

Sensorimotor learning in immersive virtual reality: a scoping literature review

Jack Ratcliffe

School of Electronic Engineering and Computer Science
Queen Mary University of London
London, UK
j.ratcliffe@qmul.ac.uk

Laurissa Tokarchuk

School of Electronic Engineering and Computer Science
Queen Mary University of London
London, UK
laurissa.tokarchuk@qmul.ac.uk

Abstract—The benefits and drawbacks of using immersive virtual reality (IVR) for learning are increasingly being explored, with growing evidence that a major contributor to IVR learning benefits are the sensorimotor-based affordances of the technology. However, to our knowledge, there have been no reviews of sensorimotor-based IVR learning studies for academic or cognitive learning. In order to provide an overview of the field, we present a scoping review based on a comprehensive search that identified 14 documents reporting on experimental sensorimotor-focused learning studies. The review found universally positive learning outcomes for sensorimotor-led approaches across a variety of topics and approaches, although noted multiple areas of difference and potential issues across studies, including differing measures for learning success, potential common confounding factors, a lack of longitudinal investigations, a lack of a common methods for surveying sensorimotor engagement in IVR, and a disconnect between researchers in this area.

Index Terms—Virtual reality, Human computer interaction, Educational technology, Cognition

I. INTRODUCTION

The benefits and drawbacks of using immersive virtual reality (IVR) for learning are increasingly being explored. Common explanations for why IVR might provide learning benefits point towards high levels of immersion, spatial presence, or motivation in participant learners [1] [2] [3] [4] [5] [6]. There is growing evidence, however, that a major cause of the benefits of IVR learning could be in allowing learners to enact their learning [7] [4], particularly in ways which leverage the body’s natural sensor and motor apparatus (i.e. a sensorimotor-based approach) [8]. These approaches stem from “enactive cognitive science”, which present cognition as an “interconnected system of multiple levels of sensorimotor subnetworks” [9].

While there have been multiple reviews into IVR and learning [10] [5] [11] [12] [13], to our knowledge, there have been no attempts to specifically synthesize the divergent approaches and interests of sensorimotor-based IVR learning interventions in order to present an understanding of the research area.

A review of studies exploring sensorimotor IVR learning is important, as the studies vary across numerous factors. These include (1) which learning topics are being explored; (2) how are (and what type of) sensorimotor approaches leveraged, and what hardware is enabling them; (3) what theoretical justifications or motivations are behind these decisions; (4)

what experimental process are used and how is learning gain being measured; (5) what learning results have been uncovered; and (6) what other experiential measures are also being monitored. This scoping review, therefore, seeks to fill that gap and present an overview of the characteristics listed above.

Our comprehensive search identified 14 documents reporting on experimental sensorimotor IVR learning studies which universally presented positive learning results. However, we also identified questions around differing definitions of learning gain, potential confounds (such as situation and context); and an absence of longitudinal or holistic learning studies, of sensorimotor-orientated embodiment questions, of discussion of how different IVR hardware might impact results, and of connections between sensorimotor research.

II. BACKGROUND

A. Embodied cognition and sensorimotor learning outcomes

Learning benefits linked to embodied cognition and sensorimotor-based encoding have been described in literature for almost 100 years [14]. Contemporary research has many examples of experiments where a group learning in a highly embodied scenario or with sensorimotor processes out-perform a lowly embodied control across a variety of learning subjects, including language [15] [16] [17] [18] [19] [20] [21], STEM subjects [22] [23] [24] [25] [26] [27] [28] and learning related-skills [29]. The breadth of research into the impact of embodied approaches to learning have mostly demonstrated positive learning results, with the majority of the above research focused on embodiment in terms of the sensorimotor or “action”, rather than the environmental aspect of embodied cognition.

Commonly used learning techniques already leverage embodied and sensorimotor mechanisms, such as counting on hands, which has been shown to positively influence children’s and adults’ numerical processing [25], and leads to better understanding and application of more advanced concepts later in their learning process [26]. The belief that enhanced embodiment leads to better learning is not universal, however, and some argue that embodied learning approaches can either be not optimal or detrimental to learning outcomes [30] [31].

There is also physiological evidence that sensorimotor processes are tightly linked to learning and cognition. Brain responses to sensorimotor activity and learning have shown that areas of the brain responsible for the motor system are activated during language comprehension, suggesting a link between reading words and activity in parts of the brain associated with engaging in that action. For example, reading action words (like kick or throw) activate areas of the brain responsible for carrying out those actions [32], while reading odor words (like the names of spices) prompt olfactory-related brain activity [33].

Sensorimotor benefits have also been shown to exist in both non-immersive and immersive virtual reality. Participants in a Kinect-based condition significantly outperformed their counterparts in the less embodied mouse-based condition immediate recall and retention recall tests [29]; Edge [34] found users enacting a sequence of movements to complete a foreign-language movement instruction performed better than a control; Macedonia [35] had participants imitating a pedagogical agent's gestures and visually successfully learn words accompanied by gestures; and Repetto [36] found that when recognizing novel words, participants made less errors for words encoded with gestures compared to words encoded with pictures.

More immersively, a Mixed Reality physics learning study saw embodied students improved their performance by 76% on the second trial compared with 51% for those who used the simulation without bodily cues [27]; and in an IVR experiment, learning was significantly better for users who had to move their body than not [28]. Additionally, students in the "high embodiment" condition retained their learning better when retested a week later.

B. Existing IVR learning reviews

Previous systematic reviews and meta-analyses into IVR and learning provide some useful definitions of different learning-related scopes of IVR studies. These were scopes of IVR hardware (from CAVEs, to three degree-of-freedom HMDs, to six-degree-of-freedom HMDs); scopes of learning (including vocational training, cognitive and education, skill-improvement, empathy enhancement); and scopes of sensorimotor engagement.

The broadest IVR learning meta-analysis, in terms of scope of IVR, learning and sensorimotor engagement, is Howard's meta-analysis of virtual reality hardware and software for personal development [10]. It found no evidence that input hardware, which includes but is not limited to, sensorimotor inputs, had a significant effect on cognitive development. Howard presents four potential explanations for this: (1) input hardware may have little impact on mechanisms that may subsequently influence learning outcomes; (2) current input hardware influences important mechanisms, but input hardware may not influence these mechanisms enough; (3) input hardware may influence mechanisms that have little effect on IVR intervention outcomes compared with other, more important IVR aspects; and (4) specialized input hardware may

not influence the nature of tasks enough to incur a substantive effect on outcome.

