NATIONAL TECHNICAL UNIVERSITY OF ATHENS
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING
DIVISION OF INFORMATION TRANSMISSION SYSTEMS AND MATERIAL TECHNOLOGY

Leveraging sensor data for real time recognition of engagement in adaptive serious games for health

Doctoral Thesis

of

Konstantinos Mitsis

Athens, February 2023

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ

ΣΧΟΛΗ ΗΛΕΚΤΡΟΛΟΓΩΝ ΜΗΧΑΝΙΚΩΝ ΚΑΙ ΜΗΧΑΝΙΚΩΝ ΥΠΟΛΟΓΙΣΤΩΝ

ΤΟΜΕΑΣ ΣΥΣΤΗΜΑΤΩΝ ΜΕΤΑΔΟΣΗΣ ΠΛΗΡΟΦΟΡΙΑΣ ΚΑΙ ΤΕΧΝΟΛΟΓΙΑΣ ΥΛΙΚΩΝ

Αναγνώριση της προσήλωσης χρήστη σε δυναμικά μεταβαλλόμενα παιχνίδια σοβαρού σκοπού για την υγεία με χρήση δεδομένων από αισθητήρες

Διδακτορική Διατριβή

του

Κωνσταντίνου Μήτση

Αθήνα, Φεβρουάριος 2023

NATIONAL TECHNICAL UNIVERSITY OF ATHENS

SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

DIVISION OF INFORMATION TRANSMISSION SYSTEMS AND MATERIAL TECHNOLOGY

Leveraging sensor data for real time recognition of engagement in adaptive serious games for health

Doctoral Thesis

Konstantinos Mitsis

Advisory Committee: Prof. Konstantina Nikita (Supervisor)

Prof. Dionysios-Dimitrios Koutsouris

Prof. Georgios Stamou

Approved by the exam committee on February 13, 2023.

....
Konstantina Nikita
Professor NTUA

....
Dionysios-Dimitrios Koutsouris
Professor Emeritus NTUA

....
Georgios Stamou
Professor NTUA

....
Georgios Matsopoulos
Professor NTUA

....
Panagiotis Tsanakas
Professor NTUA

....
Athanasios Voulodimos
Assistant Professor NTUA

....
Spireta Golemati
Associate Professor NKUA

Athens, February 2023

....
Κωνσταντίνος Π. Μήτσης / Konstantinos Mitsis

Διδάκτωρ Ηλεκτρολόγος Μηχανικός και Μηχανικός Υπολογιστών Ε.Μ.Π.

Abstract

The aim of the present Doctoral Thesis is the development of a novel conceptual framework for personalization in serious games (SGs) for health. The proposed framework leverages sensor data for the recognition of player engagement during interaction with SGs for health in real time. This approach aims to automatically generate game content based on player engagement, in-game performance, and health-related needs. In the present thesis, two novel SGs for health are designed and developed, aiming to promote self-health management in chronic health conditions and incorporating mechanics to facilitate procedural generation of content. A novel technique, based on a genetic algorithm, that employs heterogeneous data for procedural content generation (PCG) in SGs for health is presented. Two carefully designed experimental processes are implemented to investigate the feasibility of the proposed framework.

The experimental processes collect data from sensors and interaction with the SG for health and investigate matters of player experience, educational value of the intervention, and efficiency of the proposed PCG technique. The two employed SGs for health aim to promote food and nutrition literacy and raise awareness and promote self-health management for obstructive sleep apnea, respectively. Analysis for sensor-based real-time recognition of engagement is conducted by approximating the ground truth, in terms of perceived engagement, through continuous annotations by participants. Results indicate that the educational value of the first SG for health is similar to a traditional intervention and demonstrate the predictive capacity of features extracted from the collected data towards perceived player engagement. In addition, statistically significant differences are revealed in terms of player experience in correlation with the generated content by the PCG. Furthermore, the PCG technique's capacity to rely on clinically relevant sensor data to produce tailored game content is investigated in a pre-pilot study employing a SG for health for children with type 1 diabetes and/or obesity. Results display the PCG's effectiveness in generating useful and relevant content, tailored to player needs, based on sensor and platform interaction data.

Finally, the proposed PCG technique is evaluated with the use of deep reinforcement learning (DRL) agents for automated playing. Results indicate an overall superiority in DRL agents' training when exposed to SG content produced by the proposed PCG technique. Overall findings included in the present Doctoral Thesis advocate towards the feasibility of the introduced conceptual framework. The insights gained from the experimental processes provide convincing arguments towards creating a closed real-time engagement feedback loop in adaptive SG for health, based on sensor and interaction data.

Keywords: serious games, health, adaptivity, procedural content generation, sensors, real time recognition, engagement

Περίληψη

Στόχος της παρούσας Διδακτορικής Διατριβής είναι η ανάπτυξη ενός πρωτότυπου εννοιολογικού πλαισίου που επιτρέπει την ενίσχυση και εξατομίκευση της παρέμβασης που προσφέρουν έξυπνα παιχνίδια σοβαρού σκοπού για την υγεία (ΠΣΣΥ). Το προτεινόμενο πλαίσιο αξιοποιεί δεδομένα που συλλέγονται από αισθητήρες για την αναγνώριση των επιπέδων προσήλωσης χρήστη σε πραγματικό χρόνο κατά την αλληλεπίδρασή του με ΠΣΣΥ. Η προσέγγιση αυτή στοχεύει στη δυναμική προσαρμογή του περιεχομένου των παιχνιδιών, με βάση τα εκτιμώμενα επίπεδα προσήλωσης, την επίδοση στο παιχνίδι και τις ιδιαίτερες ανάγκες υγείας του χρήστη. Στη διατριβή, σχεδιάζονται και αναπτύσσονται δύο πρωτότυπα ΠΣΣΥ που αποσκοπούν στην ενημέρωση, στην εκπαίδευση και στην ενίσχυση της ικανότητας αυτοδιαχείρισης χρόνιων νοσημάτων και ενσωματώνουν μηχανισμούς που επιτρέπουν την αυτοματοποιημένη προσαρμογή περιεχομένου (ΑΠΠ). Προτείνεται μια πρωτότυπη μεθοδολογία ΑΠΠ που βασίζεται σε γενετικό αλγόριθμο και στοχεύει στο βέλτιστο και αυτόματο έλεγχο του περιεχόμενου ΠΣΣΥ με βάση συλλεγόμενα ετερογενή δεδομένα. Το προτεινόμενο εννοιολογικό πλαίσιο αξιολογείται σύμφωνα με δύο κατάλληλα σχεδιασμένες πειραματικές διαδικασίες.

Οι πειραματικές διαδικασίες συλλέγουν δεδομένα από αισθητήρες και την αλληλεπίδραση με ΠΣΣΥ και διερευνούν ζητήματα εμπειρίας χρήστη, εκπαιδευτικής αξίας της παρέμβασης και επίδοσης της προτεινόμενης μεθοδολογίας ΑΠΠ. Τα δύο ΠΣΣΥ που αξιοποιούνται στοχεύουν στην ενίσχυση της διατροφικής παιδείας και την ενημέρωση και ενίσχυση της αυτοδιαχείρισης της υπνικής άπνοιας αντίστοιχα. Πραγματοποιείται ανάλυση για αναγνώριση των επιπέδων προσήλωσης σε πραγματικό χρόνο αξιοποιώντας επισημειώσεις της αντίληψης προσήλωσης από το χρήστη σε πραγματικό χρόνο. Τα αποτελέσματα υποδεικνύουν ότι το πρώτο ΠΣΣΥ έχει παρόμοια εκπαιδευτική αξία με παραδοσιακή παρέμβαση και επιδεικνύουν την προβλεπτική ικανότητα χαρακτηριστικών που εξάχθηκαν από τα συλλεγόμενα δεδομένα προς τα αντιληπτά επίπεδα προσήλωσης του χρήστη. Παράλληλα, αποκαλύπτονται στατιστικά σημαντικές διαφορές στην εμπειρία χρήστη που σχετίζονται με το παραγόμενο περιεχόμενο από την ΑΠΠ. Η ικανότητα της προτεινόμενης μεθοδολογίας ΑΠΠ να συνυπολογίζει δεδομένα με κλινική αξία που συλλέγονται από αισθητήρες για να παράγει εξατομικευμένο περιεχόμενο αξιολογείται σε προ-πιλοτική μελέτη που αξιοποιεί ΠΣΣΥ για παιδιά με σακχαρώδη διαβήτη τύπου 1 ή/και παχυσαρκία. Τα αποτελέσματα παρουσιάζουν την ικανότητα της μεθοδολογίας ΑΠΠ να προβάλει χρήσιμο και σχετικό με τις ανάγκες του χρήστη περιεχόμενο παιχνιδιού, βασιζόμενη σε δεδομένα αισθητήρων και αλληλεπίδρασης.

Τέλος, η προτεινόμενη μεθοδολογία ΑΠΠ ελέγχεται με τη χρήση πρακτόρων βαθιάς μάθησης που εκπαιδεύονται να διατρέχουν το ΠΣΣΥ αυτόματα. Τα αποτελέσματα αποδεικνύουν την ανωτερότητα της εκπαίδευσης πρακτόρων που εκτίθενται στο περιεχόμενο που παράγεται αυτόματα από την προτεινόμενη μεθοδολογία ΑΠΠ. Τα συμπεράσματα που παρουσιάζονται στην παρούσα Διδακτορική Διατριβή υποστηρίζουν τη δυνατότητα υλοποίησης του προτεινόμενου εννοιολογικού πλαισίου. Τα αποτελέσματα των πειραματικών διαδικασιών συνηγορούν στη δημιουργία ενός κλειστού βρόχου ανατροφοδότησης της προσήλωσης χρήστη σε δυναμικά μεταβαλλόμενα ΠΣΣΥ, αξιοποιώντας δεδομένα αισθητήρων και αλληλεπίδρασης.

Λέξεις κλειδιά: παιχνίδια σοβαρού σκοπού, υγεία, προσαρμοστικότητα, αυτοματοποιημένη παραγωγή περιεχομένου, αισθητήρες, αναγνώριση, προσήλωση

Εκτεταμένη περίληψη

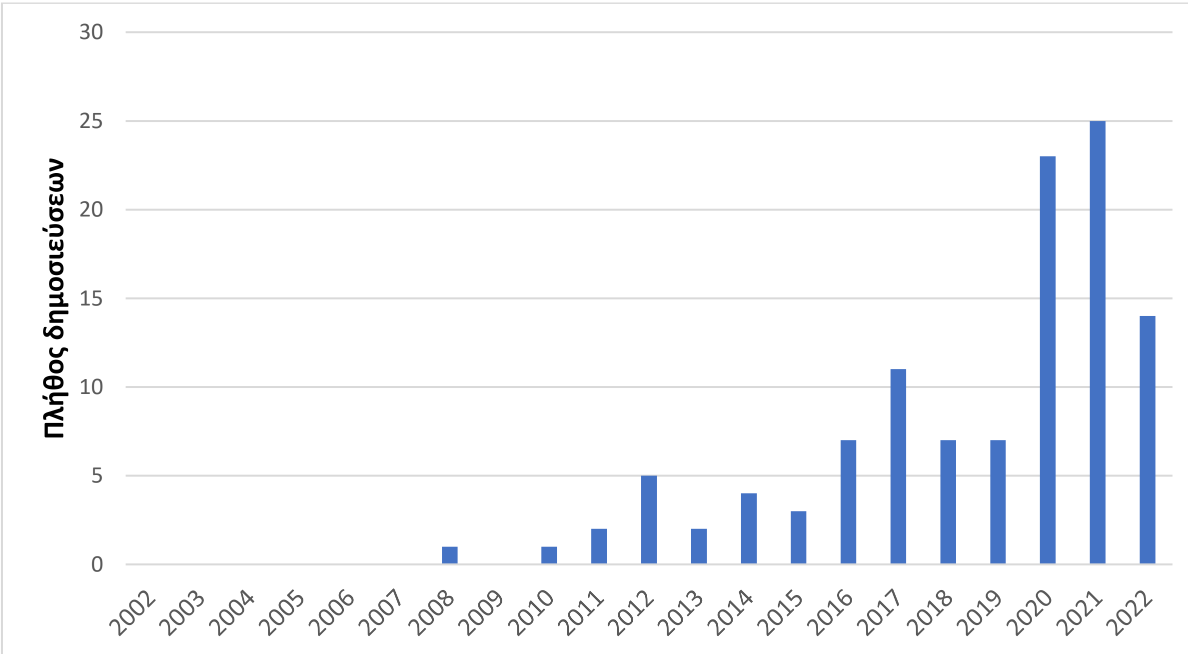
Στόχος της παρούσας Διδακτορικής Διατριβής είναι η ανάπτυξη ενός πρωτότυπου εννοιολογικού πλαισίου που επιτρέπει την ενίσχυση και εξατομίκευση της παρέμβασης που προσφέρουν έξυπνα παιχνίδια σοβαρού σκοπού για την υγεία (ΠΣΣΥ). Για τις ανάγκες του προτεινόμενου πλαισίου αξιοποιούνται δεδομένα που συλλέγονται κατά την αλληλεπίδραση του χρήστη με ΠΣΣΥ από αισθητήρες για την αναγνώριση των επιπέδων προσήλωσης χρήστη σε πραγματικό χρόνο. Η προσέγγιση αυτή στοχεύει στη δυναμική και αυτόματη προσαρμογή του περιεχομένου ΠΣΣΥ, με βάση τα εκτιμώμενα επίπεδα προσήλωσης, την επίδοση στο παιχνίδι και τις ιδιαίτερες ανάγκες υγείας του χρήστη. Για την αξιολόγηση του προτεινόμενου πλαισίου σχεδιάζονται και αναπτύσσονται δύο πρωτότυπα ΠΣΣΥ που αποσκοπούν στην ενημέρωση, στην εκπαίδευση και στην ενίσχυση της ικανότητας αυτοδιαχείρισης χρόνιων νοσημάτων. Τα παιχνίδια αυτά ενσωματώνουν μηχανισμούς που διευκολύνουν την αυτοματοποιημένη προσαρμογή περιεχομένου τους σε πραγματικό χρόνο. Παράλληλα, σχεδιάζεται και υλοποιείται πρωτότυπη μεθοδολογία αυτοματοποιημένης παραγωγής περιεχομένου (ΑΠΠ) που βασίζεται στη λειτουργία γενετικού αλγόριθμου (ΓΑ). Η μεθοδολογία αυτή αξιοποιεί ετερογενή δεδομένα για το βέλτιστο και αυτόματο έλεγχο του περιεχόμενου ΠΣΣΥ. Το προτεινόμενο εννοιολογικό πλαίσιο αξιολογείται σύμφωνα με δύο κατάλληλα σχεδιασμένες πειραματικές διαδικασίες. Στη συνέχεια, η προτεινόμενη μεθοδολογία ΑΠΠ ενσωματώνεται σε πλατφόρμα παρέμβασης που περιλαμβάνει ΠΣΣΥ και αξιολογείται με βάση δεδομένα από αισθητήρες παρακολούθησης σωματικής δραστηριότητας και διαχείρισης χρόνιων νοσημάτων. Η Διατριβή ολοκληρώνεται με τη χρήση ευφυών πρακτόρων βαθιάς ενισχυτικής μάθησης που αξιολογούν αυτόματα τη ικανότητα παραγωγής περιεχομένου σε ΠΣΣΥ.

