1): 80–91. https://doi.org/10.1016/j.jneumeth.2010.04.028.

Bezemer, J., and C. Jewitt. 2010. "Multi-Modal Analysis: Key Issues." In *Research Methods in Linguistics*, edited by L. Litosseliti, 180–193. London: Bloomsbury Publishing Plc.

Blicblau, A. S., and C. Pocknee. 2003, January. "A Multi-Modal Approach to Teaching and Learning: A Case Study of Teaching Materials and Process Engineering." In *Engineering Education for a Sustainable Future: Proceedings of the 14th Annual Conference for Australasian Association for Engineering Education and 9th Australasian Women in Engineering Forum*, Melbourne, Australia.

Boucsein, W. 2012. *Electrodermal activity*. Springer Science & Business Media.

Boucsein, W., and R. W. Backs. 2009. "The Psychophysiology of Emotion, Arousal, and Personality: Methods and Models." In *Handbook of Digital Human Modelling*, edited by V. G. Duffy, 35–38. Boca Raton, FL: CRC Press.

Boyd, D. 2007. "Why Youth Social Network Sites: The Role of Networked Publics in Teenage Social Life." In FoundatioRn Series on Digital Media And Learning -- youth, Identity And Digital Media *Volume*, edited by

D. John and C. T. MacArthur, 119–142. Buckingham. Cambridge, MA: MIT Press.
Boyd, D. M. 2008. *Taken Out of Context: American Teen Sociality in Networked Publics Doctoral Thesis Submitted to Graduate Division*. University of California, Berkeley. http://www.danah.org/papers/TakenOutOfContext.pdf.
Boyd, D., and K. Crawford. 2011. *Six Provocations for Big Data. A Decade In Internet time: Symposium on the Dynamics of the Internet and Society*.
Broughton, S. H., G. M. Sinatra, and E. M. Nussbaum 2013. ""Pluto Has Been a Planet My Whole life!" Emotions, Attitudes, and Conceptual Change in Elementary students' Learning About Pluto's Reclassification." *Research in Science Education*, 43: 529–550.
Bunt, H., R.-J. Beun, and T. Borghuis. 1998. *Multi-Modal Human-Computer Communication: Systems, Techniques, and Experiments*. Vol. 1374. New York, NY: Springer Science & Business Media.
Chejara, P., L. P. Prieto, A. Ruiz-Calleja, M. J. Rodríguez-Triana, and S. K. Shankar. 2019. "Exploring the Triangulation of Dimensionality Reduction When Interpreting Multi-Modal Learning Data from Authentic Settings." In *Transforming Learning with Meaningful technologies. EC- TEL 2019*, edited by M. Scheffel, J. Broisin, V. Pammer-Schindler, A. Ioannou, and J. Schneider, 664–667. Vol. 11722 Lecture Notes in Computer Science. Cham: Springer. https://doi.org/10.1007/978-3-030-29736-7_62.
Cope, B., and M. Kalantzis. 2009. "A grammar of multi-modality." *International Journal of Learning* 16 (2): 361–425. https://doi.org/10.18848/1447-9494/CGP/v16i02/46137.
Ćosić, K., S. Popović, M. Horvat, D. Kukolja, B. Dropuljić, B. Kovač, and M. Jakovljević. 2013. "Computer-Aided Psychotherapy Based on Multi-Modal Elicitation, Estimation and Regulation of Emotion." *Psychiatria Danubina* 25 (3): 340–346.
Creswell, J. W. 2013. *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. California: Sage publications.
Denzin, N. K. 2012. "Triangulation 2.0." *Journal of Mixed Methods Research* 6 (2): 80–88. https://doi.org/10.1177/1558689812437186.
D'mello, S. K., and A. Graesser. 2010. "Multi-Modal Semi-Automated Affect Detection from Conversational Cues, Gross Body Language, and Facial Features." *User Modeling and User- Adapted Interaction* 20 (2): 147–187. https://doi.org/10.1007/s11257-010-9074-4.
Finelli, C. 2020. "A Taxonomy for the Field of Engineering Education Research." http://taxonomy.engin.umich.edu/taxonomy/eer-taxonomy-version-1-1/.
Friedrichs, J., and F. Kratochwil. 2009. "On Acting and Knowing: How Pragmatism Can Advance International Relations Research and Methodology." *International Organization* 63 (4): 701–731. https://doi.org/10.1017/S0020818309990142.