If we subscribe to the belief that embodied cognition could play a role in learning development whatever the medium, then these questions can be re-summarised as: does existing IVR hardware and design enable embodied cognition in a form that provides comparable-or-better learning outcomes to physical-world embodied approaches?

A potential limitation in applying Howard's findings to the study of embodied sensorimotor interaction in IVR is that his scope of IVR and learning were both broad. The IVR studies he examined took a variety of less embodied and sensorimotor-enabled forms, including CAVEs or three degree-of-freedom HMDs; and his definition of cognitive learning included both academic subjects, vocational training and other forms of personal development. Other reviews and meta-analyses approached IVR learning with more tightly defined scopes, which will be discussed below.

C. Scopes of IVR hardware

Many of the existing reviews of IVR learning are focused on HMDs, rather than CAVEs, and limit their explorations to HMDs from 2016 or newer [5] [11] [12]. The reasoning for this was due to their "relatively low cost of entry and improved practicality" compared with CAVE environments; and their improved performance compared with the field-of-view limited HMDs used before 2016 [5]; or to a timeline when HMDs gained popularity and may have changed the way IVR is used in the education context [11].

There has been little distinction, however, between three degree-of-freedom IVR and six degree-of-freedom. This is an important distinction for enabling sensorimotor interactions, as six-degrees of freedom allows for navigation around a physical space and much richer bodily tracking.

D. Scopes of IVR learning

A common distinction in existing IVR learning reviews is between academic education and vocational training [13] [11] [12] [5]. Checa and Bustillo [13] noted that this was often reflected in how studies were being evaluated: the majority of studies evaluating training focused on task performance, whereas the education-dominated ones evaluated knowledge acquisition. This approach suggests that difference scopes of learning should be, and often are, treated differently. There were also distinctions in the types of academic education, with one reviews exploring studies of IVR learning in high-education contexts [11], with another focused on K-12 and higher educational subject-learning contexts [12]. Jensen & Konradsen's review found six studies, out of 21 they discussed, that specifically investigate academic education rather than vocational training [5].

E. Sensorimotor engagement and embodiment

The impact of sensorimotor engagement in IVR educational learning has not been thoroughly investigated. Jensen & Konradsen identified five studies that leveraged sensorimotor input

[5], but these were all concerned with skills acquisition - such as juggling - rather than cognitive learning. The review concluded that the prevailing question was not if HMDs should be used, but rather how and for what should HMDs be used. They also theorised that IVR learning could work successfully with educational approaches and theories such as Constructivism and active learning.

Natale et al. do not specifically mention sensorimotor engagement, but do discuss an idea similar to that of embodiment [12]. They conclude their study by mentioning that the possibility for users to feel present in IVR, to use their bodies in a natural way, and to live sensory experience similar to those in the real world could be promising as advanced learning instructional strategies, and call for further research into how these aspects could affect the learning affordances compared to less immersive ones.

Radianti et al. presented a thorough summary of the main ideas behind the existing IVR learning paradigms, including Behaviorism, Cognitivism, Constructivism, Experientialism, Connectivism, but do not specifically refer to a sensorimotor approach [11].

F. Summary

The above reviews found that IVR can support experiences that enable learning and motivate students to fulfill educational goals. Increasingly, it also seems that modern IVR HMDs, rather than CAVE systems, are considered the best place to examine this type of benefit. These reviews also present a distinction between academic education and vocational training.

There is, however, a lack of investigation into the learning possibilities and benefits afforded by leveraging sensorimotor techniques in IVR learning, especially for academic learning. While Howard touches on the area with "input hardware", his search is more interested in the hardware itself rather than its application in the learning context. Similarly, for Radianti et al. it is not mentioned in their list of identified learning paradigms. Only Natale et al. refers to the possibilities of sensorimotor-based approaches, although this is in the limitations section of their own research.

This lack of investigation, coupled with the growing number of investigations into sensorimotor IVR academic learning, suggests a dedicated review is needed.

III. SCOPING REVIEW

Our scoping review was conducted using an approach informed by Xiao and Watson [37] and consists of three stages: (1) formulating the research problem and developing the review protocol; (2) searching the literature, screening for inclusion, extracting data, analyzing and synthesizing data; and (3) reporting the results.

As this was a scoping review, and so attempted to discuss the breadth of the field, there was no requirement to formally assess paper quality [37]), however all included papers had passed a formal peer-review.

A. Formulating the research problem and developing the review protocol

Our initial proposal was to scope the field of experimental research concerning learning using sensorimotor techniques in IVR, reviewing the efficacy, experimental methodologies, theoretical approaches and hardware-use related to sensorimotor-led learning experiments inside IVRs.

Our goal was to provide a complete overview of experimental research into sensorimotor-led learning inside IVR in an attempt to identify a conceptual boundaries of the field, the size of the pool of research, types of available evidence, and any research gaps.

However, during pre-mapping we reduced the scope from all learning to only academic (or non-vocational) learning.

B. Pre-mapping

In order to understand if our area of investigation was feasible, we used a pre-mapping approach [38] to identify potential subtopics within our proposed research area. To explore this broad area, we used Google Scholar and the following search term:

"virtual reality" OR "virtual environment" OR "virtual simulation" OR vr AND immersive AND learning AND sensorimotor

As expected, Google Scholar provided many results (n = 5600). We analysed the first 250 abstracts of these results and determined an important taxonomy of subtopics: type of learning. We categorised the types of "learning" returned as being cognitive (or semantic, knowledge-based learning); skill-training; physical or mental recovery; and affective exercises (e.g. encouraging empathy or reducing anxiety). These categories are similar to those identified in Jensen & Konradsen's review of the IVR learning space, and map to the definitions and distinctions between academic and training found in other reviews.

To reduce the size of the investigation, and to align with our research interests, we were conscious to only include cognitive learning in our research questions.

C. Research question development

Our research questions were informed by criteria from previous systematic literature reviews related to IVRs and learning. In Feng et al.'s [6] IVR serious games systematic literature review, they determined two primary types of research question: pedagogical impact, concerned with learning outcomes and measures; and the behavioral impact, concerned with behavioural outcomes and measures. From the outcomes perspective, Natale [12] also recorded learning measures, learning result, effect size, motivation measure, motivation result, effect size.

Feng also presents a further question "what are the essential elements for developing IVR serious games?", with five sub-questions relevant to this study: what are the teaching meth-

ods, navigation solutions, senses simulated, narrative methods, NPC contributions.

To add sensorimotor-specific questions, we also explored whether the system was HMD or CAVE, the type of embodied controls; the conditions the study was comparing between; whether it was one-off or across multiple sessions; was it standalone or as part of a pedagogical tool; how was the data recorded; and what IVR-related experiential measures were used.

