- 1 Virtual Reality in Biology: could we become virtual naturalists?
- 2 Authors: Juliano Morimoto^{1*}, Fleur Ponton²

- 4 Author's affiliations:
- 5 1 School of Biological Sciences, University of Aberdeen, Zoology Building, Tillydrone Ave,
- 6 Aberdeen AB24 2TZ, United Kingdom
- 7 2 Department of Biological Sciences, Macquarie University, NSW 2109, North Ryde,
- 8 Australia

9

- *To whom correspondence should be addressed:
- 11 Juliano Morimoto
- 12 School of Biological Sciences, University of Aberdeen, Zoology Building, Tillydrone Ave,
- Aberdeen AB24 2TZ, United Kingdom, e-mail: juliano.morimoto@abdn.ac.uk

Abstract

Virtual Reality (VR) and derived technologies which mixes the real world with some aspect of virtuality (e.g., Augmented Reality, Mixed Reality) has emerged as a powerful allied in education. Recent technological advances made Virtual and Mixed Reality (VMR) accessible at our fingertips. However, only recently VMR has been explored for the teaching of Biology. In this review, we describe how VMR applications are useful in Biology education, discuss the caveats related to VMR use that can interfere with learning, and look into the future of VMR applications in the field. We then propose a conceptual model for the future of VMR in Biology, which envisage to combine VMR with Machine Learning and Artificial Intelligence to provide unprecedented ways to visualise how species evolve in self-sustained immersive virtual worlds, thereby transforming VMR from an educational tool to the centre of biological interest. This conceptual paper aims to stimulate debates on how new technologies can revolutionise teaching and learning across scenarios, which can be useful for improving learning outcomes of biological concepts in face-to-face, blended, and distance learning programmes.

- **Keywords**: Darwinian theory; new technology; evolutionary biology; immersive reality;
- 32 evolving algorithm

Introduction

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

64

65

66

With increasing computational power, technologies that were costly or impossible to implement in the past have now become accessible in laptops and mobile phones (Kish 2004; Waldrop 2016). These technologies are now revolutionising the ways we interact with the world, how we learn, and how we teach (Veletsianos 2010). Virtual and Mixed Reality (VMR, see Box 1 for terminology) is one of these technologies which has gained increasing attention in the academic and teaching communities (Mazuryk and Gervautz 1996). In fact, over the last decade, there has been an exponential increase in the publication of papers in topics involving Virtual and Mixed Reality in Education (Fig 1a). VMR can be defined as an alternate world filled with computer-generated entities that interact with human sensory and motor systems to cause a sense of 'presence' (psychological state) in the subject through the use of an 'immersive' technology (i.e., technology that simulates an environment that is not necessarily real) (Yoh 2001). Presence can be defined as 'a state of dissociation from reality in which people feel the subjective experience of existing in the digital environment (Slater 2003).' Although presence and immersion have been used interchangeably (Barbot and Kaufman 2020), experiences that increase presence do not necessarily increase immersive feelings, and vice versa [see meta-analysis by (Cummings and Bailenson 2016)], suggesting that, although these terms refer to the feeling of 'being there', they are not necessarily equivalent. Nevertheless, both are important in VMR applications. According to the Oxford dictionary, Virtual Reality is defined as 'images and sounds created by a computer that seem almost real to the user, who can interact with them by using sensors', which highlights that presence and immersion are key aspects of virtual reality [see e.g., (Lombart et al. 2020)]. According to Milgram et al. (1995)'s taxonomy, immersive experiences are achieved through a complex continuum in reproduction fidelity of both the real and virtual environments, whereby the limitations of approaches as well as hardware (e.g., devices and displays) can influence the degree of immersive experience and presence available to the user (Milgram et al. 1995). Note, however, that the feeling of presence and immersion may not necessarily be the ultimate goal of VMR technologies [(Milgram et al. 1995); but see (Sheridan 1992; Robinett 1992)], although the increasingly more realistic displays can eventually result in one feeling complete immersion and presence (Naimark 1991).

63

VMR has been around for decades, and it thought to have its origin when, in the 1960s, Morton Heilig created one of the first immersive multi-sensory simulator that included stimuli such as sound, scent, wind and vibration (called 'Sensorama') (Heilig 1962; Mazuryk and Gervautz

1996). Ever since, VR technology has advanced significantly to the point that today there exists many platforms for creating as well as experiencing VMR applications [e.g., (Ledermann and Schmalstieg 2005; Jerald et al. 2014; Wexelblat 2014)]. In the last decade, the use of VMR in teaching and learning has increased dramatically, and spans across a variety of subjects (Hoffman and Vu 1997; Mikropoulos and Natsis 2011; Davies, Crohn, and Treadgold 2019; Makransky, Thisgaard, and Gadegaard 2016), including biology (Makransky et al., 2016; Shim et al., 2003; Shim, Kim, & Park, 2000). A meta-analysis of sixteen studies has shown that VMR surgical simulators decrease the time to complete surgical procedures, suggesting a more efficient surgical skill acquisition (Haque and Srinivasan 2006). Furthermore, VMR improves the learning of tasks that require spatial and visual memory, observation, as well as control of emotional responses in stressful conditions (Jensen and Konradsen 2018). Importantly, autistic children have been described as having positive engagement with VMR applications in educational settings (Strickland et al. 1996; Kandalaft et al. 2013), suggesting that VMR can be used in a wide range of contexts and function as an inclusive tool for the education of students with special needs. Therefore, VMR has the potential to become an important educational tool in our century (Hoffman and Vu 1997).

83 84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

In this conceptual paper, we first provide a mini-review of VMR applications that are changing teaching and learning in biology. Next, we discuss the potential caveats associated with misuses of VMR applications and discuss how VMR use can hinder (rather than help) teaching and learning. We believe this balanced view is important to guide future developments in the field. Lastly, we look into the future of VMR technology and discuss the directions in which VMR developments can take us, revolutionising teaching of important principles of nature. We propose a conceptual model of how VMR can act as an independent, self-sustained virtual experimental world (which we call 'BioVR'). For brevity, we use a broad definition of VMR which includes all virtual types of applications, from Augmented Reality (AR), Mixed Reality (MR) through to Virtual Reality sensu stricto (see Box 1). This paper's main objectives are to provide a balanced overview of the costs and benefits of using VMR into the classroom for teaching evolutionary biology concepts and stimulate discussion about the future of educational technology by proposing a conceptual innovative application of VMR (what we called 'BioVR') in the classroom for an effective teaching of evolutionary biology that can encapsulate three domains of learning (i.e., cognitive, affective, and psychomotor). It is equally important to mention that this paper does not aim to provide a step-by-step guideline for the implementation of VMRs in the classroom, as these guidelines are available elsewhere (see our

discussions below). Having defined the scope of this conceptual paper, we hope that this paper will help guide future developments in VMR applied to biology in a constructive manner, stimulating collaborations across fields (e.g., Computer Science and Gaming) to develop new teaching technologies to facilitate and enhance students' learning experiences.