Goundar, S. 2021. "Introduction to Enterprise Systems and Technological Convergence." In *Enterprise Systems and Technological Convergence: Research and Practice*, edited by S. Goundar, 13–25. United States of America: Information age publishing, Inc.
Hammersley, M. 2008. *Advances in Mixed Methods Research*. 22–36.
Harley, J. M., F. Bouchet, and R. Azevedo. 2013. "Aligning and Comparing Data on learners' Emotions Experienced with MetaTutor." In *Lecture Notes in Computer Science Artificial Intelligence in Education*, edited by C. H. Lane, K. Yacef, J. Mostow, and P. Pavik, Vol. 7926, 61–70. Springer. https://doi.org/10.1007/978-3-642-39112-5_7.
Harley, J. M., F. Bouchet, M. S. Hussain, R. Azevedo, and R. Calvo. 2015. "A Multi- Componential Analysis of Emotions During Complex Learning with an Intelligent Multi-Agent System." *Computers in Human Behavior* 48 (1): 615–625. https://doi.org/10.1016/j.chb.2015.02.013.
Hine, C. 2015. "Mixed Methods and Multi-Modal Research and Internet Technologies." In The Oxford Handbook Of Multi-Method And Mixed Methods Research Inquiry, 503–521. Vol. *28*. New York, NY: Oxford University Press.
Husman, J., M. C. Graham, D. Chistensen, and I. Villanueva. 2019. "Prejudice and the Acceptance of Muslim Minority Practices: A Person-Centered Approach." *Society for Personality and Social Psychology* 51 (1): 1–16. https://doi.org/10.1027/1864-9335/a000380.
Järvelä, S., J. Malmberg, E. Haataja, M. Sobocinski, and P. A. Kirschner. 2021. "What Multimodal Data Can Tell Us About the students' Regulation of Their Learning Process?." *Learning & Instruction* 72: 101203.
Jewitt, C., E. Adami, A. Archer, A. Björkvall, and F. V. Lim. 2021. "Editorial." *Multi-Modality & Society* 1 (1): 3–7. https://doi.org/10.1177/2634979521992902.
Kleinginna, P. R., Jr, and A. M. Kleinginna 1981. "A Categorized List of Motivation Definitions, with a Suggestion for a Consensual Definition." *Motivation and Emotion* 5 (3): 263–291.
Köles, M. 2017. "A Review of Pupillometry for Human-Computer Interaction Studies." *Periodica Polytechnica Electrical Engineering and Computer Science* 61 (4): 320–326. https://doi.org/10.3311/PPee.10736.
Lachs, L. 2017. *Multi-Modal Perception*. Champaign, IL: DEF Publishers.
Lai, M. L., M. J. Tsai, F. Y. Yang, C. Y. Hsu, T. C. Liu, S. W. Y. Lee, S. W. Y. Lee, et al. 2013. "A Review of Using Eye-Tracking Technology in Exploring Learning from 2000 to 2012." *Educational Research Review* 10 (1): 90–115. https://doi.org/10.1016/j.edurev.2013.10.001.
Lazarus, A. A. 2005. "Multi-Modal Therapy." In *Handbook of Psychotherapy Integration*, edited by J. C. Norcross and M. R. Goldfried, 105–120. New York, NY: Oxford university press.
Lazarus, A. A. 2006. *Brief but Comprehensive Psychotherapy: The Multi-Modal Way*. New York, NY: Springer Publishing Company.
Ledin, P., and D. Machin. 2017. "Multi-Modal Critical Discourse Analysis." In *The Routledge Handbook of Critical Discourse Studies*, edited by J. Flowerdew and J. E. Richardson, 60–76. London: Routledge.
Leppänen, S., and S. Tapionkaski. 2021. "Doing Gender and Sexuality Intersectionally in Multi- Modal Social Media Practices." In *The Routledge Handbook of Language, Gender, and Sexuality*, edited by J. Angouri and J. Baxter, 543–556. Routledge. https://doi.org/10.4324/9781315514857.
Martin, P. A. 2016. "Tutorial Video Use by Senior Undergraduate Electrical Engineering Students." *Australasian Journal of Engineering Education* 21 (1): 39–47. https://doi.org/10.1080/22054952.2016.1259027.
Mills, K. A., and L. Unsworth. 2018. "The Multi-Modal Construction of Race: A Review of Critical Race Theory Research." *Language and Education* 32 (4): 313 332. https://doi.org/10.1080/09500782.2018.1434787.