1. Προσαρμοστικά παιχνίδια σοβαρού σκοπού για την υγεία

Τα παιχνίδια σοβαρού σκοπού (ΠΣΣ) αποτελούν πεδίο που συγκεντρώνει αυξανόμενο ερευνητικό ενδιαφέρον την τελευταία δεκαετία. Παράλληλα, παρατηρείται σημαντική άνθιση και σε επιχειρηματικό επίπεδο με εκτιμήσεις να προβλέπουν περαιτέρω εξέλιξή τα επόμενα χρόνια σε παγκόσμιο επίπεδο [13], [14]. Η πρώτη αναφορά του όρου ΠΣΣ αποδίδεται σε βιβλίο που εκδόθηκε το 1970 από τον Clark Abt [17]. Από τότε τα ΠΣΣ έχουν εξελιχθεί σημαντικά και ενσωματώνουν τεχνολογίες αιχμής από τομείς όπως η τεχνητή νοημοσύνη και η εικονική και επαυξημένη πραγματικότητα. Στη βιβλιογραφία απαντάται πλήθος ορισμών σχετικά με το τι συνιστά ΠΣΣ, με έναν από τους πιο διαδεδομένους να αναφέρει ότι τα ΠΣΣ αποτελούν παιχνίδια που έχουν κάποιο βασικό στόχο πέρα από αυτό της διασκέδασης [19]. Η διαδικασία της σχεδίασης και ανάπτυξης ενός ΠΣΣ θεωρείται δύσκολο και απαιτητικό έργο, το οποίο χρειάζεται πολύ-επιστημονική προσέγγιση [45]. Οι παρεμβάσεις αυτές αξιοποιούν μηχανισμούς παιχνιδιού για πετύχουν το «σοβαρό» στόχο τους ενώ παρέχουν μια εμπειρία που χαρακτηρίζεται από υψηλά επίπεδα προσήλωσης. Για την ενίσχυση της ποιότητας της παρέμβασης που παρέχουν σημαντική είναι η ύπαρξη κατάλληλου σχεδιαστικού πλαισίου που θεμελιώνει το σκοπό του παιχνιδιού με σαφή τρόπο [47]. Με την εξέλιξή τους, τα ΠΣΣ θεωρούνται πλέον αποδοτικά εργαλεία παρέμβασης που στοχεύουν, ανάμεσα σε άλλα, στην εκπαίδευση, στην απόκτηση ικανοτήτων, στην υιοθέτηση συμπεριφορών και στην ενημέρωση [1]. Τα εργαλεία αυτά βρίσκουν εφαρμογή σε πλήθος πεδίων ενδιαφέροντος, όπως η εκπαίδευση, η προσομοίωση, η υγεία, η παράδοση και ο τουρισμός [20]. Η υγεία αποτελεί έναν από τους πιο σημαντικούς τομείς εφαρμογής των ΠΣΣ [73]. Τα ΠΣΣΥ λειτουργούν ως εργαλεία κινητής και ηλεκτρονικής υγείας με πληθώρα τελικών χρηστών όπως ασθενείς, το άμεσο περιβάλλον τους, υγιή άτομα, επαγγελματίες υγείας και φοιτητές. Τα ΠΣΣ αποτελούν πλέον καινοτόμες παρεμβάσεις υγείας που στοχεύουν στην αποκατάσταση, στην ενημέρωση, στην παραγωγή συστάσεων, στη διάγνωση, στην παρακολούθηση, στην ενίσχυση συμπεριφορών, στην αυτοδιαχείριση νοσημάτων, στην εκπαίδευση και στη θεραπεία [74], [75]. Μερικά πρόσφατα παραδείγματα τέτοιων παρεμβάσεων αποτελούν ΠΣΣΥ για σχεδίαση χειρουργικών επεμβάσεων [79], τη μείωση του στίγματος σχετικά με το HIV [80] και εξατομικευμένες παρεμβάσεις ψυχικής υγείας [83].

Στο πλαίσιο της παρούσας Διδακτορικής Διατριβής πραγματοποιείται εκτενής διερεύνηση συναφούς βιβλιογραφίας και συστηματική αναζήτηση άρθρων ανασκόπησης που αναλύουν το πεδίο των ΠΣΣΥ τις τελευταίες δύο δεκαετίες. Το πλήθος των δημοσιεύσεων παρουσιάζεται στην Εικόνα 1 και υποδηλώνει την αυξητική τάση που παρατηρείται στην έρευνα στο πεδίο. Από το 2016 και έπειτα είναι φανερή η άνοδος στο πλήθος και τη συχνότητα δημοσίευσης άρθρων ανασκόπησης που αφορούν ΠΣΣΥ. Παράλληλα, πραγματοποιείται ταξινόμηση των άρθρων, ανάλογα με το επιθυμητό αποτέλεσμα των παρεμβάσεων που αναλύουν και σχολιάζουν. Συνολικά, αποτυπώνεται το γεγονός ότι τα ΠΣΣΥ αποτελούν ουσιαστικές παρεμβάσεις για πληθώρα διαφορετικών παθήσεων και αντικειμένων υγείας. Η εκπαίδευση επαγγελματιών υγείας, η ψυχική υγεία, η διαχείριση χρόνιων νοσημάτων και η αποκατάσταση φαίνεται πως αποτελούν τα πεδία που συγκεντρώνουν το μεγαλύτερο ενδιαφέρον της έρευνας. Το ενδιαφέρον που παρατηρείται στις παρεμβάσεις αυτές ενισχύει την πεποίθηση ότι έχουν τη δυνατότητα να αποτελέσουν ουσιαστικά εργαλεία που ικανοποιούν τις ολοένα και αυξανόμενες ανάγκες της υγείας για προβλεπτικές, αποτρεπτικές, εξατομικευμένες και συμμετοχικές προσεγγίσεις που αντιμετωπίζουν σύγχρονα προβλήματα. Η δυνατότητα των παρεμβάσεων αυτών να παρέχουν εξατομικευμένη αντιμετώπιση του χρήστη με αυτόματο τρόπο, εστιάζοντας σε συγκεκριμένες ανάγκες ενώ παράλληλα ενισχύουν τη συμμετοχή του στη διαχείριση της υγείας του χαρακτηρίζεται ως ιδιαίτερα σημαντική [43]. Στο πλαίσιο της διερεύνησης αναγνωρίστηκε περιορισμένος αριθμών δημοσιεύσεων που αναλύουν την ικανότητα των ΠΣΣΥ για αυτόματη προσαρμογή του περιεχομένου τους προς αυτή την κατεύθυνση. Παρ’ όλα αυτά, συνυπολογίζοντας την πρόσφατη δημοσίευσή τους, την τάση που παρατηρείται στο πεδίο των ΠΣΣ συνολικά και τις μελλοντικές κατευθύνσεις της έρευνας, όπως αποτυπώνονται στην βιβλιογραφία, η αναδυόμενη αξία από την ενσωμάτωση τέτοιας προσαρμοστικότητας σε ΠΣΣΥ αναμένεται να απασχολήσει σημαντικά τους ερευνητές τα επόμενα χρόνια.

Η προσαρμοστικότητα αποτελεί σημαντικό πεδίο έρευνας για τα ψηφιακά παιχνίδια. Η έννοια μπορεί να οριστεί ως, η δυνατότητα για κατάλληλη προσαρμογή σε μεταβαλλόμενες συνθήκες [88]. Οι τεχνικές που αξιοποιούνται για την ενσωμάτωση δυνατότητας προσαρμογής των παιχνιδιών εξελίσσονται συνεχώς και βασίζονται κυρίως στον τομέα της τεχνητής νοημοσύνης.



Εικόνα 1: Πλήθος δημοσιευμένων άρθρων ανασκόπησης σε παιχνίδια σοβαρού σκοπού για την υγεία.

Τα μοντέρνα ψηφιακά παιχνίδια έχουν τη δυνατότητα να πραγματοποιούν μοντελοποίηση των παικτών, να παράγουν δυναμικά περιεχόμενο, να μεταβάλλουν τα επίπεδα δυσκολίας τους αυτόματα και να περιλαμβάνουν χαρακτήρες με ρεαλιστική συμπεριφορά. Για την επίτευξη αυτών των δυνατοτήτων συχνά αξιοποιούνται τεχνικές ΑΠΠ. Τα πλεονεκτήματα από την ενσωμάτωση τέτοιων τεχνικών είναι σημαντικά καθώς ενισχύουν την προσήλωση του χρήστη και βελτιώνουν την εμπειρία παιχνιδιού [91]. Τα ΠΣΣ, ως εργαλεία παρέμβασης, μπορούν να εκμεταλλευτούν τα πλεονεκτήματα αυτά και να μεταβάλλουν το περιεχόμενό τους ανάλογα με τις ανάγκες του χρήστη. Με τον τρόπο αυτό δεν περιορίζονται στην ενίσχυση των επιπέδων προσήλωσής του, αλλά βελτιώνουν και την απόδοση της παρέμβασης που παρέχουν. Συγκεκριμένα για τα ΠΣΣΥ, οι ανάγκες του χρήστη μπορούν να βασίζονται και σε παραμέτρους και χαρακτηριστικά που αφορούν την υγεία του. Προς αυτή την κατεύθυνση σημαντική χαρακτηρίζεται η μοντελοποίηση των συναισθημάτων του χρήστη κατά την αλληλεπίδρασή του με ΠΣΣΥ. Τέτοιες τεχνικές μπορούν να οδηγήσουν στην αναγνώριση των επιπέδων προσήλωσής του σε πραγματικό χρόνο αξιοποιώντας δεδομένα που προκύπτουν από αισθητήρες και το ίδιο το παιχνίδι. Η προσήλωση αποτελεί σημαντική ένδειξη για την ποιότητα της εμπειρίας χρήστη και μπορεί να ενισχύσει σημαντικά την παρεχόμενη παρέμβαση [110]. Οι εξελίξεις στην τεχνολογία καθιστούν εφικτές τέτοιες προσεγγίσεις και επιτρέπουν τη δημιουργία εξατομικευμένων παρεμβάσεων υγείας που συνυπολογίζουν τις προτιμήσεις του χρήστη.

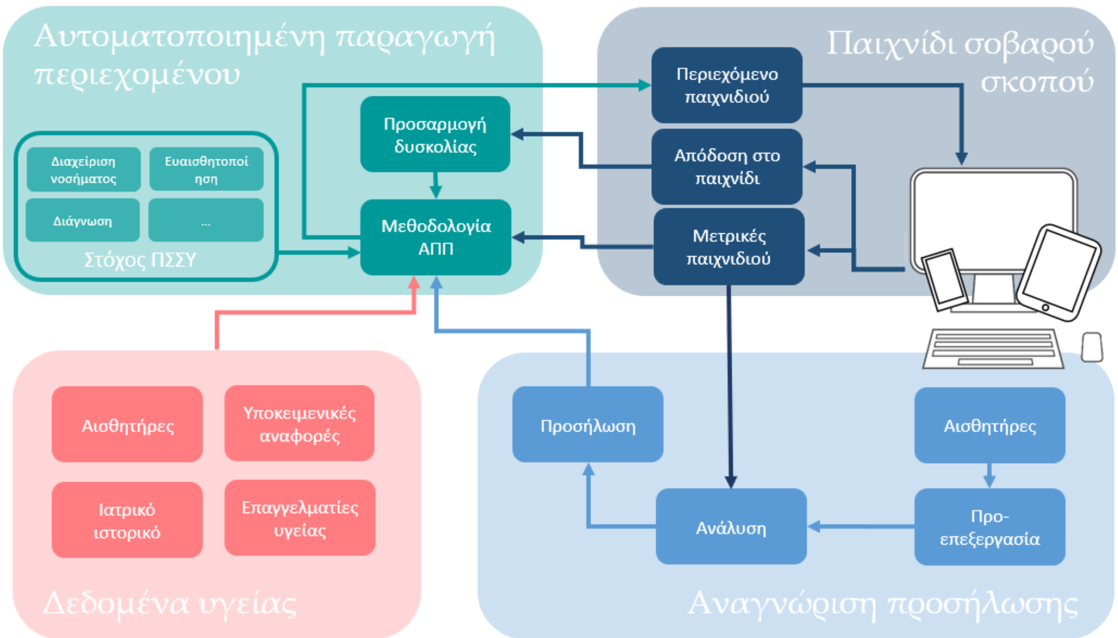
2. Εννοιολογικό πλαίσιο για προσαρμοστικότητα σε παιχνίδια σοβαρού σκοπού για την υγεία

Για τη δημιουργία έξυπνων και εξατομικευμένων παρεμβάσεων υγείας που βασίζονται σε ΠΣΣ παρουσιάζεται πρωτότυπο εννοιολογικό πλαίσιο που επιτρέπει τη δυναμική προσαρμογή του περιεχομένου τους. Το προτεινόμενο πλαίσιο αξιοποιεί καινοτόμες τεχνολογίες και υποστηρίζει τη δυναμική προσαρμογή των ΠΣΣΥ σε δύο κατευθύνσεις. Η πρώτη κατεύθυνση στοχεύει στην ενίσχυση της προσήλωσης του χρήστη στην παρέμβαση που προσφέρει το ΠΣΣΥ μέσα από τη μοντελοποίηση της σε πραγματικό χρόνο. Η δεύτερη κατεύθυνση στοχεύει στην αναγνώριση των αναγκών υγείας του χρήστη και στην παραγωγή περιεχομένου κατάλληλο για αυτές. Ένα πλαίσιο που επιτρέπει τη δυναμική προσαρμογή ενός ΠΣΣΥ βασίζεται σε τρία επίπεδα. Αρχικά δεδομένα για το χρήστη, σχετικά με το επιθυμητό περιεχόμενο παιχνιδιού, συλλέγονται κατά τη διάρκεια της αλληλεπίδρασής του με το ΠΣΣΥ. Η συλλογή των δεδομένων μπορεί να πραγματοποιηθεί με στατικό τρόπο, όπως μέσω ερωτηματολογίων ή ιατρικού φακέλου, ή με δυναμικό τρόπο, όπως μέσω αισθητήρων ή μετρικών που καταγράφονται από το παιχνίδι. Στη συνέχεια, με βάση τα συλλεγόμενα δεδομένα, κατασκευάζεται και ανανεώνεται το μοντέλο χρήστη. Το μοντέλο περιγράφει και περιλαμβάνει τις ιδιαίτερες ανάγκες και τις προτιμήσεις του χρήστη. Τέλος, σχεδιάζεται και αναπτύσσεται ευφυής πράκτορας που τροφοδοτείται από το μοντέλο χρήστη και ελέγχει την επιλογή και παραγωγή κατάλληλου περιεχομένου του ΠΣΣΥ.

Το προτεινόμενο εννοιολογικό πλαίσιο που παρουσιάζεται στην παρούσα Διδακτορική Διατριβή στοχεύει κυρίως σε ΠΣΣΥ που ενισχύουν την ικανότητα αυτό-διαχείρισης χρόνιων νοσημάτων του χρήστη, αλλά μπορεί να γενικευτεί και σε άλλες εφαρμογές. Το πλαίσιο αξιοποιεί ετερογενή δεδομένα και τεχνικές μοντελοποίησης χρήστη για να πετύχει το στόχο του. Η ενίσχυση του ενδιαφέροντος του χρήστη στην παρέμβαση που προσφέρει το ΠΣΣΥ πραγματοποιείται μέσα από ένα κλειστό βρόχο ελέγχου που βελτιστοποιεί την προσήλωσή του. Παράλληλα, το επίπεδο δυσκολίας του παιχνιδιού μεταβάλλεται δυναμικά ανάλογα με την επίδοση του χρήστη. Το προτεινόμενο πλαίσιο συνυπολογίζει στο σχεδιασμό του προκλήσεις που παρουσιάζονται σε παρεμβάσεις υγείας, όπως η μεταβλητότητα των αναγκών που παρατηρείται στη διαχείριση χρόνιων νοσημάτων και η δυσκολία στην αποδοτική προώθηση αλλαγών συμπεριφοράς. Η ενσωμάτωση της δυνατότητας στο ΠΣΣΥ να παράγει εξατομικευμένο περιεχόμενο ενισχύει την απόδοση της παρέμβασης που προσφέρεται και αντιμετωπίζει τις προκλήσεις αυτές. Για την επίτευξη αυτών των στόχων το πλαίσιο αξιοποιεί οικονομικούς και μη επεμβατικούς αισθητήρες οι οποίοι αναγνωρίζουν τα επίπεδα προσήλωσης και παρακολουθούν τις ανάγκες υγείας του χρήστη. Όπως φαίνεται στην Εικόνα 2 το προτεινόμενο εννοιολογικό πλαίσιο αποτελείται από τέσσερα επίπεδα.