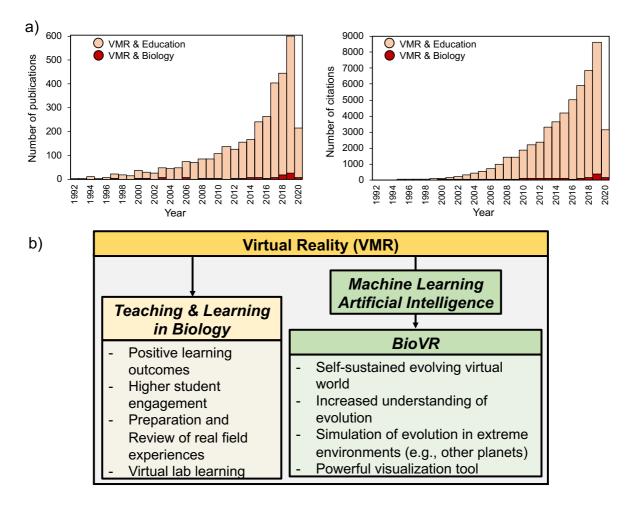


Fig 1 – VMR increasing importance in academic and educational context. (a) Web of Science Topic query of publications (left) and citations (right) that involves VMR and education (orange) and VMR and biology (red). WoS searches were conducted on 12-May-2020 with search term queries '(virtual AND augmented) reality AND education' or '(virtual AND augmented) reality AND biology'. For each search, reviews and proceedings of conferences were excluded. In total, there were 6,443 and 133 papers that fitted the selection criteria, respectively. (b) Schematic overview of the potential for VMR to impact Biology. On one hand, VMR has increasingly been used for teaching of a variety of topics within Biology. As technology advances, it may be possible to combine other cutting-edge technologies such as Machine Learning and Artificial Intelligent to create a self-sustained evolving virtual world (BioVR) that allows us to gain insights into biological processes.

VMR uses in Biology education

While VMR in education has gained exponential attention of the academic community, VMR in biology has advanced at a slower pace, comprising ~ 5% of academic publications in the field (Fig 1a). Nonetheless, VMR has gained important applications in both secondary and tertiary education biology courses (Makransky, Thisgaard, and Gadegaard 2016; Poland, Baggott la Velle, and Nichol 2003; de Jong, Linn, and Zacharia 2013; Thisgaard and Makransky 2017). A number of VMR applications attempt to reproduce the laboratory environment to students with otherwise no access to laboratory facilities, with demonstrated benefits over traditional lectures [see e.g., Labster (Bonde et al. 2014)]. Other VMR applications were designed to give the students an immersive experience of more specific biological processes such as cell structure (McClean, Slator, and White 1999), spatial orientation (Moritz and Meyer 2004), and vision formation in animals (Gochman et al. 2019). Students report higher engagement and learning outcomes with immersive experiences offered by VMR applications, which is encouraging for the use of VMR in biology education (Makransky, Terkildsen, & Mayer, 2019; Makransky et al., 2016; Mikropoulos, Katsikis, Nikolou, & Tsakalis, 2003; Mikropoulos & Natsis, 2011; Shim et al., 2003) (Fig 1b). For example, Moritz and Meyer (2004) designed an immersive interactive VMR platform for visualisation and teaching of conformation and geometry of protein crystallographic structures, whereby the test group was able to identify characteristics and regions in the samples that were obfuscated in non-immersive programs (Moritz and Meyer 2004). Thus, innovative curricula that harness the power of new technologies can provide significant benefits to the teaching and learning of biology (Eastwood and Sadler 2013; Sadler et al. 2015).

140141

142

143

144

145

146

147

148

149

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

VMR applications could help learning and teaching of ecology by simulating field expeditions in which students have to identify plants and/or animals in virtual reality, as in non-immersive virtual field trips developed previously [e.g., (Spicer and Stratford 2001; Dunleavy, Dede, and Mitchell 2009)]. Students have in fact reported that non-immersive virtual field trips provide a useful complement to the real field trip and could be a powerful tool to prepare and revise real field trips (Spicer and Stratford 2001). This could also complement units of taxonomy of plants and animals as well as provide virtual field experience to the student prior to the real task, thereby amalgamating students' learning experience. In VMR, immersive scenarios could include representative environments from different ecosystems (e.g., Amazon rainforest,

tundra, desert) in which the aim is to identify the greater number of plant species as well as the morphological traits that are shared amongst species.

It is important to mention that virtual systems have been developed to explore all aspects of biology education. For instance, pervious digital material has been designed for teaching and learning of astrobiology [for instance in the *Habitable Worlds* platform (Horodyskyj et al. 2018)], although not yet in the fully immersive platform of VMR. *Habitable Worlds* allows students to experience a inquiry-driven learning environment designed to enhance students' learning outcomes on science through observation and modelling of virtual systems (Horodyskyj et al. 2018). The results are promising as more than 70% of students had grades average or higher, and student engagement significantly increased compared to benchmark. As such, *Habitable Worlds* provides some guidelines for the design of digital platforms that could be transferable to VMR systems, including automated feedback tailored to the students' needs and student-educator interactions (both in real-time and in forums) (Horodyskyj et al. 2018). It will be interesting for future developments of *Habitable Worlds* to expand the educational content from astrobiology to other subjects within biology, as well as to include VMR experience and compare the performance of students with traditional *versus* immersive platforms.

The potential misuses of VMR

As for any new technology, we are still discovering the limitations of VMR applications as educational tools. VMR applications are attractive because they contain a wide variety of sensory stimuli that give the participant a sense of presence. However, too many stimuli – such as colours, shapes, characters, movement – can distract the participant and have detrimental effects on learning, a phenomenon that has been acknowledged in the literature and commonly referred to as *cognitive overload* (Whitelock et al. 2000). A recent study has shown that university students learned less and experienced higher cognitive overload when they experienced a science lab in a fully-mounted VMR headset as oppose to the VMR scenes played on 2D displays, in spite of higher feeling of presence in the VMR scene as opposed to the 2D screen display (Makransky, Terkildsen, and Mayer 2019). This suggests that, in some cases, the very same attributes that make VMR attractive can make VMR applications ineffective. Nonetheless, it is possible to mitigate cognitive overload by controlling the information flow that students are exposed to as part of the design of the environment, the tasks needed to be completed, as well as the degree of interactivity with the environment at any given