These were summarised into the following core research questions:

- What areas of study (topics) are being explored for cognitive sensorimotor learning in IVRs?
- What theoretical learning or teaching approaches are being used to justify these approaches, and do they depend on the topic
- How are these studies being conducted
- How is (or what type of) sensorimotor embodied is being used?
- What hardware is being used to enable this embodiment?
- What objective learning results (efficacy) have these studies demonstrated?
- What subjective learning experiences have these studies demonstrated?

D. Developing the review protocol

1) *Selection criteria:* In order to include experimental research concerning cognitive learning using sensorimotor techniques in IVRs, we outlined the following criteria for paper inclusion:

- Use immersive virtual reality (not just virtual reality)
- Explore cognitive, semantic or knowledge-based learning (not studies on skill-development, physical and mental recovery, and affective changes like anxiety or empathy)
- Designed to explore relationship between sensorimotor activity and outcomes
- Use a HMD-enabled immersive environment (not CAVE)
- Use controllers that enable some form of bodily input with the environment and learning situation beyond the user interface menus, including physical and camera-based methods
- Be experimental or quasi-experimental
- Report either objective (learning change) or subjective (experiential) measures
- Full text is accessible and available
- Full text is in English

2) *Search strategy:* Our strategy includes all peer-reviewed research documenting experimental or quasi-experimental studies related to the topic. We identified five research databases, two interdisciplinary (SCOPUS, Web of Science), and others in computer science and engineering (IEEE Xplore), education (ERIC) and psychology (PsycINFO).

As high field-of-view HMD IVR systems with embodied controls were not widely available prior to the release of modern commercial headsets, such as the HTC Vive and

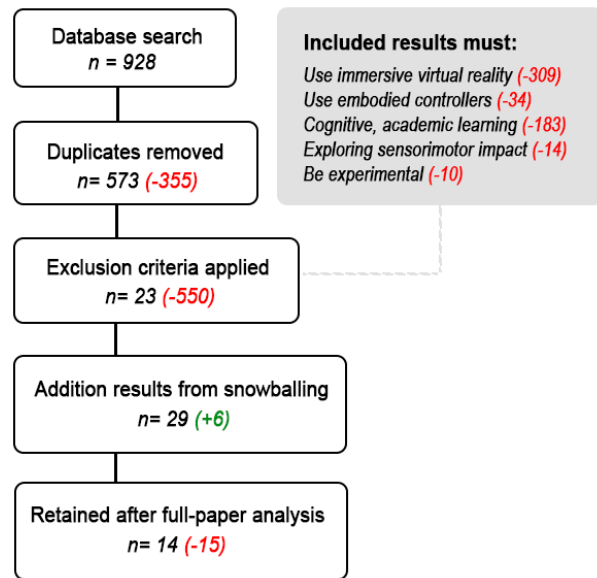


Fig. 1. Flowchart depicting the article selection process

Oculus Rift; we limited results to those published in 2016 or after (as is noted in other meta-reviews).

Each database was searched in March 2021 with keywords based on this Boolean search string:

("virtual reality" OR "virtual environment" OR "virtual simulation" OR vr OR "head-mounted display" OR "immersive environment") AND ("learning" OR "training" OR "education") AND ("sensorimotor" OR "kinematic" OR "embodied")

Papers that passed the inclusion criteria were then snowballed exhaustively in both directions, exploring papers from their reference list; and using Google Scholar to identify papers that had referenced them.

3) *Data collection:* The database listings and search results were achieved for record keeping, reproducibility, and cross-checking [39].

Each paper's abstract was examined to determine if it matched the above criteria, and if it was unclear, the conclusion section was also read [38]. If it was still unclear, studies were included for further study. The reason for exclusion was recorded for each paper.

IV. SEARCHING THE LITERATURE, SCREENING FOR INCLUSION, EXTRACTING DATA, ANALYZING AND SYNTHESIZING DATA

A. Searching the literature and screening for inclusion

Our search uncovered 928 results (ERIC, 18; PsycINFO, 30; IEEE Xplore, 118; SCOPUS, 470; Web of Science, 292). Of these, 573 remained after removing duplicates. The following removals occurred at each stage of our criteria: [A] Must use immersive virtual reality (not just virtual reality), must

use a HMD-enabled immersive environment (not CAVE) (309 removals). [B] Use controls that prioritise sensorimotor input as part of the learning process (and so behold controller or keyboard button-pressing) (34 removals). [C] Explore cognitive, semantic or knowledge-based learning (not studies on skill-development, physical and mental recovery, and affective changes like anxiety or empathy) (183 removals). [D] Designed to explore relationship between sensorimotor activity and outcomes (14 removals). [E] Be experimental or quasi-experimental and report either objective (learning change) or experiential measures (10 removals).

This left 23 papers. We snowballed references backwards and forwards and identified a further six papers. This left 29 papers for a full reading for inclusion. At this stage, we removed a further 15 papers due to the following three categories: the paper was reporting on the same experiment as another included paper; there were not recorded experimental results; or the target was not cognitive learning. This left 14 eligible papers [40] [41] [42] [43] [44] [45] [46] [47] [48] [49] [50] [51] [52] [53]. This process is illustrated in Fig. 1, and the list of included papers can be found in the appendix.

B. Extracting, analyzing and synthesizing data

We used an inductive approach to categorisation, which resulted in the research questions listed above. Responses were thematically analysed using an inductive approach based upon Braun and Clarke’s six phases of analysis [8].

V. RESULTS

A. Topics

The predominant learning topic investigated was second language education ($n = 7$); specifically word memorisation (6). Of the papers, five included noun memorisation and two included verb memorisation.

There was a diverse, if limited, range of other topics present in the papers: biology (2), mathematics (1), physics (1), computer science (1) and geography (1).

There was also a single investigation into a sensorimotor-led cognitive skill development that could provide benefits across multiple learning subjects: perspective taking, or “the ability to mentally represent a [visual] viewpoint different from one’s own”.

B. Theoretical approaches

All but one of the papers referenced a theoretical approach behind their motivation to investigate sensorimotor-based IVR learning. The most evident theory was that of embodied cognition, which was also referred to as embodied learning and learning from embodied interaction (7). Additionally, four (of the six) investigations into language learning referenced language embodiment theories.

Three papers referred to approaches that are related to embodiment, although not sensorimotor embodiment specifically. These were (1) spatial cognition, the importance of using the environment around you for meaning making, storage and processing; (2) Constructivist, specifically Interactionist, in

which interactions (but not necessarily sensorimotor ones) play an important role in learning, and (3) situated learning, which includes aspects of both spatial cognition and Interactionist perspectives.