Το πρώτο επίπεδο αντιπροσωπεύει το ΠΣΣΥ και περιλαμβάνει το περιεχόμενο του παιχνιδιού, τους μηχανισμούς αναγνώρισης της επίδοσης του χρήστη και τις μετρικές παιχνιδιού που συλλέγονται κατά τη διάρκεια της αλληλεπίδρασης. Η επίδοση του χρήστη αναγνωρίζεται με βάση μηχανισμούς επιβράβευσης που περιλαμβάνονται στο ΠΣΣΥ και παρουσιάζει ιδιαίτερη σημασία στην αξιολόγηση της επιτυχίας της παρέμβασης. Οι μετρικές παιχνιδιού αναφέρονται σε δεδομένα που αφορούν τη φύση της αλληλεπίδρασης με το παιχνίδι, όπως η διάρκεια και η συχνότητα ενασχόλησης, και χρησιμοποιούνται συχνά κατά τη μοντελοποίηση της εμπειρίας παίκτη. Το δεύτερο επίπεδο αποτελείται από τον ευφυή πράκτορα που ελέγχει την επιλογή και παραγωγή περιεχομένου παιχνιδιού. Στο επίπεδο αυτό αξιοποιούνται μέθοδοι ΑΠΠ οι οποίες στοχεύουν στην παραγωγή περιεχομένου που ταιριάζει στις ιδιαίτερες ανάγκες του χρήστη, με βάση την πληροφορία που τροφοδοτείται από τα υπόλοιπα επίπεδα. Το τρίτο επίπεδο αναγνωρίζει τα επίπεδα προσήλωσης του χρήστη σε πραγματικό χρόνο κατά την αλληλεπίδρασή του με το ΠΣΣΥ. Για την επίτευξη της αναγνώρισης αξιοποιούνται δεδομένα από αισθητήρες και πληροφορία που συλλέγονται από το επίπεδο του ΠΣΣ. Το τελευταίο επίπεδο συλλέγει ετερογενή δεδομένα που αφορούν τις ανάγκες υγείας του χρήστη και τις τροφοδοτεί στο δεύτερο επίπεδο που ελέγχει το περιεχόμενο του παιχνιδιού. Συγκρινόμενο με παρόμοια πλαίσια που αναγνωρίζονται στη βιβλιογραφία, το πλαίσιο που προτείνεται στην παρούσα Διδακτορική Διατριβή διαφέρει στη στόχευση του για προσαρμογή του ΠΣΣ σε πραγματικό χρόνο, ενώ ικανοποιεί το διπλό του σκοπό.

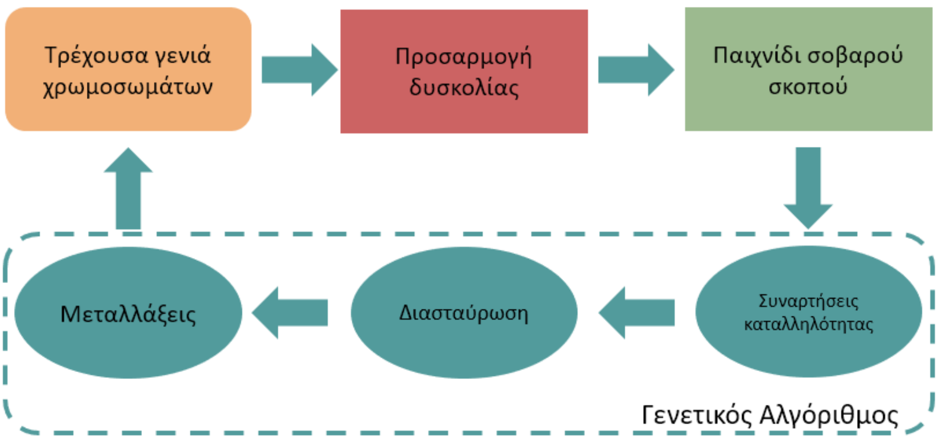
Στο πλαίσιο της παρούσας Διδακτορικής διατριβής σχεδιάζεται και αναπτύσσεται πρωτότυπη μεθοδολογία ΑΠΠ που βασίζεται σε ΓΑ. Η μεθοδολογία αυτή εστιάζει στην παραγωγή περιεχομένου για ΠΣΣΥ που αφορούν χρόνια νοσήματα, με στόχο την αξιοποίησή της για την αξιολόγηση του προτεινόμενου εννοιολογικού πλαισίου. Ο ΓΑ σχεδιάζεται για να αξιοποιεί ετερογενή δεδομένα που προκύπτουν από τα υπόλοιπα επίπεδα του πλαισίου. Παράλληλα, περιλαμβάνει τη δυνατότητα να μεταβάλλει δυναμικά τη δυσκολία παιχνιδιού, ενισχύοντας την προσαρμοστικότητα του ΠΣΣΥ στο οποίο ενσωματώνεται. Οι ΓΑ αποτελούν ευριστικές μεθόδους



Εικόνα 2: Προτεινόμενο εννοιολογικό πλαίσιο για προσαρμοστικότητα σε παιχνίδια σοβαρού σκοπού για την υγεία.

που ανήκουν στην ευρεία κατηγορία των εξελικτικών αλγορίθμων και βασίζουν τη λειτουργία τους στη θεωρία της εξελικτικής βιολογίας [139]. Οι ΓΑ χρησιμοποιούνται συχνά για παραγωγή περιεχομένου με αυτόματο τρόπο τόσο σε παιχνίδια για διασκέδαση [141], όσο και σε ΠΣΣ [147], εξελίσσοντας το περιεχόμενο παιχνιδιού ενώ αναζητούν λύση από ένα χώρο ενδεχομένων με βάση κριτήρια που ορίζονται. Παράλληλα, η δυνατότητα τους να ανανεώνουν το περιεχόμενο κατά τη διάρκεια του παιχνιδιού προσδίδει σημαντικό πλεονέκτημα στην εφαρμογή τους σε τεχνικές ΑΠΠ που λειτουργούν σε πραγματικό χρόνο παιχνιδιού.

Η ροή λειτουργίας του ΓΑ παρουσιάζεται στην Εικόνα 3. Το περιεχόμενο του ΠΣΣΥ περιγράφεται ως σειρά γονιδίων που σχηματίζουν το χρωμόσωμα της μεθοδολογίας. Τα γονίδια λαμβάνουν δυαδικές τιμές, με την τιμή «1» να αντιστοιχεί σε παρουσία του περιεχομένου στο ΠΣΣΥ και την τιμή «0» σε απουσία του. Με τον τρόπο αυτό ένα χρωμόσωμα μπορεί να δημιουργήσει ή να επιλέξει περιεχόμενο παιχνιδιού. Σε κάθε γενιά του ΓΑ ένα ή περισσότερα γονίδια επιλέγονται με βάση συναρτήσεις καταλληλότητας που σχεδιάζονται ανάλογα με το στόχο του ΠΣΣΥ. Βάρη αναθέτονται σε κάθε γονίδιο και εκπαιδεύονται με βάση ετερογενή δεδομένα που συλλέγονται από τα υπόλοιπα επίπεδα του εννοιολογικού πλαισίου. Στη συνέχεια η αξία του κάθε χρωμοσώματος, η οποία προκύπτει από τα εκπαιδευμένα βάρη, υπολογίζεται από συναρτήσεις καταλληλότητας. Τα καταλληλότερα γονίδια θα χρησιμοποιηθούν για να δημιουργήσουν την επόμενη γενιά. Η φύση της διαδικασίας ανανέωσης βαρών και των συναρτήσεων καταλληλότητας εξαρτώνται από τη περιβάλλον παιχνιδιού στο οποίο εφαρμόζονται. Τέλος, ένα σύστημα κανόνων επηρεάζει δυναμικά τη δυσκολία του ΠΣΣΥ, ανάλογα με την επίδοση του χρήστη, επιλέγοντας ποιο από τα καταλληλότερα γονίδια θα αξιοποιηθεί. Στόχος της μεθοδολογίας αυτής είναι να αξιοποιήσει πληθώρα ετερογενών δεδομένων και να ελέγξει περιεχόμενο ΠΣΣΥ με αποδοτικό και ευέλικτο τρόπο.



Εικόνα 3: Αυτοματοποιημένη παραγωγή περιεχομένου παιχνιδιού σοβαρού σκοπού με χρήση γενετικού αλγορίθμου.

3. Σχεδίαση και ανάπτυξη προσαρμοστικών παιχνιδιών σοβαρού σκοπού για την υγεία

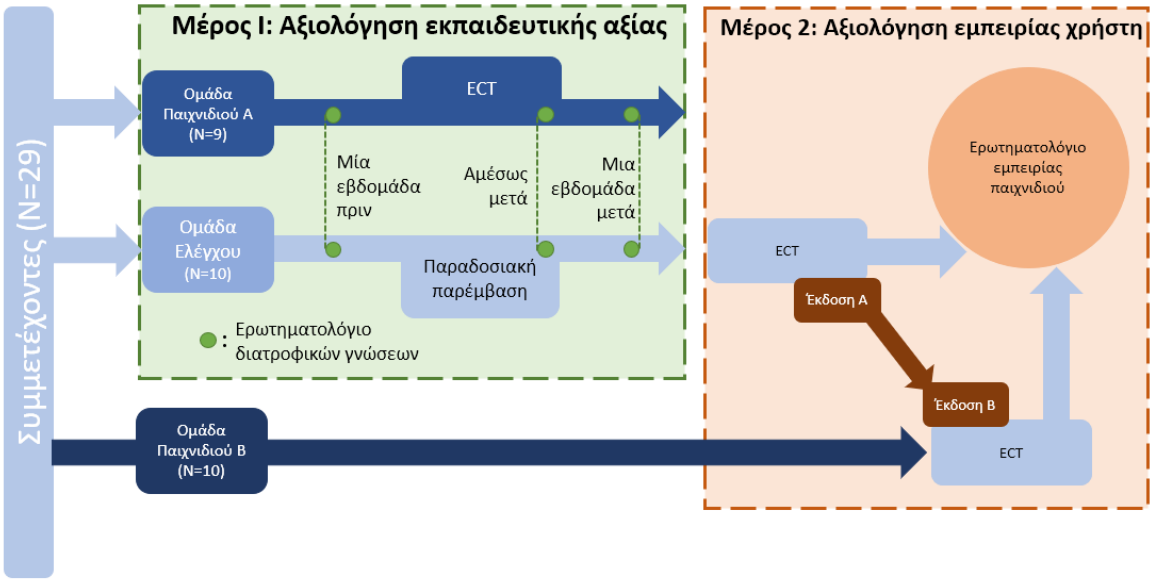
Δύο πρωτότυπα ΠΣΣΥ που στοχεύουν να παρέχουν παρέμβαση για τη διαχείριση χρόνιων νοσημάτων σχεδιάζονται και αναπτύσσονται. Το «Express cooking train» (ECT) [149] επιχειρεί να ενισχύσει τη διατροφική παιδεία σε έφηβους και νεαρούς ενήλικες, με απώτερο στόχο την υιοθέτηση σωστών διατροφικών συμπεριφορών και πιο υγιεινού τρόπου ζωής. Η έλλειψη τέτοιων συμπεριφορών έχει συνδεθεί με την εμφάνιση χρόνιων νοσημάτων, όπως ο σακχαρώδης διαβήτης, καρδιαγγειακά νοσήματα και ο καρκίνος [152], ενώ η ικανότητα και επιτυχία ατόμων στην υιοθέτησή τους περιορίζεται από την έλλειψη γνώσης σχετικά με την προετοιμασία γευμάτων, το μέγεθος μερίδων και τη διατροφική αξία υλικών. Ο σχεδιασμός του ECT αξιοποιεί ένα πρωτότυπο εννοιολογικό πλαίσιο που ενσωματώνει οντολογία συνταγών και καθοδηγείται από θεωρίες μάθησης. Σκοπός του πλαισίου αυτού είναι η βελτίωση της εμπειρίας του χρήστη και η ενίσχυση της ικανότητας του ΠΣΣΥ να λειτουργεί ως παρέμβαση υγείας. Παράλληλα, το παιχνίδι περιλαμβάνει ένα περιβάλλον προσομοίωσης κουζίνας, το οποίο τροφοδοτείται με γνώση μέσω της οντολογίας, και παρέχει δυνατότητες εξατομίκευσης μέσα από ένα ρεαλιστικό περιβάλλον παιχνιδιού. Το ECT αντλεί στοιχεία από τρία είδη μοντέρνων ψηφιακών παιχνιδιών και περιλαμβάνει μηχανισμούς και επιβραβεύσεις που μπορούν να λειτουργήσουν ως μεσολαβητές για την επίτευξη του επιθυμητού στόχου.

Το «Wake up for the future» (WuF) [125] είναι ένα ΠΣΣΥ που ενισχύει την ικανότητα αυτοδιαχείρισης και ενημερώνει σχετικά με την αποφρακτική υπνική άπνοια (ΑΥΑ). Η ΑΥΑ αποτελεί την πιο διαδεδομένη διαταραχή ύπνου και εκδηλώνεται με επαναλαμβανόμενα επεισόδια κατάρρευσης του ανώτερου αεραγωγού που οδηγούν σε μείωση ή ακόμα και διακοπή της ροής του αέρα με διάρκεια τουλάχιστον 10 δευτερολέπτων [162]. Η πάθηση θεωρείται μείζον δημόσιο ζήτημα, καθώς αντιπροσωπεύει 936 εκατομμύρια ασθενείς παγκοσμίως το 2019 [163], ενώ οι περισσότερες (80%) των περιπτώσεων παραμένουν αδιάγνωστες. Η σωστή αυτοδιαχείριση της νόσου μπορεί να ωφελήσει ασθενείς που πάσχουν από ΑΥΑ και η ευαισθητοποίηση σχετικά με τα συμπτώματα μπορεί να βοηθήσει στη μείωση της υποδιάγνωσης και στη βελτίωση της έκβασης της νόσου. Η σχεδίαση του WuF βασίζεται σε παρόμοιο πλαίσιο με αυτό που αξιοποιείται στο ECT και ενσωματώνει την προτεινόμενη μεθοδολογία ΑΠΠ. Το ΠΣΣΥ ενσωματώνει μηχανισμούς από ψηφιακά παιχνίδια με κάρτες και η σχεδίασή του στοχεύει σε ενήλικους χρήστες. Παράλληλα, το παιχνίδι χαρακτηρίζεται από σενάριο που περιλαμβάνει ανοιχτό κόσμο και προσομοιώνει αγώνες επιχειρηματολογίας με τη χρήση καρτών που περιλαμβάνουν επιχειρήματα σχετικά με την ΑΥΑ.