point in time [e.g., (Dorneich et al. 2003; Wang et al. 2019; Andersen et al. 2016; Bharathan et al. 2013)]. Other potential negative effects of VMR applications are motion sickness and dizziness caused by the immersive experience, which can preclude appropriate understanding of the learning material and hamper students' achievement of learning outcomes (McCauley and Sharkey 1992; Regan and Price 1994; Ohyama et al. 2007; Davis, Nesbitt, and Nalivaiko 2015; Palmisano, Mursic, and Kim 2017). For instance, the perception of angular and linear accelerations cause intense motion sickness both in short and long exposure to VMR applications (McCauley and Sharkey 1992; LaViola Jr 2000; Saredakis et al. 2020; Sharples et al. 2008). Yet, recent studies have demonstrated that motion-sickness can be mitigated or eliminated through appropriate use of VMR design (e.g., immersiveness of the application) and suitable hardware (e.g., display type) [see e.g., (Mittelstaedt, Wacker, and Stelling 2018; Mittelstaedt, Wacker, and Stelling 2019; Weech, Kenny, and Barnett-Cowan 2019)]. Interestingly, a recent meta-analysis suggested that the feelings of presence and motionsickness likely trade-off – with higher presence resulting in lower motion-sickness (Weech, Kenny, and Barnett-Cowan 2019), supporting the idea that motion-sickness can be mitigated (or eliminated) with thoughtful design and user-experience.

Recent literature provides comprehensive lists of fundamental characteristics of 3D virtual environments and general features that can be adjusted to increase students' engagement and learning in virtual systems [see e.g., (Dalgarno and Lee 2010; Lindgren et al. 2016; O'Connor and Domingo 2017)]. In fact, a series of learning affordances have been identified, which we reproduce in Table 1, that have positive influence on students' engagement with virtual applications and can ultimately enhance learning (Dalgarno and Lee 2010); these learning affordances are revisited in our BioVR model (see below). These affordances highlight the benefits of using VMR applications in the classroom. Here, our point is to reiterate the importance of careful design and testing of new VMR applications prior to implementation in the classroom in order to mitigate potential cognitive overload and/or motion sickness, which could significantly hamper VMR's educational potential and the achievement of learning affordances (Dalgarno and Lee 2010). Future studies will provide more detailed evidence-based guidelines to build effective VMR applications that maximise educational potential while minimising negative effects of VMR misuse (Dunleavy, Dede, and Mitchell 2009;

Bio-inspired systems and the rise of artificial evolution

Akçayır and Akçayır 2017).

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

The parallels between natural and artificial evolutionary systems have long been recognised and explored. While few artificial life systems exist, perhaps the most famous example comes from the work of Thomas Ray and the 'Tierra' system (Langton 1986; Langton 1997; Lehman, Clune, and Misevic 2018; Ray 1992). The Tierra system simulates artificial life in selfreplicating, evolving entities ('algorithms') confined within virtual computer spaces, whereby the entities can be considered as uni- or multi-cellular entities that experience errors in replication analogous to mutations in biological reproduction (Ray 1992; Thearling and Ray 1994; Ray 1994). Instead of solar energy and natural resources as in biological systems, artificial entities compete for central processing unit (CPU) and memory space [analogous to energy and spatial resource, respectively [as described in (Ray 1992)]. As a result, artificial Tierra entities become progressively more adapted to exploit one another in order to gain advantage over the use of CPU and memory (Ray 1992; Thearling and Ray 1994; Ray 1994; Ray and Xu 2001). The outcome of this self-sustained virtual evolutionary world is remarkable given that the system evolves differences in entity sizes, ecological specialisation (e.g., parasites) and population dynamics processes (e.g., extinction) (Ray, 1992, 1994; Ray & Xu, 2001; Shao & Ray, 2010). This provides an unprecedented study case to compare and understand how different shapes and forms emerge through evolutionary processes. However, visualisation of evolution in the Tierra system is not straightforward and largely inaccessible to a broader audience due to the highly technical language underlying the system. This poses a significant barrier to biologists with limited computational expertise and it is, to some extent, visually unappealing for students of biological sciences and related disciplines. Consequently, it is difficult (though not impossible) to use artificial model systems such as Tierra as an effective educational tool in the classroom while keeping the attention span and interest of students.

Can VMR and Artificial Intelligence (AI) revolutionise artificial evolutionary systems?

As discussed above, VMR is a powerful and appealing educational technology to teach biology. This is because students and educators respond rationally as well as emotionally to the educational material in the immersive experience, which can accentuate learning (de Jong, Linn, and Zacharia 2013; Thisgaard and Makransky 2017; Dunleavy, Dede, and Mitchell 2009; Harley et al. 2016; Riva et al. 2007). Thus, VMR can be an appropriate way to overcome accessibility problems of artificially evolving systems while increasing visual appealing to specialists and general audience.

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

The technological advances that allowed VMR to become an accessible tool has also allowed for the feasibility of powerful statistical models of Machine Learning and Artificial Intelligence (AI). Machine Learning are algorithms that process and learn with huge amounts of data in order to perform a task without necessarily being explicitly programmed to do so (Bishop 2006). AI attempts to simulate human intelligence in machine systems; this includes machine learning but also (bio-inspired) robotics, ethics and philosophy associated with AI development (Russell and Norvig 2002). Importantly, AI advances have recently demonstrated that machines can learn from data beyond human capabilities (Chouard 2016; Gibney 2016). Furthermore, a new area on the interface between VMR and AI aims to integrate AI to entities in VMR (Luck and Aylett 2000; Laukkanen et al. 2004; Augusto et al. 2013). As a result, a key question emerges: can we combine Machine Learning and AI with VMR to create a selfsustained evolving virtual world (a 'BioVR')? If so, why should we combine VMR with AI? The answer to the first question is, in our opinion, a sounding 'yes'. We strongly believe that future technological advances have the potential to create an immersive virtual world that reproduces the forces of evolution, which can allow us to visualise and measure how species have evolved, how ecosystems are formed, how species adapt to their environment, how we can anticipate effects of adverse climatic conditions across ecosystems in our changing world. In a sense, we could become 'virtual naturalists' that explore evolution in a simulated (virtual) world in the same sense that naturalists explore the natural (real) world. The learning benefits are unprecedented given that students can experience inaccessible and inhospitable environments, observe evolution, adaptation, trophic interaction, parasitism and many more biological processes without stepping outside the classroom (Learning affordance #2; Table 1). Moreover, the freedom given to the students within these BioVRs forms the perfect ground for inquiry-based learning and engagement, where the students will observe and explore the environment, measuring and experiencing the virtual environment to inquiry about the underlying virtual biological phenomena (Horodyskyj et al. 2018) (Learning affordance # 3, Table 1). Ultimately, the BioVR could be used to supplement and stimulate collaborative learning tasks, where students may compare and contrast the evolution of individual BioVR evolutions and identify similarities (e.g., parallel evolution) and divergences (e.g., specialisation traits) between virtual entities (Learning affordance #5; Table 1). The BioVR could then eliminate the need for complex computational expertise (at least from the users' point of view) and provide a fully immersive, artificial world upon which entities evolve following basic principles of biological evolution in our and other planets, while students can explore the environment and learn from their own virtual experience.