One paper justified its approach from a motivation-perspective, suggesting that the motivation benefits of IVR interaction should cause enhanced learning outcomes.

C. Research motivations

There was a notable distinction between the research motivations for language learning investigations and those of other topics. Four of the language learning investigations were motivated by a desire to provide empirical evidence for an existing theory of embodied benefits to language learning; and whether these extend into IVRs.

The major justification for non-language studies was to improve on perceived shortcomings of real-world sensorimotor teaching approaches, including practicality and cost. More detailed examples of these include whether a sensorimotor IVR experience could adequately replace the experience of re-constructing canine skeletons (compared with a real box of bones); whether seeing inside a hand can improve understanding of hand biology over cadaver dissection (and also remove the requirement of having access to a severed hand); or rebalance gender-related interest in Computer Science learning through more a active and interesting approach.

Three papers cited the motivation to see if a sensorimotor approach could maintain learning whilst also improving motivation and interest in the subject.

D. Experiment methodologies

The studies presented five methodologies for measuring learning gain. The most popular approach was a pre- and post-exposure test of participants’ knowledge, with the performance change between the two measures showing the learning gain (6). Other approaches examined the participants’ response times to questions, with quicker responses presented as evidence of better learning outcomes. One study monitored participant movement via HMD data to understand if the user was performing an action that represented the target learning (bowing at the start of interactions as a cultural Japanese learning process). Another monitored performance in a game to determine progress of the learning material - better game performance being indicative of a better desired learning outcome. Finally, one study simply asked participants to report if they felt they had achieved better learning via the IVR system.

Eleven of the studies compared the sensorimotor IVR condition with a control, although the types of control varied. The most prominent comparison was with a non-sensorimotor IVR alternative (3). There were also comparisons with incongruent sensorimotor activation (2); non-IVR sensorimotor interactions (2, such as a physical skeleton arrangement task versus a virtual one); non-IVR and non-sensorimotor digital setting (2); low-interaction (flashcards); non-interaction (watch-only) and AR embodiment.

One study compared non-tangible IVR sensorimotor activation with a tangible version, using a bespoke device to provide physical feedback to the virtual interaction.

E. Types of sensorimotor

The studies used a variety of different approaches for enabling participants to interact in a sensorimotor-embodied way. These could loosely be broken down into two categories: (1) sensorimotor interaction with objects in an environment, and (2) sensorimotor-as-input method. For the former, participants would manipulate 3D objects in a contextual setting with varying levels of feedback. For the latter, participants moved their bodies in novel ways and the system would recognise and react to these.

Among studies engaging sensorimotor interaction with the environment, the dominant form was allowing participants to pickup virtual objects, move them around and conduct actions with them. For example, picking up a jug and pouring it so water comes up in order to learn the verb "to pour" or the noun "jug".

Two studies allowed participants to pick up and move objects, but do not allow them to be used in an interactional way with the environment. For example, a jug could be picked and moved, but not poured with an interactive outcome from the system.

One further study allowed users to pick and move objects, but rather than provide a real-world equivalent action, instead presented an outcome only possible in IVRs. In this case, users could pick up bones and move them in the air, where they stayed suspended as the used constructed a zero-gravity canine skeleton.

The types of activation in the sensorimotor-as-input method studies were more varied. In one, participants would move their own hand to see a digitised version (and an enlarged digitised version) move in the same way but with added visualisations, to demonstrate how the typically hidden hand tendons worked. Another, using a custom suspension-rig, allowed participants to experience a situation similar to micro-gravity in order for them to experience how the related forces work. One studied used dance-based input in order to organise a virtual computer programming experience, while another had users slash words with virtual lightsabers in order to make active selections of multiple options.

Among these types of sensorimotor implementation, there was an even split between types of head-mounted display (HTC and Oculus), with bodily inputs coming from HTC "wands" (4), hand-tracking (2), bespoke hardware (2), camera-based full-body tracking (1), head movement (1) and Oculus "touch" controllers.

F. Learning efficacy

The conclusions of all of the examined papers supported the use of IVR sensorimotor approaches for learning. However, the criteria used to justify that support varied between studies. The criteria was either: (1) learning occurred, as opposed to no learning or negative learning; (2) learning occurred at a

TABLE I

SHOWING THE NON-LEARNING MEASUREMENTS USED IN THE STUDIES

Type	Measure details
Usability	2x SUS, 1x MEEGA+, 3x Custom/not given
Sickness & Comfort	1x SSQ, 1x SSQ adaptation, 3x Custom/not given
Presence	2x IPQ, 1x Nowak & Biocca, 1x Custom/not given
Motivation	1x MEEGA+, 1x Custom/not given
Enjoyment	2x Custom/not given
Embodiment	1x Gonzalez-Franco
Task Load	1x NASA TLX
Flow	1x Custom/not given
Intrinsic Interest	1x Custom/not given
Concentration	1x Custom/not given
Satisfaction	1x Custom/not given
Enthusiasm	1x Custom/not given
Interest in subject	1x Custom/not given
Perceived Usefulness	1x Custom/not given
Perception of Time	1x Custom/not given
Fun	1x Custom/not given
Preference	1x Custom/not given
None	3x no measures presented

System Usability Survey (SUS) [54], Model for the Evaluation of Educational Games (MEEGA+) [55], Simulator Sickness Questionnaire (SSQ) [56], Igroup Presence Questionnaire (IPQ) [57], Task Load Index (NASA TLX) [58], Nowak & Biocca [59], Gonzalez-Franco [60]

similar rate to the control; or (3) learning out-performed the control.

The majority of studies defined "learning occurring" as knowledge gain based on pre- and post- test result changes. However, one study examined the speed of participant responses, with faster response speeds demonstrating improved learning. Another study measured learning by the self-reported "learning experience", i.e. did participants feel they had learned during the experience?

There was also variance between studies around how many tests showed learning gain. Of the papers, six noted significant learning benefits in the IVR sensorimotor condition across all of their reported learning gain measures.

However, two noted learning gain benefits only on some of their reported indicators of learning gain. Interestingly, these two studies were both on word memorisation, and there were contrary outcomes between them. One showed only a significant benefit of sensorimotor encoding in an immediate test, but not in a retention test taken one-week later; while the other only showed a significant in a retention test taken one week later, and not in the immediate test.

Two studies noted similar results between the IVR sensorimotor condition and an alternative.

No papers reported learning gains worse than the control in all reported conditions.