Η προτεινόμενη μεθοδολογία για ΑΠΠ ενσωματώνεται επίσης σε πλατφόρμα παρέμβασης [172] για παιδιά που πάσχουν από σακχαρώδη διαβήτη τύπου 1 ή/και παχυσαρκία η οποία περιλαμβάνει ΠΣΣΥ. Πέρα από το ΠΣΣΥ, η πλατφόρμα αποτελείται από δύο εφαρμογές για γιατρούς και γονείς και αξιοποιεί δεδομένα που συλλέγονται από πληθώρα αισθητήρων και την αλληλεπίδραση με αυτή. Στόχος της πλατφόρμας αποτελεί η ενίσχυση της ικανότητας αυτό-διαχείρισης των νοσημάτων μέσα από την παραγωγή εξατομικευμένου περιεχομένου σε μορφή μηνυμάτων και περιεχομένου παιχνιδιού. Το ΠΣΣΥ περιλαμβάνει εκπαιδευτικές αποστολές, οι οποίες σχεδιάστηκαν με τη βοήθεια επαγγελματιών υγείας, όπως επίσης και αποστολές για διασκέδαση. Παράλληλα ο χρήστης συλλέγει υλικά φαγητού και νομίσματα, τα οποία μπορεί να χρησιμοποιήσει μέσα στο παιχνίδι για να ετοιμάσει γεύματα και να αλλάξει την εμφάνιση του χαρακτήρα του. Μηνύματα που αφορούν την πρόοδο και επίδοσή του και εκπαιδεύουν το χρήστη σχετικά με το νόσημά του προβάλλονται σε κατάλληλο χώρο του παιχνιδιού. Τέλος, η έκθεση του παιδιού στο παιχνίδι σε ημερήσιο επίπεδο περιορίζεται μέσα από τη δυνατότητά του να παίζει μόνο δύο διαθέσιμες αποστολές.

4. Αναγνώριση προσήλωσης σε παιχνίδια σοβαρού σκοπού για την υγεία

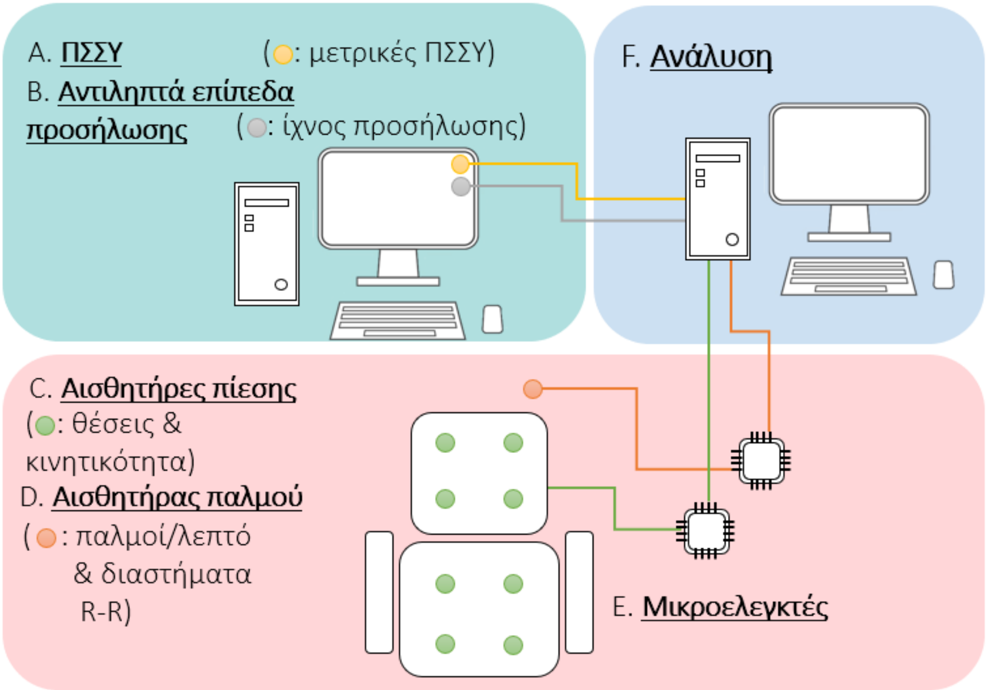
Για τη διερεύνηση της δυνατότητας αναγνώρισης των επιπέδων προσήλωσης κατά τη διάρκεια αλληλεπίδρασης με ΠΣΣΥ σε πραγματικό χρόνο σχεδιάζεται και υλοποιείται κατάλληλη πειραματική διαδικασία. Δεδομένα από αισθητήρες και το παιχνίδι, όπως περιγράφεται στο προτεινόμενο εννοιολογικό πλαίσιο, συλλέγονται κατά την αλληλεπίδραση είκοσι εννέα συμμετεχόντων με έκδοση του ECT. Παράλληλα αξιολογείται η αποτελεσματικότητα του ΠΣΣΥ ως προς τον εκπαιδευτικό στόχο του, όπως επίσης και η εμπειρία χρήστη. Στο πρώτο μέρος της πειραματικής διαδικασίας, το ECT συγκρίνεται με εναλλακτική παρέμβαση που βασίζεται στη μελέτη υλικού βασισμένου σε κείμενο. Κατάλληλα σχεδιασμένα ερωτηματολόγια που αφορούν σε διατροφικές γνώσεις απαντώνται πριν, αμέσως μετά και μία εβδομάδα μετά τις παρεμβάσεις. Στο δεύτερο μέρος της αξιολόγησης, η εμπειρία του χρήστη αποτυπώνεται μέσω του ερωτηματολογίου εμπειρίας παιχνιδιού. Κατά τη διάρκεια του παιχνιδιού συλλέγονται δεδομένα



Εικόνα 4: Σχεδίαση πειραματικής διαδικασίας.

από αισθητήρα καρδιακών παλμών, έξυπνη καρέκλα με αισθητήρες πίεσης και το ίδιο το παιχνίδι. Για τις ανάγκες της πειραματικής διαδικασίας οι συμμετέχοντες χωρίζονται σε τρεις ομάδες, όπως φαίνεται στην εικόνα 4. Κατά τη διάρκεια της πειραματικής διαδικασίας πραγματοποιούνται βελτιώσεις στο ECT με βάση απαντήσεις στο ερωτηματολόγιο εμπειρίας χρήστη και συστάσεις που δίνονται από την ομάδα Α και την ομάδα ελέγχου. Η επίπτωση των αλλαγών στην εμπειρία χρήστη αποτυπώνεται από την ομάδα συμμετεχόντων Β που αλληλεπιδρούν με την ανανεωμένη έκδοση του παιχνιδιού. Τέλος, η αντιληπτή από χρήστη προσήλωση χρησιμοποιείται ως προσέγγιση της βασικής αλήθειας. Μετά την ολοκλήρωση της αλληλεπίδρασης με το ECT οι συμμετέχοντες δηλώνουν τα επίπεδα προσήλωσής τους με συνεχή τρόπο καθώς παρακολουθούν καταγραφή της οθόνης παιχνιδιού και χρησιμοποιώντας κατάλληλο εργαλείο [183]. Η σχεδίαση της πειραματικής διαδικασίας παρουσιάζεται στην εικόνα 4. Μια ομάδα έξι επιπλέον συμμετεχόντων επιστρατεύεται για τη βαθμονόμηση των αισθητήρων της έξυπνης καρέκλας. Αναπτύσσεται κατάλληλη εφαρμογή για τον υπολογισμό των κατωφλιών ενεργοποίησης των αισθητήρων πίεσης της έξυπνης καρέκλας με στόχο την αναγνώριση συγκεκριμένων καθιστικών θέσεων που παρατηρούνται κατά τη διάρκεια της πειραματικής διαδικασίας.

Η ανάλυση για την αναγνώριση των επιπέδων προσήλωσης του χρήστη σε πραγματικό χρόνο πραγματοποιείται σε δύο επίπεδα. Αρχικά διερευνώνται συσχετίσεις μεταξύ καθιστικών στάσεων σώματος που αναγνωρίζονται μέσω της έξυπνης καρέκλας και τα επίπεδα της αντιληπτής από το χρήστη προσήλωσης. Σε δεύτερο επίπεδο, αξιολογείται η προβλεπτική ικανότητα χαρακτηριστικών που εξάγονται από το πλήθος ετερογενών δεδομένων που συλλέγονται προς χαρακτηριστικά που προκύπτουν από την αντιληπτή από το χρήστη προσήλωση. Για την ανάλυση αυτή κατασκευάζονται κατάλληλα παράθυρα παρακολούθησης τα οποία αντιπροσωπεύουν διαφορετικούς τύπους εμπειρίας παιχνιδιού στην έκδοση του ECT που χρησιμοποιήθηκε και γεγονότα μέσα στο παιχνίδι που αναμένεται να επηρεάσουν τα επίπεδα προσήλωσης χρήστη. Το σύνολο των διερευνώμενων χαρακτηριστικών και η καθιστικές θέσεις υπολογίζονται για τα παράθυρα παρακολούθησης που δημιουργούνται. Η προβλεπτική ικανότητα χαρακτηριστικών υπολογίζεται μέσω δυαδικών συντελεστών συσχέτισης με βάση τις μεταβολές τους ανάμεσα σε συναπτά παράθυρα παρακολούθησης. Αξιοποιείται μεθοδολογία ψηφοφορίας για την εξαγωγή πολυτροπικού χαρακτηριστικού. Η ανάλυση πραγματοποιείται με γραμμικό τρόπο χρησιμοποιώντας στατιστικά εργαλεία, όπως συντελεστές συσχέτισης και Student’s t-tests.



Εικόνα 5: Πειραματική διαδικασία για αναγνώριση επιπέδων προσήλωσης σε πραγματικό χρόνο.

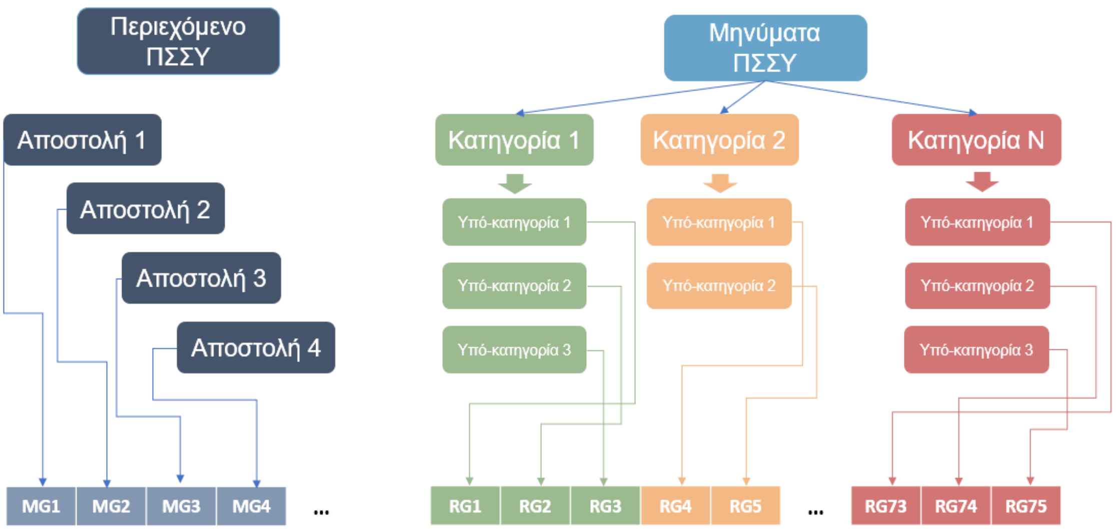
Η μεθοδολογία για τη μελέτη της δυνατότητας αναγνώρισης των επιπέδων προσήλωσης σε πραγματικό χρόνο παρουσιάζεται στην Εικόνα 5. Τα αποτελέσματα της μελέτης δείχνουν ότι τόσο το σοβαρό παιχνίδι όσο και η παρέμβαση ελέγχου ενισχύουν τις διατροφικές γνώσεις του χρήστη (p-value = 0,002, 0,025 αντίστοιχα). Η σύγκριση μεταξύ των δύο ομάδων δεν παρουσιάζει στατιστικά σημαντικές διαφορές ανάμεσα στις δύο παρεμβάσεις (p-value = 0,25). Αυξημένα επίπεδα ικανότητας, εμβάπτισης, ροής και θετικού συναισθήματος δηλώνονται στο ερωτηματολόγιο εμπειρίας παιχνιδιού, γεγονός που καταδεικνύει την ελκυστικότητα του ΠΣΣΥ. Τέλος, τα αποτελέσματα καταδεικνύουν σημαντικές συσχετίσεις και προβλεπτική ικανότητα προς την αντίληψη των επιπέδων προσήλωσης των συμμετεχόντων από όλες τις πηγές δεδομένων που ερευνώνται. Ο πολυτροπικός συνδυασμός χαρακτηριστικών που διερευνάται εμφανίζει υπεροχή ως προς την προβλεπτική του ικανότητα έναντι των υπόλοιπων μονοτροπικών χαρακτηριστικών. Τα αποτελέσματα αυτά συνηγορούν υπέρ της σκοπιμότητας της αναγνώρισης σε πραγματικό χρόνο της προσήλωσης σε προσαρμοστικά ΠΣΣΥ χρησιμοποιώντας την παρουσιαζόμενη προσέγγιση.

5. Αυτοματοποιημένη παραγωγή περιεχομένου σε παιχνίδια σοβαρού σκοπού για την υγεία

Για την αξιολόγηση της δυνατότητας παραγωγής εξατομικευμένου περιεχομένου σε ΠΣΣΥ με αυτόματο τρόπο σχεδιάζεται και υλοποιείται κατάλληλη πειραματική διαδικασία με εκδόσεις του WuF που ενσωματώνουν τη βασισμένη σε ΓΑ μεθοδολογία ΑΠΠ. Στο πλαίσιο της διαδικασίας αυτής διερευνάται η επίπτωση του παραγόμενου περιεχομένου στην εμπειρία χρήστη. Για το σκοπό αυτό αναπτύσσονται τρεις εκδόσεις (Α, Β και Γ) του WuF στις οποίες η παραγωγή περιεχομένου πραγματοποιείται με διαφορετικές παραμέτρους και στοχεύει στην πρόκληση διαφορετικής εμπειρίας παιχνιδιού. Οι εκδόσεις Α και Β ενσωματώνουν τη μεθοδολογία που βασίζεται ΓΑ, σε συνδυασμό με το μηχανισμό για αυτοματοποιημένη προσαρμογή δυσκολίας παιχνιδιού. Η έκδοση Α σχεδιάζεται για να παρέχει εμπειρία με υψηλά επίπεδα δυσκολίας και να δημιουργεί αίσθηση απογοήτευσης στο χρήστη. Η έκδοση Β αντίθετα, προσαρμόζει ομαλά τη δυσκολία τη δυσκολία του παιχνιδιού και ενισχύει την κατάσταση ροής του χρήστη. Η έκδοση Γ δεν χαρακτηρίζεται από δυνατότητα αυτόματης προσαρμογής του περιεχομένου της ανάλογα με την επίδοση του χρήστη, παρά χρησιμοποιεί απλό σύστημα κανόνων για την παραγωγή περιεχομένου και μεταβολή της δυσκολίας της. Οι τρεις εκδόσεις δεν παρουσιάζουν καμία οπτική διαφορά προς το χρήστη και η αξιοποίησή τους στοχεύει στη διερεύνηση της επίπτωσης που θα έχουν στην εμπειρία του.