- Practical implementation of BioVR by experts could in theory be achieved through the following steps:
 - 1. Simulate an artificial 'planet' whereby entities will interact, compete, and evolve. In this artificial planet, the 'biotic' rules are established, such as the basic environmental conditions (e.g., temperature), habitat (e.g., marine vs terrestrial landscapes), resource distribution and so on [similar in concept to the 'soup' in Tierra and the concept of virtual environments in (Luck and Aylett 2000)].
 - 2. Design the ancestor entity, defining the rules of reproduction, mutation, and ecological interactions that govern the outcomes of interactions between entities as well as with the environment in the planet. This will likely require that the entities follow theoretical models that relate trait expression and fitness, as well as fitness changes over lifetime and through generations [see for example models in (Edelaar and Bolnick 2019)]. We envisage that the ancestor entity is the 'building block' for artificial life to evolve in BioVR and without it, the system does not have the evolving entity. The ancestor entity is equivalent to the ancestor species which gave origin to life on Earth, and is a common feature of artificial life systems [e.g., (Ray 1994)]. In other words, the ancestral entity is the first 'living' inhabitant of the virtual planet.
 - 3. Gather a large empirical dataset of environment-traits-species interactions as a basic starting-point for determining how different species evolve in different ecosystems (e.g., evolutionary convergences, divergences, character displacement) – this could be called 'rules of evolution'. This will allow the system to 'know' which adaptations are more likely to yield fitness advantages in a given environment. For instance, heat tolerance (or traits related to coping with high temperatures) is a likely to increase fitness of populations living in virtual habitats that resemble arid regions. The actual evolution of traits following the 'rule of evolution' will depend on the variability present in the ancestral entity populations and we envisage that, in a BioVR simulation, this parameter can adjustable according to the purpose or the needs of the virtual world. One way in which evolutionary rules could be extracted from this dataset is using, for example, supervised learning and/or clustering algorithms (see Box 2) to extract general rules as to how species evolve (morphologically and behaviourally) across different environments, commonality between functional traits across species in the same environments, as well as the number, distribution, and behaviour of different species within the same environment. For instance, species living in warm habitats are likely

- to share similar adaptations to high temperatures and thus, one could expect that virtual entities inhabiting warm virtual habitats should resemble follow similar patterns. Once these rules of evolution are estimated (or guessed) across all habitats of the virtual world, ancestral entities can evolve and differentiate accordingly.
- 4. Ideally, BioVRs are self-sustained, and thus it would be interesting to have the changes and adaptations in one time point to feedback into the system for the next time point. For example, imagine that a species evolves a remarkable adaptation to convert virtual resource A into B. This transformation should feedback into the system so as to allow new evolutionary rules, perhaps favouring other species to adapt and utilise virtual resource B (which is being produced) instead of virtual resource A (Fig 2).
- 5. Given this self-sustained cycle of interaction between entities and the environments, and the iterative system that modulates virtual evolutionary rules, BioVR can become an artificial ecosystem, fully accessible for exploration through VMR in inquiry-based learning quests. This allows students and researchers to experience and study evolution in this immersive environment, comparing the outcomes of evolutionary forces within different environments within a BioVR and across BioVRs with different setups. Furthermore, since data visualisation is key for understanding biological processes [e.g., (Karr and Brady 2000)] and is an essential component of affective learning, the use of VR to create BioVR worlds will allow VR to transcend the status of an educational tool that helps learning and teaching in Biology to become the main technology for experiencing and learning about virtual biological phenomena.

We provided these steps in order to quick-start ideas around the practical challenges necessary to realise the conceptual proposition made in this paper. We are not presumptuous of our knowledge boundaries and understand that experts may have better implementation methods and other tools that we are unaware. Having said that, these steps are aimed to foster open discussions that can generate international collaborations which may make possible for BioVR to reach the classroom as an effective educational tool of the future. It is important to mention that while the idea of BioVRs may seem now allusive, attempts to merge the fields of VMR, artificial life, and AI have been around for decades (Luck and Aylett 2000), with more recent efforts emerging from the astonishing 'boost' in computer power of our generation (Petrović 2018; Zhang et al. 2018). We are also aware that virtual environments, AI and semi-autonomous VMR agents have been developed for other purposes such as direct or assist users into tasks [see e.g., (Luck and Aylett 2000)]. To our knowledge, concepts similar to the one of

BioVR as presented in this paper have never been conceptualised let alone tested, which underscores the importance of opening this avenue of communication around the concept of BioVR for future developments that aid education of Biology. Future research and discussion should therefore aim at assessing the feasibility of the concepts proposed here.

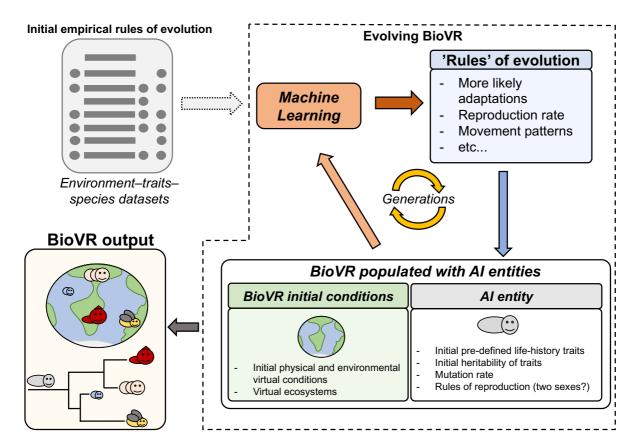


Fig 2 –Conceptual overview of the steps to build a BioVR. A supervised machine learning algorithm is implemented to empirical environment-trait-species datasets in order to extract the patterns (or 'rules') of evolution across environments. Meanwhile, the initial settings for the BioVR world and the ancestral AI entity are also set. The settings include physical and environmental conditions, as well as patterns of lifespan, movement, and reproduction of the AI entity. Next, the 'rules of evolution' are incorporated into the BioVR and AI entity with original settings, and the BioVR is allowed to evolve. Note that the evolution patterns in the BioVR are then fed-back to the machine learning model, which is updated. This way, the only input from empirical data is at the initial states, and BioVR are allowed to evolve independently afterwards. As a result, we can measure and visualise species evolution as it happens, in an immersive experience of the BioVR.