G. Experiential measures

The studies featured a wide range of experiential measures (see Table 1). The most commonly used measures were sickness/comfort, usability and presence. Broadly, the measures could be categorised into the following categories: properties strongly linked to IVR (presence, embodiment, sickness); measures of the system usability (usability); general engagement properties (enjoyment, excitement, motivation, interest/subject

interest, preference); and experience in the session (flow, NASA task load index, concentration).

Interestingly, there was only one study that specifically measured IVR embodiment [42], using the Gonzalez-Franco and Peck embodiment questionnaire [60]. We found no questions regarding experiences of the sensorimotor activity.

VI. DISCUSSION

We made the following observations about different aspects of the surveyed papers, and thus, we believe, the field of sensorimotor IVR investigations for academic learning.

A. Conceptual consensus and confounds

Unsurprisingly, the IVR sensorimotor investigations found were predominantly grounded within embodied theoretical approaches. Perhaps more interesting are the studies that grounded their explorations in other conceptual approaches, as these could help highlight potential mediating factors for investigations looking to support or repudiate sensorimotor and embodied cognition approaches in IVR (and also the inverse).

One of these approaches was motivation [50], which as a factor in improving learning outcomes is well-established [61] [62]. There is also evidence that sensorimotor learning activities increase motivation [63] [64] [65] [66] [67]. However, the idea that sensorimotor interaction IVRs prompting motivation to be the key learning driver of these interventions, is unproven. Indeed, Ratcliffe [46] found no evidence of motivation serving as a mediating factor for embodied learning.

Another approach was situated learning [48] [49]. Despite being contextualised by their authors inside a situated cognition framework, these studies deeply deployed sensorimotor interactivity; one in which physically bowing was a key interactive input [48], while the other involved pointing-and-grabbing objects [49]. In fact, some degree of situatedness (such as realistic and congruent learning contexts for interactions), was part of all of the surveyed IVR experiences.

The risk of situatedness being a confound for a pure sensorimotor vs. non-sensorimotor comparison was present in the majority of studies, as only some of the papers compared (congruent) sensorimotor interaction against either non-sensorimotor or non-congruent sensorimotor interaction in the same IVR [40] [45] [46] [51].

If we are to truly understand the contribution of sensorimotor interaction, we need to be sure we are controlling other IVR-induced variables, especially those that have a theoretical backing as potentially contributing to learning outcomes.

B. Language dominance

There was a clear language dominance in the studies we surveyed. The prevalence of language acquisition in these studies is likely due to the long history [68] and encouraging results [69] of using sensorimotor teaching approaches in applied linguistics outside of IVR. The ability for sensorimotor engagement to enhance word memorisation has been widely explored (and evidenced) in this field, and so exploring

whether these benefits transfer into embodied virtual interactions is a logical next-step.

That said, even these studies are not close to giving us a clear understanding of the impact of sensorimotor interaction on language learning as a whole. Word memorisation is an important but subsidiary part of what is considered language teaching or second language acquisition, and we need to be able to move beyond these initial examinations if we are to be able to understand if and how sensorimotor engagement is beneficial for the myriad of language skills (although we have one example of this here, with the attempt to teach appropriate bowing etiquette [48]).

C. What is "success"?

While all papers reported successful learning outcomes from their sensorimotor approach, the definition of what success was varied across papers. Learning improvement over a control; learning improvement in a specific condition over a control; the same performance as a control; testing response speed; "experience" or motivation improvements were all presented as evidence for the success of a sensorimotor approach.

For studies that use multiple measures of success, such as retention tests after a set period of time, or knowledge gain and answer speed, it is unclear how contradictory results between different tests should be used to promote or refute a learning "success".

While the definition of success is certainly learning intervention-specific, perhaps it would be useful for the field to define types of success or potential metrics for success, in order to allow thorough future meta-analyses of the field.

D. Learning "session" variation

The prevalence of word memorisation studies unveiled another variable that might benefit from codification and standardisation in future - the scale of the learning to be achieved in an intervention. The discrepancies between amount of learning material are evident in the language-based studies, which asked participants to learn either six [40], 12 [49], 14 [51], 20 [46] [50] [53] words. Potentially, six words might be too few to detect a difference in learning outcomes (although in this case, response time, not accuracy, was the performance indicator), while 20 may introduce too much confusion for optimal learning performance. This difference also adds another confounding factor for cross-analysis.

E. Longitudinal lacking

None of the studies surveyed were longitudinal, nor involved a longer curriculum of learning. These is a known problem with experimental learning studies [70], and one that might be exasperated by the additional development requirements that sensorimotor IVR studies require. With all the papers reporting some form of learning gain from engaging these systems, however, a longitudinal study seems like a logical and pressing next-step to understand how useful IVR sensorimotor learning is for practical teaching applications.

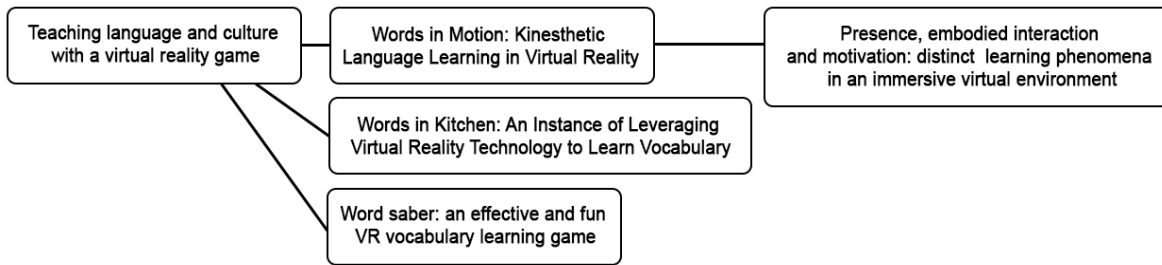


Fig. 2. Diagram depicting the lack of referencing between papers found in review. Three papers referenced "Teaching language and culture with a virtual reality game"; one further paper referenced one of those

F. Sickness and presence above others

There was a wide variety of experiential measures used across the studies. The frequency of measures of sickness and presence show that these two factors are starting to become well-established as important aspects of IVR studies. A notable absence, however, was any form of measurement of sensorimotor activity, and the single measurement of embodiment. Perhaps these studies would have benefited from surveying how embodied or sensorimotor-activated the participants felt in their sessions. Questionnaire-based embodied surveys for IVR are now emerging and being validated [60], although it remains to be seen if they are suitable for measuring sensorimotor interactions in IVR.