Η προτεινόμενη μεθοδολογία για ΑΠΠ ενσωματώνεται στις εκδόσεις Α και Β του WuF και δημιουργεί τους αντιπάλους του χρήστη στους αγώνες επιχειρηματολογίας που προσομοιώνονται στο παιχνίδι. Κάθε τέτοιος αντίπαλος χαρακτηρίζεται από συνήθειες και πεποιθήσεις που επηρεάζουν την εμφάνιση και την έκβαση της ΑΥΑ. Τα χαρακτηριστικά αυτά σχηματίζουν το βιογραφικό του αντιπάλου, που παρέχεται στον παίκτη, και τις κάρτες που θα πρέπει να αντιμετωπίσει στο παιχνίδι. Για την ενσωμάτωσή τους στο ΓΑ τα χαρακτηριστικά χωρίζονται σε κατηγορίες, κάθε μία εκ των οποίων αντιπροσωπεύεται από ένα γονίδιο στο χρωμόσωμα του ΓΑ. Τα βάρη των γονιδίων εκπαιδεύονται με διαφορετικό τρόπο στις εκδόσεις Α και Β, ανάλογα με την ικανότητα του χρήστη να αντιμετωπίζει σωστά την εκάστοτε γνωσιακή κατηγορία στο παιχνίδι. Συναρτήσεις καταλληλότητας σχηματίζονται για να επιλέγουν τα χρωμοσώματα που θα σχηματίσουν τον επόμενο πληθυσμό του ΓΑ και το χρωμόσωμα που θα δημιουργήσει τον αντίπαλο για τον επόμενο αγώνα επιχειρηματολογίας. Η έκδοση Γ του παιχνιδιού χρησιμοποιείται ως βάση σύγκρισης για τις εκδόσεις Α και Β, με τους αντιπάλους της να σχηματίζονται με τυχαίο τρόπο. Ο μηχανισμός προσαρμογής της δυσκολίας του παιχνιδιού χρησιμοποιεί ένα σκορ για κάθε χρήστη το οποίο μεταβάλλει ανάλογα με την επίδοσή του. Το σκορ αυτό επηρεάζει το πλήθος των χαρακτηριστικών του αντιπάλου και κατ' επέκταση τη δυσκολία του παιχνιδιού.

Στο πλαίσιο της πειραματικής διαδικασίας, 42 συμμετέχοντες επιστρατεύονται και χωρίζονται σε δύο ομάδες για να αλληλεπιδράσουν με δύο εκδόσεις του ΠΣΣΥ, χωρίς προηγούμενη γνώση των διαφορών μεταξύ τους. Η πρώτη ομάδα αλληλεπιδρά με τις εκδόσεις Α και Γ, ενώ η δεύτερη με τις εκδόσεις Β και Γ, σε τυχαία σειρά, ενώ οι συνεδρίες λαμβάνουν χώρα σε δύο συναπτές ημέρες. Μετά από κάθε συνεδρία, εφαρμόζεται το ερωτηματολόγιο εμπειρίας παιχνιδιού. Τα αποτελέσματα των ερωτηματολογίων αποκαλύπτουν στατιστικά σημαντικές διαφορές σχετικά με την αντίληψη των χρηστών όσον αφορά την αίσθηση ικανότητας, την πρόκληση και την αρνητική εμπειρία για τις εκδόσεις Α και Γ αντίστοιχα, και την αίσθηση ικανότητας και την αρνητική εμπειρία για τις εκδόσεις Β και Γ αντίστοιχα. Η έκδοση Β πετυχαίνει καλύτερες βαθμολογίες στην εμπειρία χρήστη από την έκδοση Γ ενώ η έκδοση Α χειρότερες βαθμολογίες στην εμπειρία χρήστη από την έκδοση Γ. Από τα αποτελέσματα γίνεται σαφές ότι η προτεινόμενη μεθοδολογία ΑΠΠ μπορεί να παραμετροποιηθεί με τρόπο που επηρεάζεται σημαντικά η εμπειρία του χρήστη προς την επιθυμητή κατεύθυνση. Τέλος, οι συμμετέχοντες δήλωσαν ότι δεν μπόρεσαν να αντιληφθούν διαφορές ανάμεσα στις εκδόσεις.



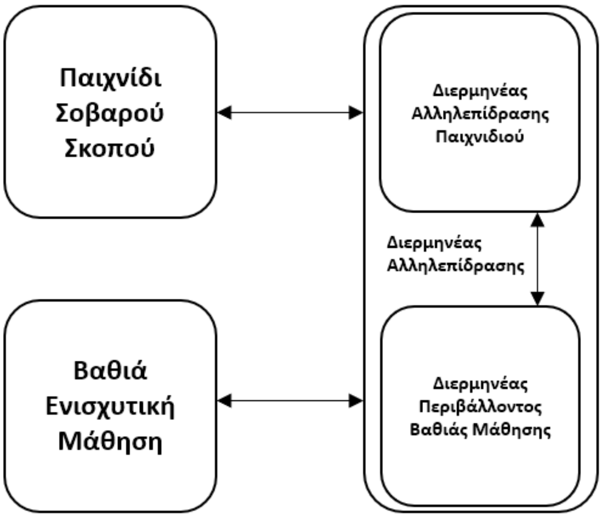
Εικόνα 6: Διαδικασία σχηματισμού του γονιδίου από το περιεχόμενο του παιχνιδιού σοβαρού σκοπού για την υγεία και τα μηνύματα προς τους χρήστες.

Στη συνέχεια, η ικανότητα της προτεινόμενης μεθοδολογίας ΑΠΠ να αξιοποιεί δεδομένα με κλινική αξία που συλλέγονται από αισθητήρες για την επίτευξη εξατομίκευσης σε ΠΣΣΥ διερευνάται σε προ-πιλοτική μελέτη. Η μελέτη περιλαμβάνει 20 παιδιά που πάσχουν από παχυσαρκία και αλληλεπιδρούν με πλατφόρμα παρέμβασης που περιλαμβάνει ΠΣΣΥ για μια περίοδο δώδεκα εβδομάδων. Ο ΓΑ ενσωματώνεται στην πλατφόρμα και παρέχει εξατομικευμένο περιεχόμενο με τη μορφή προσαρμοσμένων μηνυμάτων και αποστολών του ΠΣΣΥ, λαμβάνοντας υπόψη δεδομένα που συλλέγονται από αισθητήρες και την αλληλεπίδραση των χρηστών με την πλατφόρμα. Για τη διευκόλυνση της λειτουργίας του ΓΑ το περιεχόμενο παιχνιδιού και τα μηνύματα χωρίζονται σε κατηγορίες και σχηματίζουν τα γονίδια του χρωμοσώματος όπως φαίνεται στην Εικόνα 6. Τα βάρη των γονιδίων που αφορούν το περιεχόμενο του παιχνιδιού εκπαιδεύονται με βάση την επίδοση του χρήστη σε αυτό. Τα βάρη των γονιδίων που αφορούν τα μηνύματα που προβάλλονται εκπαιδεύονται με βάση δεδομένα αλληλεπίδρασης με την πλατφόρμα και αισθητήρες που παρακολουθούν σωματική δραστηριότητα και ποιότητα ύπνου. Η εκπαίδευση των βαρών πραγματοποιείται ανά ημέρα και ο ΓΑ παράγει νέες γενιές χρωμοσωμάτων σε εβδομαδιαία βάση. Το χρωμόσωμα που επιλέγεται από τη συνάρτηση καταλληλότητας καθορίζει το περιεχόμενο της πλατφόρμας για την εβδομάδα που ακολουθεί.

Η ανάλυση που ακολουθεί αξιολογεί την ικανότητα της προτεινόμενης ΑΠΠ να παράγει περιεχόμενο σχετικό με τα δεδομένα που συλλέγονται από τους αισθητήρας, όπως επίσης και η αντίληψη της σχετικότητας του παραγόμενου περιεχόμενου από τους χρήστες μέσα από ερωτηματολόγια Τα αποτελέσματα συνηγορούν προς την αποτελεσματικότητα της ΑΠΠ όσον αφορά τη δημιουργία εξατομικευμένου περιεχομένου με επαρκή συνάφεια και χρησιμότητα. Αποκαλύπτεται στατιστικά σημαντική συσχέτιση μεταξύ του περιεχομένου που παράγεται από το ΓΑ και των δεδομένων που συλλέγονται από τους αισθητήρες. Μείωση στους δείκτες που αφορούν συχνή σωματική δραστηριότητα και καλή ποιότητα ύπνου συνδέονται με την εμφάνιση μηνυμάτων που προωθούν συναφείς αλλαγές στη συμπεριφορά. Το γεγονός αυτό υπογραμμίζει τις δυνατότητες της προτεινόμενης μεθοδολογίας ΑΠΠ στον εντοπισμό και την αντιμετώπιση των εξατομικευμένων αναγκών του χρήστη. Παράλληλα, με βάση τις απαντήσεις σε σχετικές ερωτήσεις αποτυπώνεται η αντίληψη της καταλληλότητας του παραγόμενου περιεχομένου από τους χρήστες.

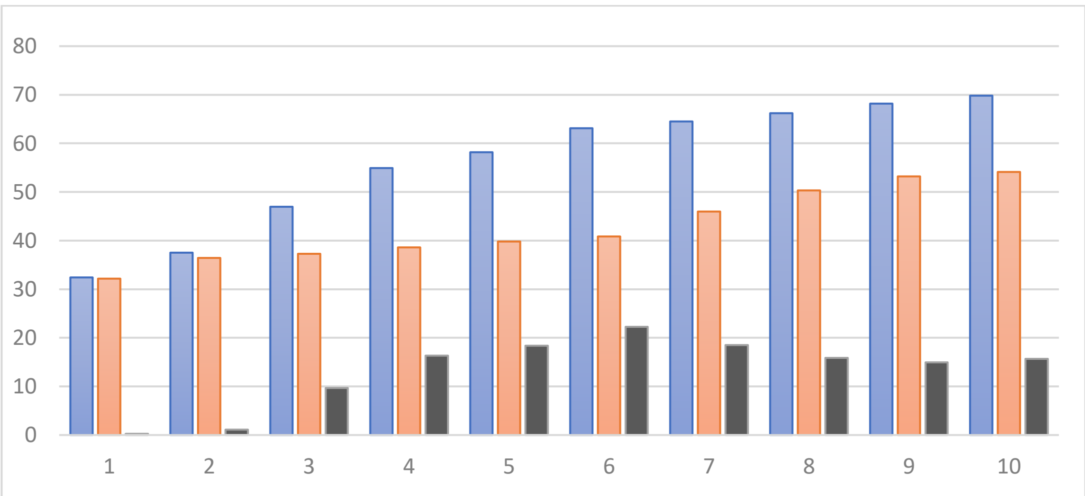
6. Αυτοματοποιημένη δοκιμή παιχνιδιού σοβαρού σκοπού για την υγεία

Η προτεινόμενη μεθοδολογία για ΑΠΠ που βασίζεται σε ΓΑ αξιολογείται με τη χρήση πρακτόρων βαθιάς ενισχυτικής μάθησης για αυτοματοποιημένη δοκιμή παιχνιδιών. Τεχνικές αυτοματοποιημένης δοκιμής παιχνιδιών χρησιμοποιούνται πλέον κατά κόρον κατά τη διάρκεια της σχεδίασης και ανάπτυξης ψηφιακών παιχνιδιών. Οι τεχνικές αυτές παρουσιάζουν σημαντικά πλεονεκτήματα λόγω της δυνατότητας τους για ταχύτατη δοκιμή παιχνιδιών με επαναληπτικό και οικονομικό τρόπο. Για την αξιολόγηση της προτεινόμενης μεθοδολογίας πράκτορες βαθιάς ενισχυτικής μάθησης εκπαιδεύονται να παίζουν αυτόνομα το WuF. Με τον τρόπο αυτό γίνεται δυνατή η πλήρως αυτοματοποιημένη αξιολόγηση της ενσωματωμένης μεθοδολογίας ΑΠΠ, χωρίς την ανάγκη επιστράτευσης ανθρώπινων συμμετεχόντων. Για να επιτευχθεί αυτοματοποίηση και αυτονομία στη δοκιμή, σχεδιάζεται και υλοποιείται παραλλαγή του WuF η οποία στη συνέχεια γενικεύεται και μεταφράζεται εν μέρει σε περιγραφική γλώσσα παιχνιδιού [228]. Μέσω της διαδικασίας αυτής δημιουργείται κατάλληλο περιβάλλον που επιτρέπει στους πράκτορες βαθιάς ενισχυτικής μάθησης να αλληλεπιδρούν με το ΠΣΣΥ μέσω ερωτημάτων. Η περιγραφική γλώσσα παιχνιδιού επιτρέπει την αναπαράσταση άγνωστων παιχνιδιών παρέχοντας πληροφορίες σχετικά με τους κανόνες τους. Για την αυτόματη δοκιμή χρησιμοποιούνται δύο εκδόσεις του WuF, μία που ενσωματώνει τη μεθοδολογία ΑΠΠ και μία που παράγει περιεχόμενο με απλό σύστημα κανόνων. Στόχος της διαδικασίας αποτελεί η διερεύνηση διαφορών στην εκπαίδευση και επίδοση των πρακτόρων όταν εκτίθενται στο περιεχόμενο των διαφορετικών εκδόσεων. Το πλαίσιο αυτόματης δοκιμής φαίνεται στην εικόνα 7.



Εικόνα 7: Πλαίσιο για αυτόματη δοκιμή παιχνιδιού σοβαρού σκοπού για την υγεία.

Για την αυτόματη δοκιμή των δύο εκδόσεων του παιχνιδιού εκπαιδεύονται είκοσι ζευγάρια ξεχωριστών πρακτόρων Proximal Policy Optimization (PPO) [230] σε κάθε έκδοση του ΠΣΣΥ. Η ανάλυση της απόκρισης των πρακτόρων στις διαφορετικές εκδόσεις διεξάγεται υπολογίζοντας το μέσο ποσοστό κέρδους κάθε πράκτορα και τη συνολική ανταμοιβή που του απονέμεται κάθε 100.000 βήματα εκπαίδευσης. Τα αποτελέσματα από τη διαδικασία εκπαίδευσης καταδεικνύουν μια συνολική υπεροχή των πρακτόρων PPO που εκπαιδεύονται όταν εκτίθενται σε περιεχόμενο που παράγεται από την προτεινόμενη μεθοδολογία ΑΠΠ, σε σχέση με τους πράκτορες που εκπαιδεύονται στην απλή έκδοση του WuF. Το μέσο ποσοστό νίκης βρέθηκε να είναι σημαντικά υψηλότερο κατά μέσο όρο για όλα τα σενάρια εκπαίδευσης μετά από 300.000 εποχές εκπαίδευσης, όπως φαίνεται στην Εικόνα 8. Τρεις κατηγορίες σεναρίων ορίστηκαν μετά την ανάλυση σύμφωνα με την απόδοση εντός του παιχνιδιού. Οι πράκτορες που ανήκουν στην πρώτη κατηγορία παρουσιάζουν σημαντικά καλύτερη απόδοση από την αρχή της διαδικασίας εκπαίδευσης. Οι πράκτορες της δεύτερης κατηγορίας έχουν ελαφρώς καλύτερη απόδοση, η οποία εμφανίζεται αργότερα στη διαδικασία εκπαίδευσης και οι πράκτορες της τρίτης κατηγορίας έχουν παρόμοια επίδοση και στις δύο εκδόσεις του WuF. Δεν παρατηρούνται περιπτώσεις όπου οι πράκτορες που εκτέθηκαν στο περιεχόμενο που παράγει ο ΓΑ έχουν χαμηλότερη απόδοση σε σύγκριση με τους πράκτορες που εκπαιδεύονται στην απλή έκδοση του WuF. Τα αποτελέσματα αυτά αποτελούν ισχυρή ένδειξη προς την ικανότητα της προτεινόμενης μεθοδολογίας ΑΠΠ να προσαρμόζει το περιεχόμενο ανάλογα με την ανάγκη του παίκτη.



Εικόνα 8: Μέσο ποσοστό νικών ανά 100.000 βήματα εκπαίδευσης.