373 Conclusion

- 374 The use of VMR has provided promising results for consolidating learning across secondary
- and tertiary biology education. With increasing technology, the combination of VMR with
- 376 Machine Learning and AI has the potential to create a self-sustained evolving virtual world
- 377 (BioVR) that allow us to uniquely explore how life as we know evolves and responds to
- 378 extreme climatic conditions. Thus, the use of new technologies and innovative applications can
- provide better learning outcomes and student experiences in face-to-face, blended, and distance
- learning contexts by amalgamating biological principles in immersive classrooms.

382 Funding details

381

384

395

396 397

398

399 400

401 402

403

404 405

406

407

383 This study was not funded by any funding agency.

385 Conflict of interest

386 The authors have no conflict of interests to declare.

387 References

- Akçayır, Murat, and Gökçe Akçayır. 2017. "Advantages and Challenges Associated with Augmented Reality for Education: A Systematic Review of the Literature." *Educational Research Review* 20: 1–11. doi:https://doi.org/10.1016/j.edurev.2016.11.002.
- Andersen, Steven Arild Wuyts, Peter Trier Mikkelsen, Lars Konge, Per Cayé-Thomasen, and
 Mads Sølvsten Sørensen. 2016. "The Effect of Implementing Cognitive Load Theory Based Design Principles in Virtual Reality Simulation Training of Surgical Skills: A
 Randomized Controlled Trial." Advances in Simulation 1 (1). Springer: 20.
 - Augusto, Juan C, Vic Callaghan, Diane Cook, Achilles Kameas, and Ichiro Satoh. 2013. "Intelligent Environments: A Manifesto." *Human-Centric Computing and Information Sciences* 3 (1): 12. doi:10.1186/2192-1962-3-12.
 - Barbot, Baptiste, and James C Kaufman. 2020. "What Makes Immersive Virtual Reality the Ultimate Empathy Machine? Discerning the Underlying Mechanisms of Change." *Computers in Human Behavior*. Elsevier, 106431.
 - Bharathan, Rasiah, Saaliha Vali, Thomas Setchell, Tariq Miskry, Ara Darzi, and Rajesh Aggarwal. 2013. "Psychomotor Skills and Cognitive Load Training on a Virtual Reality Laparoscopic Simulator for Tubal Surgery Is Effective." *European Journal of Obstetrics & Gynecology and Reproductive Biology* 169 (2). Elsevier: 347–352.
 - Bishop, Christopher M. 2006. *Pattern Recognition and Machine Learning*. Edited by M Jordan, J. Kleinberg, and B. Scholkopf. Cambridge, UK: Springer Science Business Media.
- Bonde, Mads T, Guido Makransky, Jakob Wandall, Mette V Larsen, Mikkel Morsing, Hanne
 Jarmer, and Morten O A Sommer. 2014. "Improving Biotech Education through
 Gamified Laboratory Simulations." *Nature Biotechnology* 32 (7): 694–697.
 doi:10.1038/nbt.2955.
- Chouard, Tanguy. 2016. "The Go Files: AI Computer Wraps up 4–1 Victory against Human Champion." *Nature News*.
- Cummings, James J, and Jeremy N Bailenson. 2016. "How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence." *Media Psychology*

- 416 19 (2). Taylor & Francis: 272–309.
- Dalgarno, Barney, and Mark J W Lee. 2010. "What Are the Learning Affordances of 3-D Virtual Environments?" *British Journal of Educational Technology* 41 (1). John Wiley & Sons, Ltd: 10–32. doi:10.1111/j.1467-8535.2009.01038.x.
- Davies, Andrew Graham, Nick J Crohn, and Laura Anne Treadgold. 2019. "Can Virtual Reality Reality Reality Be Used within the Lecture Theatre?" *BMJ Simulation and Technology Enhanced Learning* 5 (4): 234 LP 235. doi:10.1136/bmjstel-2017-000295.
- Davis, Simon, Keith Nesbitt, and Eugene Nalivaiko. 2015. "Comparing the Onset of Cybersickness Using the Oculus Rift and Two Virtual Roller Coasters." In *Proceedings of the 11th Australasian Conference on Interactive Entertainment (IE 2015)*, 27:30.
- de Jong, Ton, Marcia C Linn, and Zacharias C Zacharia. 2013. "Physical and Virtual
 Laboratories in Science and Engineering Education." *Science* 340 (6130): 305 LP 308.
 doi:10.1126/science.1230579.
- Dorneich, Michael C, Stephen D Whitlow, Patricia May Ververs, and William H Rogers.
 2003. "Mitigating Cognitive Bottlenecks via an Augmented Cognition Adaptive
 System." In SMC'03 Conference Proceedings. 2003 IEEE International Conference on
 Systems, Man and Cybernetics. Conference Theme-System Security and Assurance (Cat.
 No. 03CH37483), 1:937–944. IEEE.
- Dunleavy, Matt, Chris Dede, and Rebecca Mitchell. 2009. "Affordances and Limitations of Immersive Participatory Augmented Reality Simulations for Teaching and Learning."

 Journal of Science Education and Technology 18 (1): 7–22. doi:10.1007/s10956-008-9119-1.
- Eastwood, Jennifer L, and Troy D Sadler. 2013. "Teachers' Implementation of a Game-Based Biotechnology Curriculum." *Computers & Education* 66: 11–24. doi:https://doi.org/10.1016/j.compedu.2013.02.003.
- Edelaar, Pim, and Daniel I Bolnick. 2019. "Appreciating the Multiple Processes Increasing
 Individual or Population Fitness." *Trends in Ecology & Evolution* 34 (5). Elsevier: 435–
 443
- Gibney, Elizabeth. 2016. "Google AI Algorithm Masters Ancient Game of Go." *Nature News* 529 (7587): 445.
- Gochman, Samuel R, Marilyn Morano Lord, Naman Goyal, Kristie Chow, Benjamin K
 Cooper, Lauren K Gray, Stephanie X Guo, et al. 2019. "Tarsier Goggles: A Virtual
 Reality Tool for Experiencing the Optics of a Dark-Adapted Primate Visual System."
 Evolution: Education and Outreach 12 (1): 9. doi:10.1186/s12052-019-0101-6.
- Haque, S, and S Srinivasan. 2006. "A Meta-Analysis of the Training Effectiveness of Virtual
 Reality Surgical Simulators." *IEEE Transactions on Information Technology in* Biomedicine 10 (1): 51–58. doi:10.1109/TITB.2005.855529.
- Harley, Jason M, Eric G Poitras, Amanda Jarrell, Melissa C Duffy, and Susanne P Lajoie.
 2016. "Comparing Virtual and Location-Based Augmented Reality Mobile Learning:
 Emotions and Learning Outcomes." *Educational Technology Research and Development* 64 (3): 359–388. doi:10.1007/s11423-015-9420-7.
- 457 Heilig, Morton L. 1962. "Sensorama Simulator." Google Patents.
- Hoffman, H, and D Vu. 1997. "Virtual Reality: Teaching Tool of the Twenty-First Century?"
 Academic Medicine: Journal of the Association of American Medical Colleges 72 (12).
 University of California, San Diego, School of Medicine, La Jolla, USA.
- hhoffman@ucsd.edu: 1076–1081. doi:10.1097/00001888-199712000-00018.
- Horodyskyj, Lev B, Chris Mead, Zack Belinson, Sanlyn Buxner, Steven Semken, and Ariel D
 Anbar. 2018. "Habitable Worlds: Delivering on the Promises of Online Education."
- 464 Astrobiology 18 (1). Mary Ann Liebert, Inc., publishers: 86–99.
- doi:10.1089/ast.2016.1550.