G. Hardware hiccups

The surveyed studies featured a range of sensorimotor-interaction enabling hardware, which could introduce another confound. Although it unclear how the hardware enabling sensorimotor input affects input, experience or learning, initial research has found that these devices are not neutral in their impact. For example, one study found that that the Leap Motion hand tracker is preferred to the HTC wand when grabbing close objects due to its natural affordances; but also resulted in much slower performance at the selection of distance objects [71]. How we might deal with the impact of the affordances of the hardware are not clear, but studying, understanding and noting the impact of these at a system and experiment design stage could be beneficial.

H. An unconnected field

Finally, it was notable how few of these papers referred to each other. Among the papers, there were only four references to other studies found in this selection. Three of these [53] [50] [49] referred to one paper [48], while the final reference was to one of these three ([46] referencing [53]) (see Fig. 2). These references were all contained within studies investigating language learning; there were no cross-learning topic references.

This suggests that IVR sensorimotor learning has not yet established a central canon, which could bring challenges to researchers looking for similar works, understanding their outcomes and exploring and critiquing other research procedures and approaches.

VII. CONCLUSION

Our review has found that sensorimotor-based IVR learning explorations are happening across a variety of subjects, backed predominantly by theories of embodied cognition. We found two major justifications for the research: (1) to provide empirical evidence for an existing theory of learning; and (2) to understand if IVR is a suitable replacement or improvement on physical-world teaching. We also found a variety of measures of learning, as well as a vast selection of experiential variables being measured.

From these results, we summarised a number of shortcomings for the field, which resulted in suggestions for further research for future sensorimotor-based IVR learning practitioners. These are:

- (1) be mindful of the potential confound of situation and context (situation as location-in-space, whereas context is more cognitive);
- (2) engage in more longitudinal or holistic topic learning, rather than singular interventions;
- (3) define measures of success, including beyond learning gain;
- (4) identify viable learning outcomes for interventions;
- (5) establish and engage sensorimotor-orientated embodiment questionnaires;
- (6) understand and discuss how the chosen IVR hardware might impact results;
- (7) more strongly engage with other research in the field, potentially across traditional discipline divisions

REFERENCES

- [1] A. Dengel and J. Mägdefrau, "Immersive learning explored: Subjective and objective factors influencing learning outcomes in immersive educational virtual environments," in *2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*. IEEE, 2018, pp. 608–615.
- [2] M. Kozhevnikov, J. Gurlitt, and M. Kozhevnikov, "Learning relative motion concepts in immersive and non-immersive virtual environments," *Journal of Science Education and Technology*, vol. 22, no. 6, pp. 952–962, 2013.
- [3] R. Coulter, L. Saland, T. Caudell, T. E. Goldsmith, and D. Alverson, "The effect of degree of immersion upon learning performance in virtual reality simulations for medical education," *InMedicine Meets Virtual Reality*, vol. 15, p. 155, 2007.
- [4] N. Pellas, A. Dengel, and A. Christopoulos, "A scoping review of immersive virtual reality in stem education," *IEEE Transactions on Learning Technologies*, vol. 13, no. 4, pp. 748–761, 2020.