7. Σύνοψη και μελλοντική έρευνα

Στην παρούσα Διδακτορική Διατριβή παρουσιάζεται πρωτότυπο εννοιολογικό πλαίσιο για αναγνώριση των επιπέδων προσήλωσης κατά τη διάρκεια αλληλεπίδρασης με έξυπνο ΠΣΣΥ σε πραγματικό χρόνο. Το πλαίσιο αξιοποιεί αισθητήρες, δεδομένα από το παιχνίδι και δεδομένα υγείας για να δημιουργήσει μια εξατομικευμένη παρέμβαση με στόχο την ενίσχυση των επιπέδων προσήλωσης και την στόχευση των αναγκών υγείας του χρήστη. Τα συμπεράσματα που παρουσιάζονται στην παρούσα Διδακτορική Διατριβή υποστηρίζουν τη δυνατότητα υλοποίησης του προτεινόμενου εννοιολογικού πλαισίου. Τα αποτελέσματα των πειραματικών διαδικασιών συνηγορούν στη δημιουργία ενός κλειστού βρόχου ανατροφοδότησης της προσήλωσης παίκτη σε δυναμικά μεταβαλλόμενα ΠΣΣΥ, αξιοποιώντας δεδομένα αισθητήρων και αλληλεπίδρασης. Η μελλοντική έρευνα περιλαμβάνει την περαιτέρω διερεύνηση της ικανότητας της προτεινόμενης μεθοδολογίας ΑΠΠ για γενίκευση, όπως επίσης και την αξιολόγηση της αναγνώρισης των επιπέδων προσήλωσης με μη γραμμικό τρόπο. Τέλος, το προτεινόμενο εννοιολογικό πλαίσιο μπορεί να αξιολογηθεί στο σύνολό του με την ενσωμάτωσή του σε διαφορετικά είδη ΠΣΣΥ.

Acknowledgements

With the completion of my Doctoral Thesis, I would like to extend my deepest gratitude to my supervisor, Prof. Konstantina S. Nikita, for her continuous support and valuable mentorship throughout these past years. From the moment she entrusted me with the assignment of this research topic to this date, through tireless effort, her supervision and advice have motivated and empowered my research work and helped me grow and evolve both professionally and personally.

I would also like to offer my gratitude to Prof. Dimitrios - Dionysios Koutsouris, Prof. Andreas-Georgios Stafylopatis and Prof. Giorgos Stamou for the guidance they provided me with as part of my dissertation committee, as well as Prof. George Matsopoulos, Prof. Panagiotis Tsanakas, Asst. Prof. Athanasios Voulodimos and Assoc. Prof. Spyretta Golemati for the time they devoted and valuable insights as part of my examination committee.

Furthermore, I would like to extend my gratitude to Dr. Konstantia Zarkogianni, for her constant instruction and valuable guidance throughout the entirety of my studies. I express my sincere thanks to my colleague, Dr. Maria Athanasiou, for our impeccable collaboration and all the assistance and support she has provided through these years. I'm also extremely grateful to Eleutherios Kalafatis for his assistance, his support, and our wonderful cooperation. I would also like to thank all the members of the Biomedical Simulations and Imaging Laboratory of National Technical University of Athens for their support and assistance.

This endeavour would not have been possible without the constant presence and support from my wife, my father, and all my family and friends. Finally, I would like to dedicate my thesis to the loving memory of my mother.

Table of contents

Abstract 6
Περίληψη 8
Εκτεταμένη περίληψη 9
Acknowledgements 21
List of tables 24
List of figures 25
List of abbreviations 27
Glossary 28
1. Adaptive serious games for health 29
1.1 Introduction 29
1.2 Serious games 29
1.2.1 Definition and terminology 30
1.2.2 What is a serious game? 31
1.2.3 Design and validation considerations 33
1.2.4 Fields of interest 34
1.3 Serious games for health 36
1.4 Adaptivity in serious games for health 40
1.4.1 Affective player modeling 41
1.4.2 Procedural content generation 44
1.5 Thesis structure 45
2. Conceptual framework for adaptivity in serious games for health 47
2.1 A proposal for adaptivity in SGs for health 48
2.1.1 Serious game layer 49
2.1.2 Procedural content generation layer 50
2.1.3 Engagement recognition layer 50
2.1.4 Health related data layer 50
2.2 A genetic algorithm approach for procedural content generation 51
3. Design and development of adaptive serious games for health 55
3.1 A serious game for food and nutrition literacy 55
3.2 A serious game for obstructive sleep apnea 60
3.3 A serious game for type 1 diabetes and obesity 66
4. Modeling engagement in serious games for health 68
4.1 Experimental process 69
4.1.1 Serious game version 70
4.1.2 Measurement of the intervention’s efficiency 71

4.1.3 Annotation of perceived engagement ... 73
4.1.4 In-game metrics ... 73
4.1.5 Sensors ... 74
4.1.6 Data analysis ... 76
4.2 Results and discussion ... 79
4.2.1 Measurement of the intervention's efficiency ... 79
4.2.2 Game experience ... 80
4.2.3 Real time recognition of engagement ... 82
5. Procedural content generation in serious games for health ... 90
5.1 Incorporation of the genetic algorithm technique in a SG for obstructive sleep apnea ... 91
5.1.1 Experimental process ... 93
5.1.2 Results and discussion ... 94
5.2 Incorporation of the genetic algorithm in a SG for type 1 diabetes and obesity ... 95
5.2.1 Experimental process ... 98
5.2.2 Results and discussion ... 98
6. Automated testing in serious games for health ... 101
6.1 Deep reinforcement learning in games ... 102
6.2 Automated testing of a serious game for health ... 103
6.2.1 Conceptual framework for testing ... 103
6.2.2 Serious game ... 104
6.2.3 Interaction interpreter ... 106
6.2.4 Deep reinforcement learning ... 106
6.2.5 Analysis ... 107
6.3 Results and discussion ... 107
7. Conclusions and future work ... 113
7.1 Adaptive serious games for health ... 114
7.2 Player modeling in serious games for health ... 117
7.3 Future work ... 118
References ... 120

List of tables

Table 1: Search strategy........ 38
Table 2: Summary of extracted features. 70
Table 3: Participants of the experimental process 72
Table 4: Set of sitting postures and their activation patterns. 75
Table 5: Number of pairwise comparisons included in different types of observation frames. 78
Table 6: Results for GEQ for the two versions of Express Cooking Train........ 80
Table 7: Information about serious games applying GEQ identified in the literature........ 81
Table 8: Correlation of annotation features and features extracted from posture sensors, heart rate sensor and in-game metrics for 10s reactive frames. Significant values (P<0.05) are depicted in bold. Highly significant values (P<0.01) are denoted by (*). 87
Table 9: Correlation of annotation features and features extracted from posture sensors, heart rate sensor and in-game metrics for 30s reactive frames. Significant values (P<0.05) are depicted in bold. Highly significant values (P<0.01) are denoted by (*). 87
Table 10: Correlation of annotation features and features extracted from posture sensors, heart rate sensor and in-game metrics for continuous frames. Significant values (P<0.05) are depicted in bold. Highly significant values (P<0.01) are denoted by (*). 88
Table 11: An example chromosome featuring seven non-player character attributes linked to obstructive sleep apnea........ 91
Table 12: Results in terms of game experience for groups 1 and 2. 94
Table 13: Post-intervention questions for user acceptance........ 99
Table 14: Pearson's correlation between delivered messages and sensor data. 99

List of figures

Figure 1-1: What is a serious game. 32
Figure 1-2: Serious game application domains and design elements combined with further concepts, technologies, and disciplines. 35
Figure 1-3 : Classification of serious games for health according to player and objective. 38
Figure 1-4: Number of publications of reviews about serious games for health per year. 39
Figure 1-5: Classification of identified review articles. 40
Figure 2-1: Conceptual framework for adaptivity in serious games for health 49
Figure 2-2: Framework based on a genetic algorithm for procedural content generation in serious games for health. 52
Figure 2-3: Translation of serious game content into the genetic algorithm's chromosome structure. 53
Figure 3-1: Express cooking train conceptual framework. 56
Figure 3-2: Express cooking train example recipe (a) and sample ingredients (b) 57
Figure 3-3: Express cooking train virtual kitchen environment. 57
Figure 3-4: Express cooking train. 58
Figure 3-5: Reward, progress and achievement screen of Express Cooking Train. 59
Figure 3-6: Conceptual framework for the design of A Marcus Tale [164]. 61
Figure 3-7: Conceptual framework for Wake Up for the Future. 62
Figure 3-8: Scope of Wake Up for the Future. 62
Figure 3-9: Cards in Wake Up for the Future. 63
Figure 3-10: Open world of Wake Up for the Future. 63
Figure 3-11: Non-player character biographies in Wake Up for the Future. 64
Figure 4-1: Conceptual framework for real-time recognition of engagement. 69
Figure 4-2: Setup of the experimental process. 70
Figure 4-3: Screenshots from the three phases of Mission-1 and the monster: (a) Tutorial; (b) Gameplay; (c) Review; (d Monster 71
Figure 4-4: Design of the evaluation study for educational value and game experience of Express Cooking Train. 72
Figure 4-5: (a) The interface of the annotation tool along with the playthrough recording as shown in 73
Figure 4-6: (a) Smart chair with initial arrangement of 12 pressure sensors. PulseSensor (placed in the user's ear lobe) is marked by the green circle.; (b) Microcontroller setup for acquisition of data from the smart chair. 74
Figure 4-7: Accuracy results towards the determination of sensor activation threshold for posture identification. 76
Figure 4-8: Continuous and reactive frames 77
Figure 4-9: Flowchart for data analysis. 78
Figure 4-10: Average scores in Knowledge Questionnaire for Game Group A and Control Group. 79
Figure 4-11: GEQ Core module dimensions for version A and B of Express Cooking Train, along with comparison with five serious games identified in the literature. 81
Figure 4-12: GEQ Post-game module dimensions for versions A and B of "Express Cooking Train", along with comparison with three serious games identified in the literature. 81
Figure 4-13: GEQ dimensions with significant differences among Clusters 1 (low interaction), 2 (intermediate interaction) and 3 (high interaction). 82
Figure 4-14: Distribution of postures identified for Tutorial, Gameplay, Review, and Mission-2. 83
Figure 4-15: Contingency tables for continuous frames: (a) Tutorial; (b) Gameplay; (c) Review; (d) Mission-2. 83

Figure 4-16: Percentage of posture transitions identified in each frame. 84
Figure 4-17: Whisker boxplots for Tutorial, Gameplay, Review and Mission-2, for (a) mean value of perceived engagement, (b) area of annotation trace, (c) amplitude of engagement, (d) average gradient of engagement. 84
Figure 4-18: Distribution of postures identified during reactive frames. 85
Figure 4-19: Contingency tables for reactive frames: (a) Frame prior to event (10 s); (b) Frame after event (10 s); (c) Frame prior to event (30 s); (d) Frame after event (30 s) 85
Figure 4-20: Percentage of posture transitions identified in each frame. 86
Figure 4-21: Whisker boxplots for reactive frames, for (a) mean value of perceived engagement, (b) area of annotation trace, (c) amplitude of engagement, (d) average gradient of engagement. 86
Figure 5-1: Conceptual framework for procedural content generation. 91
Figure 5-2: Chromosome formation, with genes controlling missions (MG) and genes controlling messages (RG). 96
Figure 5-3: Chromosome crossover. (a) Two random points, cM and cR, split the chromosomes. (b) Two new gene sequences for MG and two for RG are generated. 97
Figure 6-1: Conceptual framework for automated testing in serious games. 104
Figure 6-2 : Screenshot from the core game aspect of Wake Up for the Future. 105
Figure 6-3 : Screenshot from the meta game aspect of Wake Up for the Future. 105
Figure 6-4: Overall win rate of an example training in category 1. 108
Figure 6-5: Cumulative reward of an example training in category 1. 108
Figure 6-6: Overall win rate of an example training in category 2. 109
Figure 6-7: Cumulative reward of an example training in category 2. 109
Figure 6-8: Overall win rate of an example training in category 3. 110
Figure 6-9: Cumulative reward of an example training in category 3. 110
Figure 6-10: Average performance for all training scenarios for the adaptive version of the serious game (A) and the simple version of the serious game (B) every 10.000 training steps. 111
Figure 6-11: Overall win rate for all training scenarios. 111
Figure 7-1 : The proposed conceptual framework for adaptivity in serious games for health. 113
Figure 7-2: Game mechanics facilitating the incorporation of procedural content generation techniques in the employed serious games: (a) ingredient selection in Express Cooking Train; (b) card selection in Wake Up for the Future. 114

List of abbreviations

ADHD	Attention Deficit Hyperactivity Disorder
AI	Artificial Intelligence
AMT	A Marcus Tale
ANOVA	Analysis of Variance
BMI	Body Mass Index
DRL	Deep Reinforcement Learning
ECT	Express Cooking Train
EEG	Electroencephalogram
EMG	Electromyogram
FDA	United States Food and Drug Administration
FL	Food Literacy
GA	Genetic Algorithm
GDL	Game Description Language
GEQ	Game Experience Questionnaire
ICT	Information and communications technology
LFF	Losing Fitness Function
MG	Mission Genes
MHealth	Mobile Health
ML	Machine Learning
NL	Nutrition Literacy
NPC	Non-player characters
OSA	Obstructive Sleep Apnea
PC	Personal Computer
PCG	Procedural Content Generation
PPG	Photoplethysmography
PPO	Proximal Policy Optimization
RG	Recommendation Genes
SCT	Social Cognitive Theory
SG	Serious Game
USD	United States Dollar
VR	Virtual Reality
WFF	Winning Fitness Function
WFS	Winning Fitness Score
WuF	Wake Up for the Future

Glossary

Adaptivity	Προσαρμοστικότητα
Affective Computing	Συναισθηματική Υπολογιστική
Artificial Intelligence	Τεχνητή Νοημοσύνη
Body Mass Index	Δείκτης Μάζας Σώματος
Chromosome	Χρωμόσωμα
Deep Learning	Βαθιά Μάθηση
Dynamic Difficulty Adjustment	Δυναμική Προσαρμογή Δυσκολίας
Electroencephalogram	Ηλεκτροεγκεφαλογράφημα
Electromyogram	Ηλεκτρομυογράφημα
Engagement	Προσήλωση
Fitness Function	Συνάρτηση Καταλληλότητας
Game Experience	Εμπειρία Παιχνιδιού
Game Testing	Δοκιμή Παιχνιδιού
Gamification	Παιχνιδοποίηση
Gene	Γονίδιο
Genetic Algorithm	Γενετικός Αλγόριθμος
Machine Learning	Μηχανική Μάθηση
Non-player Character	Χαρακτήρας Παιχνιδιού
Obstructive Sleep Apnea	Αποφρακτική Υπνική Άπνοια
Personalization	Εξατομίκευση
Photoplethysmography	Φωτοπληθυσμογραφία
Player Modeling	Μοντελοποίηση Παίκτη
Procedural Content Generation	Αυτοματοποιημένη Παραγωγή Περιεχομένου
Reinforcement Learning	Ενισχυτική Μάθηση
Serious Game	Παιχνίδι Σοβαρού Σκοπού

1. Adaptive serious games for health

1.1 Introduction

Serious games (SGs) have been a topic of growing interest during the last decade. SGs are considered effective tools that provide means for training skills, education, raising awareness, as well as promoting behavioral change. The field of health is considered one of the most prevalent fields of application for SG interventions. These intervention tools can address several health-related challenges such as raising awareness, training of health professionals, education, rehabilitation, disease monitoring and diagnosis, promotion of behavioral lifestyle changes, and management of mental health [1], [2]. However, despite the observed advancements in the field, the application of SGs in health still encounters various limitations, mostly in terms of the employed design approaches and the systematic validation of their effectiveness [3]. Additionally, research for personalizing persuasive game design to tailored players' needs is still considered nascent, despite the benefits that have been reported [4]. Relevant literature identifies differences in receptivity of employed persuasive strategies in SGs for health, among a variety of user types, indicating that intuitive, one-size-fits-all approaches in SG design are not always effective. This finding, in combination with the ambiguous results reported regarding the efficiency of SGs, provides additional motivation for research towards enhancing SG adaptivity [5]. The importance of delivering personalized content in SGs is also highlighted in a recent review study [6], indicating that personalization in these interventions does not only empower game experience, but also addresses specific user needs linked to the game's serious purpose, like knowledge acquisition or task performance. As a result, adaptivity in SGs for health is expected to enhance the efficiency of the delivered intervention, as well as promote adherence to it. In this manner, SGs for health can be augmented in tools capable of delivering adaptive interventions, which provide personalized content tailored to the players particular health related needs.