- Jensen, Lasse, and Flemming Konradsen. 2018. "A Review of the Use of Virtual Reality 466 Head-Mounted Displays in Education and Training." Education and Information 467 468 Technologies 23 (4). Springer: 1515–1529.
- 469 Jerald, Jason, Peter Giokaris, Danny Woodall, Arno Hartbolt, Anish Chandak, and Sebastien 470 Kuntz. 2014. "Developing Virtual Reality Applications with Unity." In 2014 IEEE 471 *Virtual Reality (VR)*, 1–3. IEEE.
- 472 Kandalaft, Michelle R, Nyaz Didehbani, Daniel C Krawczyk, Tandra T Allen, and Sandra B 473 Chapman. 2013. "Virtual Reality Social Cognition Training for Young Adults with High-Functioning Autism." Journal of Autism and Developmental Disorders 43 (1). 474 475 Springer: 34–44.
- Karr, Timothy L, and Rachael Brady. 2000. "Virtual Biology in the CAVE." Trends in 476 Genetics 16 (5). Elsevier: 231–232. doi:10.1016/S0168-9525(00)01996-X. 477
- Kish, L B. 2004. "Moore's Law and the Energy Requirement of Computing versus 478 479 Performance." IEE Proceedings - Circuits, Devices and Systems 151 (2): 190–194. 480 doi:10.1049/ip-cds:20040434.
- 481 Langton, Christopher G. 1986. "Studying Artificial Life with Cellular Automata." *Physica D*: 482 Nonlinear Phenomena 22 (1): 120–149. doi:https://doi.org/10.1016/0167-483 2789(86)90237-X.
- 484 Langton, Christopher G. 1997. Artificial Life: An Overview. Mit Press.
- 485 Laukkanen, Seppo, Ilkka Karanta, Ville Kotovirta, Jussi Markkanen, and Jukka Rönkkö. 486 2004. "Adding Intelligence to Virtual Reality." In Proceedings of the 16th European 487 Conference on Artificial Intelligence, 1136–1141. Citeseer.
- 488 LaViola Jr, Joseph J. 2000. "A Discussion of Cybersickness in Virtual Environments." ACM 489 Sigchi Bulletin 32 (1). ACM New York, NY, USA: 47–56.
 - Ledermann, Florian, and Dieter Schmalstieg. 2005. APRIL: A High-Level Framework for Creating Augmented Reality Presentations. IEEE.
 - Lehman, Joel, Jeff Clune, and Dusan Misevic. 2018. "The Surprising Creativity of Digital Evolution." In *Artificial Life Conference Proceedings*, 55–56. MIT Press.
- Lindgren, Robb, Michael Tscholl, Shuai Wang, and Emily Johnson. 2016. "Enhancing 494 495 Learning and Engagement through Embodied Interaction within a Mixed Reality 496 Simulation." Computers & Education 95: 174–187. 497 doi:https://doi.org/10.1016/j.compedu.2016.01.001.
- Lombart, Cindy, Elena Millan, Jean-Marie Normand, Adrien Verhulst, Blandine Labbé-Pinlon, and Guillaume Moreau. 2020. "Effects of Physical, Non-Immersive Virtual, and 500 Immersive Virtual Store Environments on Consumers' Perceptions and Purchase Behavior." Computers in Human Behavior. Elsevier, 106374.
 - Luck, Michael, and Ruth Aylett. 2000. "Applying Artificial Intelligence to Virtual Reality: Intelligent Virtual Environments." Applied Artificial Intelligence 14 (1). Taylor & Francis: 3–32. doi:10.1080/088395100117142.
- 505 Makransky, Guido, Thomas S Terkildsen, and Richard E Mayer. 2019. "Adding Immersive Virtual Reality to a Science Lab Simulation Causes More Presence but Less Learning." 506 507 Learning and Instruction 60. Elsevier: 225–236.
- Makransky, Guido, Malene Warming Thisgaard, and Helen Gadegaard. 2016. "Virtual 508 509 Simulations as Preparation for Lab Exercises: Assessing Learning of Key Laboratory Skills in Microbiology and Improvement of Essential Non-Cognitive Skills." PLOS 510 ONE 11 (6). Public Library of Science: e0155895. 511
- https://doi.org/10.1371/journal.pone.0155895. 512

492 493

498 499

501

502 503

- Mazuryk, Tomasz, and Michael Gervautz. 1996. "Virtual Reality-History, Applications, 513 Technology and Future." Citeseer. 514
- 515 McCauley, Michael E, and Thomas J Sharkey. 1992. "Cybersickness: Perception of Self-