- [5] L. Jensen and F. Konradsen, "A review of the use of virtual reality head-mounted displays in education and training," *Education and Information Technologies*, vol. 23, no. 4, pp. 1515–1529, 2018.
- [6] Z. Feng, V. A. González, R. Amor, R. Lovreglio, and G. Cabrera-Guerrero, "Immersive virtual reality serious games for evacuation training and research: A systematic literature review," *Computers & Education*, vol. 127, pp. 252–266, 2018.
- [7] G. Makransky, N. K. Andreasen, S. Baceviciute, and R. E. Mayer, "Immersive virtual reality increases liking but not learning with a science simulation and generative learning strategies promote learning in immersive virtual reality," *Journal of Educational Psychology*, 2020.
- [8] M. C. Johnson-Glenberg, "Immersive vr and education: Embodied design principles that include gesture and hand controls," *Frontiers in Robotics and AI*, vol. 5, p. 81, 2018.
- [9] F. J. Varela, E. Thompson, and E. Rosch, *The embodied mind: Cognitive science and human experience*. MIT press, 2016.
- [10] M. C. Howard, "Virtual reality interventions for personal development: A meta-analysis of hardware and software," *Human-Computer Interaction*, vol. 34, no. 3, pp. 205–239, 2019.
- [11] J. Radianti, T. A. Majchrzak, J. Fromm, and I. Wohlgenannt, "A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda," *Computers & Education*, vol. 147, p. 103778, 2020.
- [12] A. F. Di Natale, C. Repetto, G. Riva, and D. Villani, "Immersive virtual reality in k-12 and higher education: A 10-year systematic review of empirical research," *British Journal of Educational Technology*, vol. 51, no. 6, pp. 2006–2033, 2020.
- [13] D. Checa and A. Bustillo, "A review of immersive virtual reality serious games to enhance learning and training," *Multimedia Tools and Applications*, vol. 79, no. 9, pp. 5501–5527, 2020.
- [14] J. Dewey, "My pedagogic creed," *Journal of Education*, vol. 101, no. 18, pp. 490–490, 1925.
- [15] A. M. Glenberg, T. Gutierrez, J. R. Levin, S. Japuntich, and M. P. Kaschak, "Activity and imagined activity can enhance young children's reading comprehension," *Journal of educational psychology*, vol. 96, no. 3, p. 424, 2004.
- [16] A. M. Glenberg, A. B. Goldberg, and X. Zhu, "Improving early reading comprehension using embodied cai," *Instructional Science*, vol. 39, no. 1, pp. 27–39, 2011.
- [17] A. Glenberg, J. Willford, B. Gibson, A. Goldberg, and X. Zhu, "Improving reading to improve math," *Scientific Studies of Reading*, vol. 16, no. 4, pp. 316–340, 2012.
- [18] M. P. Kaschak, C. M. Connor, and J. L. Dombek, "Enacted reading comprehension: using bodily movement to aid the comprehension of abstract text content," *PLoS one*, vol. 12, no. 1, p. e0169711, 2017.
- [19] L. Q. Allen, "The effects of emblematic gestures on the development and access of mental representations of french expressions," *The Modern Language Journal*, vol. 79, no. 4, pp. 521–529, 1995.
- [20] M. Macedonia and T. R. Knösche, "Body in mind: How gestures empower foreign language learning," *Mind, Brain, and Education*, vol. 5, no. 4, pp. 196–211, 2011.
- [21] M. Macedonia, "Embodied learning: Why at school the mind needs the body," *Frontiers in Psychology*, vol. 10, p. 2098, 2019.
- [22] R. E. Núñez, L. D. Edwards, and J. F. Matos, "Embodied cognition as grounding for situatedness and context in mathematics education," *Educational studies in mathematics*, vol. 39, no. 1-3, pp. 45–65, 1999.
- [23] D. H. Clements, "'concrete' manipulatives, concrete ideas," *Contemporary issues in early childhood*, vol. 1, no. 1, pp. 45–60, 2000.
- [24] T. Martín and D. L. Schwartz, "Physically distributed learning: Adapting and reinterpreting physical environments in the development of fraction concepts," *Cognitive science*, vol. 29, no. 4, pp. 587–625, 2005.
- [25] A. Badets and M. Pesenti, "Creating number semantics through finger movement perception," *Cognition*, vol. 115, no. 1, pp. 46–53, 2010.
- [26] J. Boaler, L. Chen, C. Williams, and M. Cordero, "Seeing as understanding: The importance of visual mathematics for our brain and learning," *Journal of Applied & Computational Mathematics*, vol. 5, no. 5, pp. 1–6, 2016.
- [27] D. Abrahamson and R. Lindgren, "Embodiment and embodied design," 2014.
- [28] M. C. Johnson-Glenberg, C. Megowan-Romanowicz, D. A. Birchfield, and C. Savio-Ramos, "Effects of embodied learning and digital platform on the retention of physics content: Centripetal force," *Frontiers in psychology*, vol. 7, p. 1819, 2016.
- [29] K.-J. Chao, H.-W. Huang, W.-C. Fang, and N.-S. Chen, "Embodied play to learn: Exploring kinesthetic-facilitated memory performance," *British Journal of Educational Technology*, vol. 44, no. 5, pp. E151–E155, 2013.
- [30] N. D. Finkelstein, W. K. Adams, C. Keller, P. B. Kohl, K. K. Perkins, N. S. Podolefsky, S. Reid, and R. LeMaster, "When learning about the real world is better done virtually: A study of substituting computer simulations for laboratory equipment," *Physical review special topics-physics education research*, vol. 1, no. 1, p. 010103, 2005.
- [31] W. T. Pouw, T. Van Gog, and F. Paas, "An embedded and embodied cognition review of instructional manipulatives," *Educational Psychology Review*, vol. 26, no. 1, pp. 51–72, 2014.
- [32] O. Hauk, I. Johnsrude, and F. Pulvermüller, "Somatotopic representation of action words in human motor and premotor cortex," *Neuron*, vol. 41, no. 2, pp. 301–307, 2004.
- [33] J. González, A. Barros-Loscertales, F. Pulvermüller, V. Meseguer, A. Sanjuán, V. Belloch, and C. Ávila, "Reading cinnamon activates olfactory brain regions," *NeuroImage*, vol. 32, no. 2, pp. 906–912, 2006.
- [34] D. Edge, K.-Y. Cheng, and M. Whitney, "Spatial ease: learning language through body motion," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2013, pp. 469–472.
- [35] M. Macedonia, C. Repetto, A. Ischebeck, and K. Mueller, "Depth of encoding through observed gestures in foreign language word learning," *Frontiers in psychology*, vol. 10, 2019.
- [36] C. Repetto, E. Pedrolì, and M. Macedonia, "Enrichment effects of gestures and pictures on abstract words in a second language," *Frontiers in psychology*, vol. 8, p. 2136, 2017.
- [37] Y. Xiao and M. Watson, "Guidance on conducting a systematic literature review," *Journal of Planning Education and Research*, vol. 39, no. 1, pp. 93–112, 2019.
- [38] P. Brereton, B. A. Kitchenham, D. Budgen, M. Turner, and M. Khalil, "Lessons from applying the systematic literature review process within the software engineering domain," *Journal of systems and software*, vol. 80, no. 4, pp. 571–583, 2007.
- [39] B. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," 2007.
- [40] C. L. Gordon, T. M. Shea, D. C. Noelle, and R. Balasubramaniam, "Affordance compatibility effect for word learning in virtual reality," *Cognitive science*, vol. 43, no. 6, p. e12742, 2019.
- [41] J. H. Seo, B. M. Smith, M. Cook, E. Malone, M. Pine, S. Leal, Z. Bai, and J. Suh, "Anatomy builder vr: Applying a constructive learning method in the virtual reality canine skeletal system," in *International Conference on Applied Human Factors and Ergonomics*. Springer, 2017, pp. 245–252.
- [42] P. Saalfeld, A. Albrecht, W. D'Hanis, H.-J. Rothkötter, and B. Preim, "Learning hand anatomy with sense of embodiment," in *VCBM*, 2020, pp. 43–47.
- [43] Y. Bian, Q. Wang, C. Zhou, G. Qi, C. Yang, X. Meng, and C. Shen, "Comparison of 3d display technologies for embodied interaction in virtual hands-on experiential learning," International Society of the Learning Sciences, Inc.[ISLS], 2018.
- [44] K. Tamaddon and D. Stiefs, "Embodied experiment of levitation in microgravity in a simulated virtual reality environment for science learning," in *2017 IEEE virtual reality workshop on K-12 embodied learning through virtual & augmented reality (KELVAR)*. IEEE, 2017, pp. 1–5.
- [45] J. S.-K. Chang, G. Yeboah, A. Doucette, P. Clifton, M. Nitsche, T. Welsh, and A. Mazalek, "Evaluating the effect of tangible virtual reality on spatial perspective taking ability," in *Proceedings of the 5th symposium on spatial user interaction*, 2017, pp. 68–77.
- [46] J. Ratcliffe and L. Tokarchuk, "Presence, embodied interaction and motivation: Distinct learning phenomena in an immersive virtual environment," in *Proceedings of the 28th ACM International Conference on Multimedia*, 2020, pp. 3661–3668.
- [47] D. Parmar, J. Isaac, S. V. Babu, N. D'Souza, A. E. Leonard, S. Jörg, K. Gundersen, and S. B. Daily, "Programming moves: Design and evaluation of applying embodied interaction in virtual environments to enhance computational thinking in middle school students," in *2016 IEEE Virtual Reality (VR)*. IEEE, 2016, pp. 131–140.
- [48] A. Cheng, L. Yang, and E. Andersen, "Teaching language and culture with a virtual reality game," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, 2017, pp. 541–549.
- [49] T. Jia and Y. Liu, "Words in kitchen: An instance of leveraging virtual reality technology to learn vocabulary," in *2019 IEEE International*

Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct). IEEE, 2019, pp. 150–155.