The main contribution of the present Doctoral Thesis is to propose and investigate the potential of a conceptual framework for adaptivity in SGs for health, based on a real-time engagement feedback loop, which employs sensors able to collect clinically relevant data. The purpose of this framework is to facilitate the development of adaptive SGs for health, which target personalized player needs by adjusting their content and improving adherence to the intervention. Two SGs for health, aiming at different health outcomes for a variety of health conditions, were designed and developed to support the feasibility of the proposed framework. These intervention tools were designed for adaptivity, incorporating PCG techniques specific to their purpose. In addition, the SGs were developed to be scalable, with the capacity to be generalized to support diagnosis, raising awareness, and disease self-management in various conditions. A third SG, aiming to promote disease self-management in children suffering from obesity and/or type 1 diabetes, was also employed to evaluate the generalization capability of the proposed framework, as well as its capacity to incorporate health-related sensor data in the process of generating adaptive game content. Two experimental procedures were carefully designed and implemented, along with data from a pre-pilot study, to assess the effectiveness of the developed tools, evaluate user acceptance, and collect data from in-game metrics and sensors. The feasibility of real-time identification of perceived engagement during SG playing was investigated, by employing a continuous annotation tool. The PCG technique incorporated in the SGs was also evaluated in terms of validity of the generated content and effect to player experience. Results from the conducted experimental procedures and the analysis of the collected data advocate towards the feasibility of the proposed conceptual framework for adaptivity in SGs for health.

1.2 Serious games

Nowadays SGs constitute an established field of research that gathers continuously increasing interest [7]. Discussion regarding games' potential for purposes besides entertainment can be traced to the work of Plato [8]. Plato argues that by reinforcing behaviours during play at young age, these behaviours would consequently be reinforced in adulthood. In recent decades, the field of SGs has grown and evolved, a development that is often accredited to simulation-based learning that has been applied in military training exercises [8]. SGs now refer almost exclusively to digital games, developed for a variety of operating systems and including virtually all game genres [9]. With the evolution of the gaming industry and the immersion of novel technologies that facilitate game development and lower production costs, SGs are now easier and cheaper to develop. Initial doubts and concerns regarding the "serious" potential of games, voiced by researchers in the beginning of the century [10], have recently subsided and were mostly accredited to the poor quality of the SGs produced at that time [11]. SGs have the potential to provide communication and interaction with peers and mentors, be flexible, portable, and scalable, while maintaining a high quality of efficiency in terms of the intervention they are providing [12]. After decades of growth, SGs are now widely considered a valuable tool in many fields of application.

Besides the increased focus on SGs in the field of research, SGs have also displayed a significant presence in the market, gradually forming their own industry in the past years. According to the summary of a recent report by Allied Market Research [13], the value of the global SG market was estimated at 5.94 billion USD in 2020 and is expected to reach 32.72 billion USD by 2030. Another report, by Mordor Intelligence [14], estimated the market's worth in 2020 at 6.29 billion USD, expected to reach 25.54 billion USD by 2026. The COVID-19 pandemic, along with the rising demand for remote learning it brought about, has been identified as one of the main reasons driving this rapid growth by both reports. Various commercial SGs are now available to experts and the public, targeting multiple fields of interest. A large number of companies have shifted focus to accommodate the needs for the design and development of SGs, by employing professionals from a variety of backgrounds. Mobile gaming appears to have penetrated the SG industry, in accordance with the trends recorded in the entertainment game industry [15], while novel technologies such as procedural content generation (PCG), player modeling, virtual and augmented reality are being increasingly employed in commercial SGs [16].

1.2.1 Definition and terminology

SGs have been in the research spotlight over the last two decades, however, the first reference to the term SG can be found in a book titled "Serious Games" by Clark Abt [17], which was first published in 1970 [18]. As the field of SGs evolved, multiple definitions regarding SGs have been reported in the literature, with one of the most widely used stating that SGs are digital games with a main purpose other than pure entertainment [19]. Another similar and more recent definition, describes SGs as digital games created with the intention to entertain and to achieve at least one additional goal [20]. The main difference between these two definitions exists in the intention of the game designer, and whether a digital game can constitute a SG if it was not designed to be one. This can be considered a significant difference, as there have been many examples of commercial games that were designed and developed without any "serious" purpose in mind, however resulted in having significant impact in player wellbeing, reduction of depression, and positive emotional regulation in general, as well as training of skills, such as learning of second languages, development of communication skills, improvement of spatial cognition and many other [21]. This sparks a debate that is perhaps outside the scope of the present Doctoral Thesis, regarding what is the crossing line between an entertainment game and a serious and how the purpose of the designer can be safely interpreted. A secondary difference between the two definitions lies in the distinction between entertainment and "serious" purpose, and whether one of them should be characterized as primary or main. This difference is more evident in the definition describing SGs as games that do not have entertainment, enjoyment or fun as their primary purpose [22]. There is an abundance of other definitions of SGs, some of them implying that the interventions can only be conducted in formal settings (e.g., educational classrooms). To this day, despite the advancement of the field, no universally accepted definition of the term SG exists.

In reality some would argue that the field evolution has led to the surfacing of additional terms describing intervention tools close to SGs in relevant literature, such as game-based learning [23], edutainment [24], playful learning [25], games for learning [26], games with a purpose [27]. These terms share common elements and characteristics in the way they are defined, both with SGs, but also amongst each other. However, most of these terms refer mostly to educational applications for games, whereas a SG is a digital game that can be employed for a multitude of purposes besides learning [20]. In addition, many of these terms do not refer solely to digital games, or even games in general. As a result, the term SG is considered to be overarching and is currently the most widely accepted. However, this abundance of terms and definitions has reportedly led to confusion in the field and amongst researchers, with various attempts of clearing the matter found in relevant literature.

Finally, SGs are not to be confused with gamification and applications that incorporate such techniques. Gamification, as a term, made a much later appearance, in 2008, and was not widely used until 2010 [28]. Gamification refers to the use of design elements that are characteristic to games in non-game contexts [28]. As such, gamification refers to the transfer of mechanics typically found in game settings in other applications, mostly digital, whereas a SG satisfies all the necessary conditions to constitute an actual game. Gamification rapidly gathered much attention, and the field is now considered to be among the top technology and software trends [29]. Gamification aims to enhance the applications it encompasses, by increasing user engagement, usability, functionality and producing positive business impact [30], [31]. These goals are similar to those of SGs; however, they are achieved in a different manner and applied in different settings. Despite these differences, gamification and SGs are often considered similar fields, and as such they are studied and researched together, as is evident from relevant literature [30], [32], [33], [34]. In addition, certain attempts to align gamification and SGs have been identified, however, the use of game elements in the two fields has been reported to influence learning through different processes [35].

1.2.2 What is a serious game?

To better understand what constitutes a SG, the first step is to define what a game is. According to the Oxford Learner's Dictionaries, by Oxford University Press, a game is "an activity that you do to have fun, often one that has rules and that you can win or lose" [36]. Comparing this definition with the ones identified in the literature for SGs, the main difference that arises is the fact that SGs also serve another purpose besides entertainment. SGs by definition constitute games, and as such, they are governed by rules and feature winning and losing conditions. In addition, SGs are designed to be fun to play, otherwise they forfeit one of their greater assets, their capacity to induce engagement during interaction.

Games employ various mechanics and elements to increase engagement during play. Engagement is in fact one of the most defining factors in the success of a commercial entertainment game. However, engagement during play is often considered in different ways and holds multiple meanings according to context. In a recent review study that investigates engagement in SGs a three-fold structure to define engagement [37]. These definitions are based on different uses of the term engagement during SG playing, identified during the review process. The first refers to the player "engaging" with the SG. The second to engagement referring to the state of the player as being "engaged", and the third as a property of the SG that aims to be "engaging". According to another study, engagement during play can be modelled in different stages based on player interaction, namely, point of engagement, sustained engagement, disengagement and re-engagement [38]. Despite various attempts, no universal consensus appears to exist regarding the definition of engagement during play [39], but the importance of creating an engaging experience is highlighted across the literature [40].

SGs benefit from this aspect of gaming greatly, as they aim to achieve their "serious" goals while providing an engaging experience. To this end, the entertaining elements that are incorporated in the SGs are considered key factors to their success. SGs strive to be impactful intervention tools, able educate, promote attitude or behaviour change and skill acquisition, while maintaining a balance with their entertainment side [41]. Additionally, SGs employ various other game mechanics that provide the player with active, experiential, problem based tasks and give immediate feedback [42], elements of critical importance for effective learning according to modern theories. In order to better describe the association between games and SGs, based on what defines and separates them, Figure 1-1 [43] employs seven broad categories of components. According to the author, six of them exist in both games and SGs, but one is exclusive to SGs. This categorization refers to all games in general, and of course includes digital game as well.

The first component is the set of rules, gameplay mechanics and systems that govern the game. The second one refers to the challenge that is presented by the game and the reward that is obtained from overcoming them. The third one involves the interaction of the player with the game and all the means present in a game to achieve that. The final two components both refer to the game objective, which refers to something that, through player actions and decisions, they intend to achieve or attain. According to this approach, in entertainment games only an explicit objective is present, entertainment. On the other hand, in SGs the objectives are also implicit, and include for example, skill improvement, knowledge and experience acquisition, recovery and physical improvement. This description is in accordance with most of the definitions of SGs that were presented, however the question remains, whether there exist games that do not present the player with at least one implicit objective. Whatever the case, for a SG to be effective, the presentation of these implicit objectives must be clear and concise. To this end, the design process must be taken into very careful consideration. The fact that this is often not the case, along with a lack of systematic methodologies for the evaluation of the impact of the SGs is still hindering a wider adoption of them as intervention tools [44].

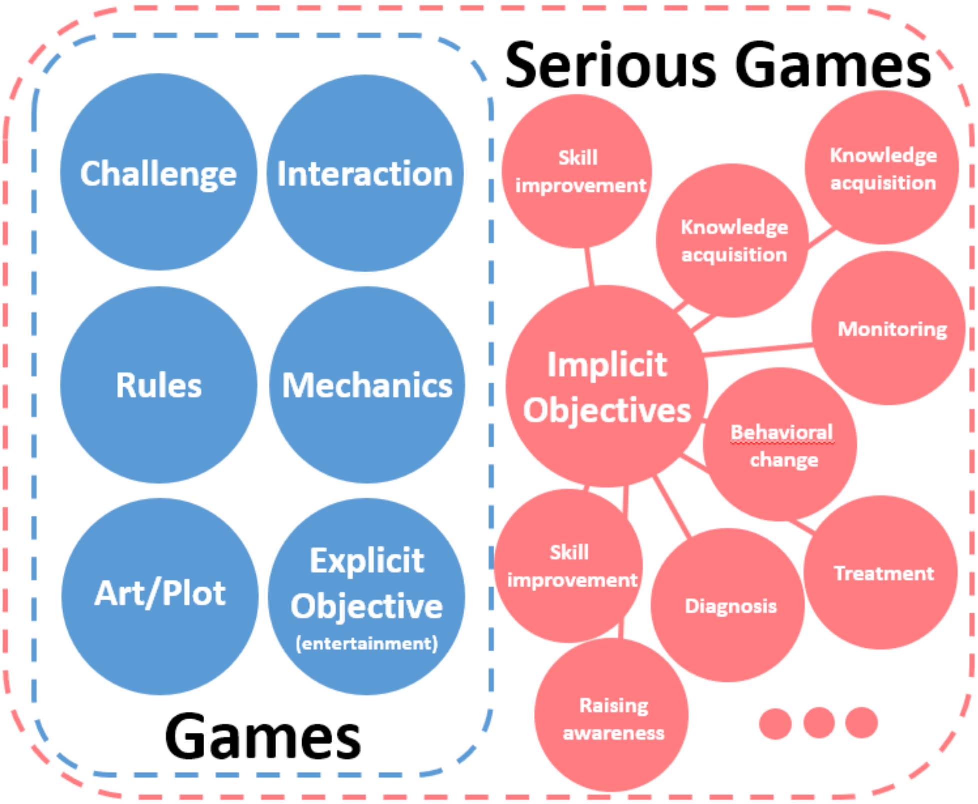


Figure 1-1: What is a serious game.

1.2.3 Design and validation considerations

Having given a somewhat less vague definition of what constitutes a SG, a question that presents itself is how one goes about designing and validating a SG. The design and development process of a successful SG is considered a difficult and demanding task, requiring a multidisciplinary approach and contribution of experts from various fields [45]. This is to be expected, as the same principle applies in the design and development process of a successful entertainment digital game. A SG, as presented in its definition, adds additional layers of complexity on top of traditional game design and its success relies heavily on its ability to produce an engaging and fun experience. Various approaches in SG design have been identified in the literature; however, no universally accepted one appears to exist [46]. Researchers have proposed frameworks and guidelines in an effort to improve SG design standards and advance the field. In addition, various attempts have been made towards establishing theory- and evidence-based design frameworks for SGs. These frameworks usually employ behavioral theories to enhance the impact and capacity of the SGs. Based on the findings of a recently published literature review, the effectiveness of SGs often relies on the successful employment of such a theory-driven design approach [47]. The most commonly used theoretical basis is reported to be the Social Cognitive Theory (SCT) [48], [49]. Based on insight gained form the Social Cognitive Theory, in order to develop effective SGs, the game design should be structured on the basis of a conceptual design framework that takes into account user experience and interaction. Another approach, that is often applied in conjunction with behavioral theories [50], is a user-centred design logic [51], in an effort to better understand user needs and incorporate them in the SGs. This is in contrast with designer-centred design logic, that promotes designer decisions over needs identified by players, implying that the expert has a better understanding of their needs. With user experience becoming increasingly important in modern interactive systems in general, user-centred design approaches have been gaining ground in the game industry as a whole [52].

Regardless of the preferred design approach an important element of SG design is the proper application of a conceptual framework that represents and communicates the SG's design in a coherent and non-conflicting manner [45]. Adding to that, the underlying framework must not lose focus from the balance that needs to be maintained between the fun and engagement element and the "serious" purpose of the SG [53]. Entertainment should never be sacrificed in an attempt to achieve the main, by some definitions, objective of the SG, as that would instantly defeat the purpose. Finally, another important aspect of SG design is the application of game metrics and analytics throughout the whole process of creating a game. Metrics and analytics have been an important part of software design and development since its beginning, however, measuring aspects of experience and engagement during interaction with a game presents certain difficulties. Game metrics and analytics include telemetry data that can be captured during playing and provide insight regarding player performance, game progress, bugs and potential exploits, among other things [54]. To this end, methods and techniques that allow for this type of monitoring have become increasingly important during design and development, as they can facilitate the process and reduce designer bias. Fast prototyping and agile methodologies have also been applied successfully along with monitoring of game metrics and analytics in order to facilitate the development process, reduce costs and expedite it. The application of game metrics in SGs, besides their positive effect on the design process, have been associated with a capacity to predict the SG's educational efficiency [55], [56]. By predicting learning during interaction with the SG experts can enhance the intervention's efficiency through design and analysis.

SGs are applied in various fields to achieve a number of different goals, such as learning and behavior change, however, their efficacy in those tasks is sometimes contradictory. Arguments exist in the literature that the procedures employed in most validation studies of SGs may not be optimal in their attempt to measure their potential [57]. There have been several attempts towards introducing systematic approaches capable of validating the efficiency of SGs. These attempts typically include ways to model the receptivity of players regarding the serious content of the intervention in a data-driven manner [4]. A large volume of taxonomies, validation protocols, and methodologies can be found in relevant literature [58], [59], [60], [61], [62], however, as is evident in multiple occasions, there is not much agreement regarding the optimal approach when validating a SG. This introduces strong obstacles that impact SG research, design, development and deployment. In addition, the fact that systematic approaches towards validation in SGs are still considered nascent has allowed for scepticism regarding their actual potential. Despite having a multitude of research studies advocating towards their capabilities, there have been several researchers doubting the "revolution" brought by SGs [62]. With the growing volume of results indicating that SGs can in fact serve as intervention tools with a variety of goals, these opinions appear to be shifting and subsiding in the last years.