542 543

544

545

546 547

548 549

550

551

552

553554

555

556

557

- Motion in Virtual Environments." *Presence: Teleoperators and Virtual Environments* 1 (3). MIT Press: 311–318. doi:10.1162/pres.1992.1.3.311.
- McClean, Phillip E, Brian M Slator, and Alan R White. 1999. "The Virtual Cell: An
 Interactive, Virtual Environment for Cell Biology." In *EdMedia+ Innovate Learning*,
 1442–1443. Association for the Advancement of Computing in Education (AACE).
- Mikropoulos, Tassos A, Apostolos Katsikis, Eugenia Nikolou, and Panayiotis Tsakalis. 2003.
 "Virtual Environments in Biology Teaching." *Journal of Biological Education* 37 (4).
 Routledge: 176–181. doi:10.1080/00219266.2003.9655879.
- Mikropoulos, Tassos A, and Antonis Natsis. 2011. "Educational Virtual Environments: A Ten-Year Review of Empirical Research (1999–2009)." *Computers & Education* 56 (3): 769–780. doi:https://doi.org/10.1016/j.compedu.2010.10.020.
- Milgram, Paul, Haruo Takemura, Akira Utsumi, and Fumio Kishino. 1995. "Augmented
 Reality: A Class of Displays on the Reality-Virtuality Continuum." In *Proc.SPIE*. Vol. 2351. doi:10.1117/12.197321.
- Mitchell, R S, J G Michalski, and T M Carbonell. 2013. *An Artificial Intelligence Approach*.
 Berlin, Germany: Springer,.
- 532 Mittelstaedt, Justin Maximilian, Jan Wacker, and Dirk Stelling. 2019. "VR Aftereffect and 533 the Relation of Cybersickness and Cognitive Performance." *Virtual Reality* 23 (2). 534 Springer: 143–154.
- 535 Mittelstaedt, Justin, Jan Wacker, and Dirk Stelling. 2018. "Effects of Display Type and 536 Motion Control on Cybersickness in a Virtual Bike Simulator." *Displays* 51. Elsevier: 537 43–50.
- Moritz, E, and J Meyer. 2004. "Interactive 3D Protein Structure Visualization Using Virtual
 Reality." In *Proceedings. Fourth IEEE Symposium on Bioinformatics and* Bioengineering, 503–507. doi:10.1109/BIBE.2004.1317384.
 - Naimark, Michael. 1991. "Elements of Real-Space Imaging: A Proposed Taxonomy." In *Stereoscopic Displays and Applications II*, 1457:169–179. International Society for Optics and Photonics.
 - O'Connor, Eileen A, and Jelia Domingo. 2017. "A Practical Guide, with Theoretical Underpinnings, for Creating Effective Virtual Reality Learning Environments." *Journal of Educational Technology Systems* 45 (3). SAGE Publications Sage CA: Los Angeles, CA: 343–364.
 - Ohyama, Seizo, Suetaka Nishiike, Hiroshi Watanabe, Katsunori Matsuoka, Hironori Akizuki, Noriaki Takeda, and Tamotsu Harada. 2007. "Autonomic Responses during Motion Sickness Induced by Virtual Reality." *Auris Nasus Larynx* 34 (3): 303–306. doi:https://doi.org/10.1016/j.anl.2007.01.002.
 - Palmisano, Stephen, Rebecca Mursic, and Juno Kim. 2017. "Vection and Cybersickness Generated by Head-and-Display Motion in the Oculus Rift." *Displays* 46. Elsevier: 1–8.
 - Petrović, V M. 2018. "Artificial Intelligence and Virtual Worlds Toward Human-Level AI Agents." *IEEE Access* 6: 39976–39988. doi:10.1109/ACCESS.2018.2855970.
 - Poland, Roger, Linda Baggott la Velle, and Jon Nichol. 2003. "The Virtual Field Station (VFS): Using a Virtual Reality Environment for Ecological Fieldwork in A-Level Biological Studies—Case Study 3." *British Journal of Educational Technology* 34 (2). John Wiley & Sons, Ltd: 215–231. doi:10.1111/1467-8535.00321.
- Ray, Thomas S. 1992. Evolution, Ecology and Optimization of Digital Organisms. Citeseer.
- Ray, Thomas S. 1994. "Evolution, Complexity, Entropy and Artificial Reality." *Physica D: Nonlinear Phenomena* 75 (1): 239–263. doi:https://doi.org/10.1016/0167-2789(94)90286-0.
- Ray, Thomas S, and Chenmei Xu. 2001. "Measures of Evolvability in Tierra." *Artificial Life and Robotics* 5 (4): 211–214. doi:10.1007/BF02481504.

580

581

582

583

584 585

586

587 588

589

590 591

592

593 594

595 596

597

598

603

- Regan, E C, and K R Price. 1994. "The Frequency of Occurrence and Severity of Side Effects of Immersion Virtual Reality." *Aviation, Space, and Environmental Medicine*.
 Aerospace Medical Assn.
- Riva, Giuseppe, Fabrizia Mantovani, Claret Samantha Capideville, Alessandra Preziosa,
 Francesca Morganti, Daniela Villani, Andrea Gaggioli, Cristina Botella, and Mariano
 Alcañiz. 2007. "Affective Interactions Using Virtual Reality: The Link between
 Presence and Emotions." CyberPsychology & Behavior 10 (1). Mary Ann Liebert, Inc. 2
 Madison Avenue Larchmont, NY 10538 USA: 45–56.
- Robinett, Warren. 1992. "Synthetic Experience: A Proposed Taxonomy." *Presence: Teleoperators & Virtual Environments* 1 (2). MIT Press: 229–247.
- Russell, Stuart, and Peter Norvig. 2002. Artificial Intelligence: A Modern Approach. Upper
 Saddle River New Jersey: Pearson Education.
 - Sadler, TROY D, WILLIAM L Romine, DEEPIKA Menon, RICHARD E Ferdig, and LEONARD Annetta. 2015. "Learning Biology Through Innovative Curricula: A Comparison of Game- and Nongame-Based Approaches." *Science Education* 99 (4). John Wiley & Sons, Ltd: 696–720. doi:10.1002/sce.21171.
 - Saredakis, Dimitrios, Ancret Szpak, Brandon Birckhead, Hannah A D Keage, Albert Rizzo, and Tobias Loetscher. 2020. "Factors Associated with Virtual Reality Sickness in Head-Mounted Displays: A Systematic Review and Meta-Analysis." *Frontiers in Human Neuroscience* 14. Frontiers Media SA.
 - Serino, Maeve, Kyla Cordrey, Laura McLaughlin, and Ruth L Milanaik. 2016. "Pokémon Go and Augmented Virtual Reality Games: A Cautionary Commentary for Parents and Pediatricians." *Current Opinion in Pediatrics* 28 (5). Wolters Kluwer: 673–677.
 - Shao, Jie, and Thomas S Ray. 2010. "Maintenance of Species Diversity by Predation in the Tierra System." In *ALife*, 533–540.
 - Sharples, Sarah, Sue Cobb, Amanda Moody, and John R Wilson. 2008. "Virtual Reality Induced Symptoms and Effects (VRISE): Comparison of Head Mounted Display (HMD), Desktop and Projection Display Systems." *Displays* 29 (2). Elsevier: 58–69.
 - Sheridan, Thomas B. 1992. "Musings on Telepresence and Virtual Presence." *Presence: Teleoperators & Virtual Environments* 1 (1). MIT Press: 120–126.
 - Shim, K C, H S Kim, and Y C Park. 2000. *Application of Multi-Media in Biology Education, Summer Program of Korean Society of Biology Education*. Edited by null. Vol. null. Null.
- Shim, Kew-Cheol, Jong-Seok Park, Hyun-Sup Kim, Jae-Hyun Kim, Young-Chul Park, and
 Hai-II Ryu. 2003. "Application of Virtual Reality Technology in Biology Education."
 Journal of Biological Education 37 (2). Routledge: 71–74.
 doi:10.1080/00219266.2003.9655854.
 - Slater, Mel. 2003. "A Note on Presence Terminology." *Presence Connect* 3 (3). Citeseer: 1–5.
- Spicer, John, and J Stratford. 2001. "Student Perceptions of a Virtual Field Trip to Replace a
 Real Field Trip." *Journal of Computer Assisted Learning* 17 (December): 345–354.
 doi:10.1046/j.0266-4909.2001.00191.x.
- Strickland, Dorothy, Lee M Marcus, Gary B Mesibov, and Kerry Hogan. 1996. "Brief
 Report: Two Case Studies Using Virtual Reality as a Learning Tool for Autistic
 Children." *Journal of Autism and Developmental Disorders* 26 (6). Springer: 651–659.
- Thearling, Kurt, and Thomas S Ray. 1994. "Evolving Multi-Cellular Artificial Life." In
 Artificial Life IV, 283–288.
- Thisgaard, Malene, and Guido Makransky. 2017. "Virtual Learning Simulations in High School: Effects on Cognitive and Non-Cognitive Outcomes and Implications on the Development of STEM Academic and Career Choice ." Frontiers in Psychology.