National Science Foundation. 2016. *Convergence Research at NSF*. National Science Foundation. https://www.nsf.gov/od/oia/convergence/index.jsp.

Obaidellah, U., M. Al Haek, and P. C. H. Cheng. 2018. "A Survey on the Usage of Eye-Tracking in Computer Programming." *ACM Computing Surveys (CSUR)* 51 (1): 1–58. https://doi.org/10.1145/3145904.

Palmer, S. 2006. "Multi-Modal Therapy." In *The SAGE Handbook of Counselling and Psychotherapy*, edited by C. Feltham and I. Horton, 322–329. London: SAGE publications.

Stegmaier, P. 2009. "The rock' n' Roll of Knowledge Co-Production." *EMBO Reports* 10 (2): 114–119. https://doi.org/10.1038/embor.2008.253.

Stein, P. 2007. *Multi-Modal Pedagogies in Diverse Classrooms: Representation, Rights and Resources*. Routledge. https://doi.org/10.4324/9780203935804.

Tranel, D. 2000. "Electrodermal activity in cognitive neuroscience: Neuroanatomical and neuropsychological correlates." In *Series in Affective Science Cognitive Neuroscience of Emotion*, edited by R. D. Lane and L. Nadel, 192–224. New York, NY: Oxford University Press.

Turk, M. 2014. "Multi-modal interaction: A review." *Pattern Recognition Letters* 36:189–195. https://doi.org/10.1016/j.patrec.2013.07.003.

Van Gog, T., and K. Scheiter. 2010. "Eye Tracking as a Tool to Study and Enhance Multimedia Learning." *Learning and Instruction* 20 (2): 95–99. https://doi.org/10.1016/j.learninstruc.2009.02.009.

Villanueva Alarcón, I., E. M. Zorrilla, J. Husman, and M. Graham 2021. "Human-Technology Frontier: Measuring Student Performance-Related Responses to Authentic Engineering Education Activities via Physiological Sensing." Paper presented at the *International Conference on Applied Human Factors and Ergonomics*, United States of America.

Villanueva, I., and S. Anwar. 2022. "Situating Multi-Modal Approaches in Engineering Education Research." *Journal of Engineering Education* 111 (2): 277–282. https://doi.org/10.1002/jee.20460.

Villanueva, I., B. D. Campbell, A. C. Raikes, S. H. Jones, and L. G. Putney. 2018. "A Multi-Modal Exploration of Engineering students' Emotions and Electrodermal Activity in Design Activities." *Journal of Engineering Education* 107 (3): 414 441. https://doi.org/10.1002/jee.20225.

Villanueva, I., M. Di Stefano, L. Gelles, P. V. Osoria, and S. Benson. 2019. "A Race Re-Imaged, Intersectional Approach to Academic Mentoring: Exploring the Perspectives and Responses of Womxn in Science and Engineering Research." *Contemporary Educational Psychology* 59 (1): 1–17. article no. 101786. https://doi.org/10.1016/j.cedpsych.2019.101786.

Villanueva, I., W. Goodridge, N. J. A. Wang, M. Valladares, B. S. Robinsonu, and K. Jordan 2014. "Hormonal & Cognitive Assessment of Spatial Ability & Performance in Engineering Examination Activities." Paper presented at the *44th Annual Meeting of the Society of Neuroscience*, Washington, DC.

Villanueva, I., J. Husman, D. Christensen, K. Youmans, M. Khan, P. Vicioso, S. Lampkins, and M. C. Graham. 2019. "A Cross-Disciplinary and Multi-Modal Experimental Design for Studying Near-Real-Time Authentic Examination Experiences." *Journal of Visualized Experiments* 151 (1): 1–10. https://doi.org/10.3791/60037-v.

Villanueva, I., M. G. Valladares, and W. Goodridge. 2016. "Use of Galvanic Skin Responses, Salivary Biomarkers, and Self-Reports to Assess Undergraduate Student Performance During a Laboratory Exam Activity." *Journal of Visualized Experiments* 108 (1): 1–8. article no. e53255. https://doi.org/10.3791/53255.

Wert, E., J. Grifski, S. Luo, and Z. Atiq 2021. "A Multi-Modal Investigation of Self-Regulation Strategies Adopted by First-Year Engineering Students During Programming Tasks." Paper presented at the *17th ACM Conference on International Computing Education Research*, Virtual. https://doi.org/10.1145/3446871.3469795