- [50] J. Hartfill, J. Gabel, D. Neves-Coelho, D. Vogel, F. Räthel, S. Tiede, O. Ariza, and F. Steinicke, “Word saber: an effective and fun vr vocabulary learning game,” in *Proceedings of the Conference on Mensch und Computer*, 2020, pp. 145–154.
- [51] O. Fuhrman, A. Eckerling, N. Friedmann, R. Tarrasch, and G. Raz, “The moving learner: Object manipulation in virtual reality improves vocabulary learning,” *Journal of Computer Assisted Learning*, vol. 37, no. 3, pp. 672–683, 2021.
- [52] J. H. Madden, A. S. Won, J. P. Schuldt, B. Kim, S. Pandita, Y. Sun, T. Stone, and N. Holmes, “Virtual reality as a teaching tool for moon phases and beyond,” *arXiv preprint arXiv:1807.11179*, 2018.
- [53] C. Vázquez, L. Xia, T. Aikawa, and P. Maes, “Words in motion: Kinesthetic language learning in virtual reality,” in *2018 IEEE 18th International Conference on advanced learning technologies (ICALT)*. IEEE, 2018, pp. 272–276.
- [54] J. R. Lewis, “The system usability scale: past, present, and future,” *International Journal of Human–Computer Interaction*, vol. 34, no. 7, pp. 577–590, 2018.
- [55] G. Petri, C. G. von Wangenheim, and A. F. Borgatto, “Meega+: an evolution of a model for the evaluation of educational games,” *INCoD/GQS*, vol. 3, pp. 1–40, 2016.
- [56] R. S. Kennedy, N. E. Lane, K. S. Berbaum, and M. G. Lilienthal, “Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness,” *The international journal of aviation psychology*, vol. 3, no. 3, pp. 203–220, 1993.
- [57] T. Schubert, F. Friedmann, and H. Regenbrecht, “The experience of presence: Factor analytic insights,” *Presence: Teleoperators & Virtual Environments*, vol. 10, no. 3, pp. 266–281, 2001.
- [58] S. G. Hart, “Nasa task load index (tlx). volume 1.0; paper and pencil package,” 1986.
- [59] K. L. Nowak and F. Biocca, “The effect of the agency and anthropomorphism on users’ sense of telepresence, copresence, and social presence in virtual environments,” *Presence: Teleoperators & Virtual Environments*, vol. 12, no. 5, pp. 481–494, 2003.
- [60] M. Gonzalez-Franco and T. C. Peck, “Avatar embodiment. towards a standardized questionnaire,” *Frontiers in Robotics and AI*, vol. 5, p. 74, 2018.
- [61] J. H. Corpus, M. S. McClintic-Gilbert, and A. O. Hayenga, “Within-year changes in children’s intrinsic and extrinsic motivational orientations: Contextual predictors and academic outcomes,” *Contemporary Educational Psychology*, vol. 34, no. 2, pp. 154–166, 2009.
- [62] M. Jurišević, S. A. Glažar, C. R. Pučko, and I. Devetak, “Intrinsic motivation of pre-service primary school teachers for learning chemistry in relation to their academic achievement,” *International Journal of Science Education*, vol. 30, no. 1, pp. 87–107, 2008.
- [63] W.-J. Lee, C.-W. Huang, C.-J. Wu, S.-T. Huang, and G.-D. Chen, “The effects of using embodied interactions to improve learning performance,” in *2012 IEEE 12th International Conference on Advanced Learning Technologies*. IEEE, 2012, pp. 557–559.
- [64] H. Sun and Y. Gao, “Impact of an active educational video game on children’s motivation, science knowledge, and physical activity,” *Journal of Sport and Health Science*, vol. 5, no. 2, pp. 239–245, 2016.
- [65] J. H. Ho, S. Z. Zhou, D. Wei, and A. Low, “Investigating the effects of educational game with wii remote on outcomes of learning,” in *Transactions on Edutainment III*. Springer, 2009, pp. 240–252.
- [66] C.-Y. Lin, Y.-H. Jen, L.-C. Wang, H.-H. Lin, and L.-W. Chang, “Assessment of the application of wii remote for the design of interactive teaching materials,” in *International Conference on Information and Management Engineering*. Springer, 2011, pp. 483–490.
- [67] M. Shakroum, K. W. Wong, and C. C. Fung, “The influence of gesture-based learning system (gbls) on learning outcomes,” *Computers & Education*, vol. 117, pp. 75–101, 2018.
- [68] J. J. Asher, “The learning strategy of the total physical response: A review,” *The modern language journal*, vol. 50, no. 2, pp. 79–84, 1966.
- [69] M. Tellier, “The effect of gestures on second language memorisation by young children,” *Gesture*, vol. 8, no. 2, pp. 219–235, 2008.
- [70] A. M. Glenberg and M. P. Kaschak, “Grounding language in action,” *Psychonomic bulletin & review*, vol. 9, no. 3, pp. 558–565, 2002.
- [71] L. Figueiredo, E. Rodrigues, J. Teixeira, and V. Teichrieb, “A comparative evaluation of direct hand and wand interactions on consumer devices,” *Computers & Graphics*, vol. 77, pp. 108–121, 2018.

APPENDIX A: SUMMARY OF 14 INCLUDED PAPERS

Ref.	Paper Title	Learning Topic
[40]	Affordance Compatibility Effect for Word Learning in Virtual Reality	Word recognition
[41]	Anatomy builder VR: Applying a constructive learning method in the virtual reality canine skeletal system	Canine anatomy
[42]	Learning Hand Anatomy with Sense of Embodiment	Hand anatomy
[43]	Comparison of 3d display technologies for embodied interaction in virtual hands-on experiential learning	Maths
[44]	Embodied experiment of levitation in microgravity in a simulated virtual reality environment for science learning	Microgravity, physics
[45]	Evaluating the Effect of Tangible Virtual Reality on Spatial Perspective Taking Ability	3D Perspective taking
[46]	Presence, embodied interaction and motivation: distinct learning phenomena in an immersive virtual environment	10 noun, 10 verbs
[47]	Programming moves: Design and evaluation of applying embodied interaction in virtual environments to enhance computational thinking in middle school students	Computer science
[48]	Teaching language and culture with a virtual reality game	Cultural actions
[49]	Words in Kitchen: An Instance of Leveraging Virtual Reality Technology to Learn Vocabulary	12 nouns (kitchen)
[50]	Word saber: an effective and fun VR vocabulary learning game	20 nouns (food)
[51]	The moving learner: Object manipulation in virtual reality improves vocabulary learning.	14 nouns (daily objects)
[52]	Virtual reality as a teaching tool for moon phases and beyond	Moon phases
[53]	Words in Motion: Kinesthetic Language Learning in Virtual Reality	20 transitive verbs