629

630

631 632

633

634 635

636

637 638

639

- https://www.frontiersin.org/article/10.3389/fpsyg.2017.00805.
- Veletsianos, George. 2010. Emerging Technologies in Distance Education. Athabasca
 University Press.
- Waldrop, M Mitchell. 2016. "The Chips Are down for Moore's Law." *Nature News* 530 (7589): 144.
- Wang, Danding, Qian Yang, Ashraf Abdul, and Brian Y Lim. 2019. "Designing Theory Driven User-Centric Explainable AI." In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–15.
- Weech, Séamas, Sophie Kenny, and Michael Barnett-Cowan. 2019. "Presence and
 Cybersickness in Virtual Reality Are Negatively Related: A Review ." Frontiers in
 Psychology . https://www.frontiersin.org/article/10.3389/fpsyg.2019.00158.
- Wexelblat, Alan. 2014. Virtual Reality: Applications and Explorations. Academic Press.
 - Whitelock, Denise, Daniela Romano, Anne Jelfs, and Paul Brna. 2000. "Perfect Presence: What Does This Mean for the Design of Virtual Learning Environments?" *Education and Information Technologies* 5 (4). Springer: 277–289.
 - Yoh, Myeung-Sook. 2001. "The Reality of Virtual Reality." In *Proceedings of the Seventh International Conference on Virtual Systems and Multimedia (VSMM'01)*, 666. VSMM '01. USA: IEEE Computer Society.
 - Zhang, Z, B Cao, J Guo, D Weng, Y Liu, and Y Wang. 2018. "Inverse Virtual Reality: Intelligence-Driven Mutually Mirrored World." In 2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), 735–736. doi:10.1109/VR.2018.8446260.
 - Zhu, Xiaojin, and Andrew B Goldberg. 2009. "Introduction to Semi-Supervised Learning." *Synthesis Lectures on Artificial Intelligence and Machine Learning* 3 (1). Morgan & Claypool Publishers: 1–130.
- Zsila, Ágnes, Gábor Orosz, Beáta Bőthe, István Tóth-Király, Orsolya Király, Mark Griffiths,
 and Zsolt Demetrovics. 2018. "An Empirical Study on the Motivations Underlying
 Augmented Reality Games: The Case of Pokémon Go during and after Pokémon Fever."
 Personality and Individual Differences 133: 56–66.
 doi:https://doi.org/10.1016/j.paid.2017.06.024.

Box 1 – The reality-virtuality continuum

In a highly influential paper, Milgram, Takemura, Utsumi, & Kishino (1995) proposed the reality-virtuality continuum (see Fig 3) to classify Virtual, Augmented, and Mixed reality technologies. On one side of the spectrum is the real world (reality) and, on the other side of the spectrum, the fully virtual world (virtuality) where Virtual Reality (VR) in its strict sense resides. In between the extremes, stands Augmented Reality (AR) – which relies mostly on real world elements but with the addition of virtual entities; the best known (and controversial) example of AR has been Pokemon Go! (Serino et al. 2016; Zsila et al. 2018) – and Augmented Virtuality (AV) with the opposite of AR, that is, mostly virtual world but with the addition of 'real' entities. AR and AV are cases of Mixed Reality (MR), where real and virtual elements are intertwined within the application (see Fig 3). For the purpose of this paper and for simplicity, we refer to AR, AV, and VR all as *virtual and mixed reality* (VMR) applications because they all have some degree of virtual components introduced into the application.

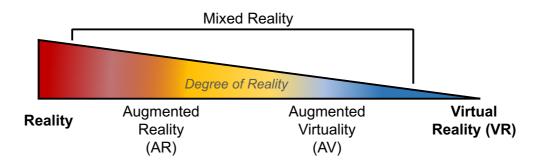


Fig 3 – The reality-virtuality continuum. AR – augmented reality; AV – augmented virtuality; VR – virtual reality (based on Milgram et al., 1995).

Box 2 – Supervised and unsupervised machine learning.

Machine learning models can be broadly classified into *supervised* or *unsupervised* learning algorithms, depending on the structure of the data (Mitchell, Michalski, and Carbonell 2013) (Note: there are intermediate cases called *semi-supervised learning* which we will not consider here, see e.g., (Zhu and Goldberg 2009) for details). Unsupervised learning algorithms use data in which the outcome is not yet labelled or identified, and therefore the algorithm cannot 'know' the outcomes in advance. The algorithm then learns how to classify and predict the outcome from new observations based on the inherent structure of the data at hand. An example of unsupervised learning is the clustering of groups within a dataset (Fig 4a). Conversely, supervised learning algorithms uses data in which the outcome is known, and the algorithm learns how to predict the outcome of future observations based on what was learnt from the information and outcomes obtained from previous data. An example of supervised learning is the classification (or prediction, in the case of regression models) of a new observation between two categories based on *n* number of characteristics or variables (Fig 4b).

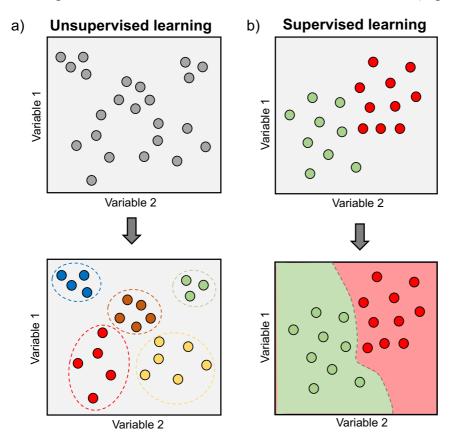


Fig 4– Supervised and unsupervised machine learning. a) Schematic representation of an unsupervised learning model. Unlabelled data is used in unsupervised learning algorithms for

clustering. b) Schematic representation of a supervised learning model. Labelled data are used
 in supervised learning algorithms for classification.