Finally, during the past few years, there has been a growing number of suggestions and directions in review articles that appear to drive and formalize the process of validation in SGs, advancing the field in the process. Comparative studies between SGs with the same goal, but borrowing elements from different genres and governed by different mechanics, is of paramount importance to investigate the effect that different game styles have on SG efficiency [62]. Furthermore, comparative analysis between SGs and more traditional software interventions can help with SG design and application, but also help distinguish the elements that can facilitate SGs in working in unison with traditional interventions. User profiling and modeling can also facilitate the process of validation in SGs by pointing out the impact different player archetypes have on intervention receptivity. When such approaches are combined with other design elements that effect user experience, such as user interface and user experience engineering, the result can be very beneficial in the SG's capacity to foster learning and change behaviors [63]. Finally, the development of modular tools for validation, with the ability to generalize and be applied to different SG types is going to assist the process of systematically analysing their efficiency in a replicable manner. SGs are generally very different in form, content, and scope, hence it is often hard to compare and discern a SG's efficiency without the application of more generalized tools. To this end, the development of such tools, along with the formation of taxonomies that point out important characteristics of SGs, is considered crucial to the advancement of the field [64].

1.2.4 Fields of interest

Over the years there have been a few SGs that are considered prevalent and successful examples of the impact these tools can achieve and the potential acceptance they can be received with. One of the earliest examples is "Microsoft Flight Simulator" (Microsoft, New Mexico, U.S.), often mentioned as the "grandfather of serious games" and one of the few non-combat civil aviation simulators in existence, with a multitude of successors. "Microsoft Flight Simulator" is currently the longest-running software product line for Microsoft and one of the longest-running digital games of all time. "Minecraft Education Edition" (Mojang Studios, Stockholm, Sweden) is another prime example of successful SG and has been applied in numerous classrooms as an educational tool that provides accessible and effective learning in various topics such as mathematics, physics, chemistry, and biology. "Darfur is Dying" (Take Action Games, Santa Cruz, U.S.) managed to attract 800.000 players in the first two months following its release. This SGs can be classified as a news game, as it attempts to shed light on the war in Darfur and expose the following humanitarian disaster through an interactive and immersive way. "Pacific: The Leadership Game" (Arc Institute, Palo Alto, U.S.) is a SG focusing on leadership training, developed with the help of over 200 experts in the field. The SG allows the player to control a group of in-game characters lost on an island and through decisions help them survive, forging leadership skills in the process. "EndeavorRx" (Akili Interactive Labs, Boston, US) is a SG designed to treat ADHD in children between the ages of eight and twelve. This is the first SG that after seven years of testing was approved by the FDA to be prescribed by healthcare professionals to help treat the disorder. In addition, this SG features adaptive capabilities, as it can tailor the intervention provided according to each child's needs.

As evidenced by some of the most successful examples of SGs, their potential as intervention tools allows for their application in various fields and practices. As the field of SGs grows, both through research and entrepreneurship, so do the potential fields of interest. New SGs with innovative scopes and purposes emerge each year, while the incorporation of novel technologies facilitates novelty and generalization. Augmented and virtual reality, advancements in computer graphics, cloud computing, smart and affordable sensors, internet of things and artificial intelligence (AI) are but a few technologies that open up new opportunities and capabilities for SGs. SGs, characterized by a heightened capacity for customization and diversity, are nowadays considered efficient tools and new ideas for application emerge constantly. Due to this, categorization of the fields of application is becoming increasingly difficult. An approach that classifies SGs in five broad categories is presented below [20]:

- Training and simulation
- Education
- Health
- Societal and public awareness
- Cultural heritage and tourism

Training and simulation are two of the most popular application areas for SGs. Various examples and subcategories of this area exist, with the most prominent being SGs for military simulations, civil relief organizations, training environments for various experts and professionals, as well as SGs for business and management. SGs in this area of interest focus on creating realistic simulation environments, often taking advantage of virtual and augmented reality technologies, besides traditional 3D game graphics. With recent advances in game engines and rendering capabilities of target platforms, training, and simulation, SGs are becoming increasingly life-like and provide efficient educational and training tools. The employment of smart agents for the incorporation of human-like non player characters enhances the immersive game experience even further. The gamified environments produced by SGs provide various benefits for companies as well, by enhancing recruitment and retention, increasing program adoption and improving work performance in general [65].

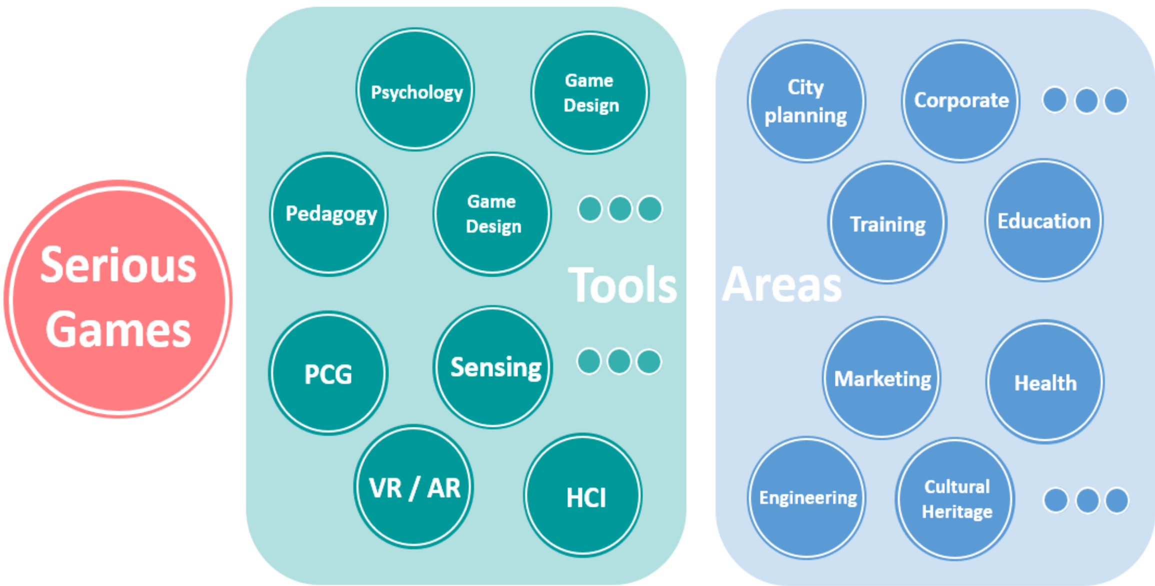


Figure 1-2: Serious game application domains and design elements combined with further concepts, technologies, and disciplines.

Education SGs are often deployed in school or home settings and aim towards playful acquisition and transfer of knowledge, as well as competence development. SGs in this application area are reported to gain increasing interest, as they exploit new visualization technologies and borrow novel game mechanics from the entertainment game industry [66]. In addition, SGs of this type tend to incorporate a variety of acoustic, tactile, and spiritual stimuli, allowing for multisensory learning [67], attract students' attention and stimulating their commitment and motivation [68].The main difference from the SGs employed for training and simulation is their target group, as educational SGs usually target students and the public. The assessment of student performances through the use of SGs is considered a difficult issue to address, however its importance is drawing attention [69]. SGs that could efficiently determine, and address personalized student learning weaknesses, by reliably assessing their performance, could prove to be incredibly valuable tools and decision support systems for students and parents.

Societal and public awareness SGs, or as they are often called games for good [20], represent a category that addresses public and societal relevant issues such as politics, religion, sexism and racism. This type of SGs is usually publicly funded or supported by non-profit organizations and charity institutions. Their content and game mechanics is usually similar to educational SGs. The category of cultural heritage and tourism SGs has many similarities to societal and public awareness SGs, sharing design characteristics for a different scope. These SGs often incorporate technologies such as virtual or augmented reality in order to simulate journeys in both space and time and immerse players in simulated environments. SGs for tourism can increase brand awareness and loyalty to the destination [70]. Innovative tools have been developed and tested by various stakeholders, such us destination management organizations, hotels, and airlines, displaying the significant potential of SGs in the field [71]. SGs for cultural heritage are considered very beneficial to the field and for this reason several frameworks that allow for a holistic design approach have been suggested [72]. Cultural heritage games are also often applied in formal educational settings and facilitate the study of history and culture, coinciding with education SGs in the process.

As SGs for health lay in the core of the present Doctoral Thesis, they are going to be presented in detail in the following sections. Finally, Figure 1-2 [20] presents an even broader spectrum of interest domains for SGs, as well as various design elements and theoretical backgrounds that are often employed for the development of SGs. This graph shows the potential of SGs to impact a large number of fields of practice and the multidisciplinary effort that goes into the creation of them. SGs are expected to impact more fields in the future as novel technologies are incorporated, the SG industry is growing, and opportunities for innovation are presented.

1.3 Serious games for health

One of the most, by some even the most [73], popular field of application for SGs and the main focus of the present Doctoral Thesis is healthcare. SGs for health act as electronic and mobile health interventions with a variety of end users including patients, their families, healthy individuals, medical professionals, and students. SGs appear to have penetrated healthcare, with innovative interventions aiming to rehabilitate, to raise awareness, provide clinical decision support systems, to diagnose and monitor conditions, to act as therapeutic interventions, to facilitate healthy behavioral change, to promote disease self-management and to educate both patients and health experts [74], [75]. This is evident through the rapid increase in relevant research publications in recent years, covering multiple medical practices, health conditions and diseases. Additionally, the development of COVID-19 has shifted medical workers and educators towards telemedicine and remote education, where SGs can constitute effective intervention tools, increasing their value and highlighting their capabilities even further [76].

One of the earliest examples of digital games designed and developed for healthcare, Captain Novolin (Acclaim Studios, Salt Lake City, U.S.), published in 1992, aimed to promote diabetes self-management in children suffering from T1DM. Captain Novolin was a platformer game released for the Nintendo console Super-NES. The plot of the SG pitted a hero against aliens that had abducted a diabetic mayor, who had enough insulin to survive only for 48 hours. Captain Novolin included advice by a doctor and dietician, and the gameplay consisted of trying to avoid aliens who had the form of junk food. The SG was funded by Novo Nordisk and the National Institutes of Health (U.S.) and was received with positive acceptance from diabetes specialists and children suffering from type 1 diabetes. This was evident in two studies [77], [78] that were conducted among children, which indicated their acceptance of the intervention and demonstrated the SG's potential to help parents speak with children about diabetes. However, Captain Novolin was accepted very poorly by video game critics and has been named since as one of the worst video games of all time. Since then, SGs for health have evolved into complex and efficient health intervention tools.

Some recent examples of SGs for health feature the incorporation of novel technologies to produce state-of-the-art interventions with significant benefits over traditional ones. In [79] virtual surgical planning, an integral part of computer-assisted surgery, is conducted through a virtual reality environment to reduce or overcome potential shortcomings through the benefits of visuospatial vision, bimanual interaction and user immersion. Results from a study indicate faster learning through the virtual reality environment and participants reported less fatigue and a more intuitive experience. In [80] a SG aiming to reduce HIV-related stigma and promote participatory gamification culture for health interventions is presented. The SG employed assets from user-generated content and a randomized trial was conducted to evaluate its performance. The game intervention was largely preferred by participants and the SG showed an advantage in reducing intimacy stigma. In [81] a SG improved cervical cancer screening attendance, as well as detection of women with high risk of cervical cancer, amongst women aged 20 to 69. The SG was a mobile application featuring educational content and was given to participants bound to have a cervical exam in the following year. In [82], a user-centred design approach was employed for a home-based exergame featuring on-body feedback sensors. Most of the participants found the SG supportive and motivating and stated that they could imagine using the SG at home and playing it in the future. In addition, the majority of participants reported that they enjoyed the on-body feedback system. Finally, in [83], a SG for people with mild to borderline intellectual disability incorporating immersive virtual reality avatars and user-centred development is presented. The SG displayed high usability among participants and results indicated potential application of similar approaches for tailored mental health therapies.

As is evidenced by the few examples presented above, design of SGs for health is nowadays highly diverse and needs to address various problems of increased complexity. A preliminary investigation of the relevant literature regarding SGs for health revealed a huge volume of scientific publications on the field. The terms used in these articles varied greatly, with multiple terms and definitions used for SGs, the objective of the reported health interventions, and the desired outcomes. A large number of review articles has been identified, however, cohesion in the research field of SGs for health appears to be lacking, as researchers from a multitude of backgrounds meet and report their findings in different manners. The limitations described in the sections above regarding definition and clarity in the research field of SGs appear to be present in the literature relevant to the application of SGs in healthcare as well. In fact, these limitations seem to be greater in this case, as the field of health is highly extensive and diverse, with multiple possible intervention objectives and target groups of players. An attempt [43] to classify SGs for health based on target groups and purpose can be seen in Figure 1-3. If elements of game design and genre, along with different types of technologies employed, and particular medical fields of practice are added to the attempt to properly classify SGs for health, it quickly becomes evident that the limitations mentioned before might indeed be justified. SGs for health create a complex field of research and innovation, with many limitations in terms of design, development, validation and application, but also ripe for successful attempts to create new and novel health interventions.

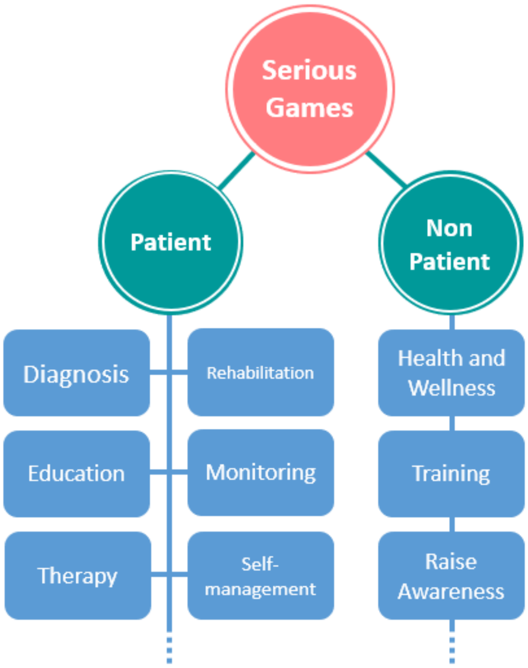


Figure 1-3 : Classification of serious games for health according to player and objective.

An extensive investigation of the relevant literature has been conducted during the present Doctoral Thesis. In an effort to map the field of SGs for health, a search for review papers published in the previous two decades investigating SGs for health was conducted in a semi-systematic manner. Review papers were selected instead of research articles in order to provide a basis for discussion about the direction of the field and current application status. The search engines included in this overview were Google Scholar, PubMed and ScienceDirect. After examination of the relevant literature and terminology the following search strategy was selected (Table 1). The key words and concepts were selected based on a similar search strategy presented in a recent review article [75] that focused on health education. The conducted search did not include terms relevant to "simulation", as simulations do not constitute necessarily games. Identified publications that revolved around simulations that incorporated gamified elements were included in the search results. The term gamification was included in the search strategy due to its conceptual ambiguity with the term SG [84]. Only peer-reviewed review papers written in English were included in the results. In addition, only publications dealing solely with SG interventions, and not e-health or ICT tools, were included. Results that included the term gamification were included if they reviewed articles that presented SGs as well. Finally, duplicates were removed and findings from investigation outside the search strategy were included.

Table 1: Search strategy

Concepts	Keywords
Game	Game OR Serious Game OR Video Game OR Digital Game OR Gamification OR Gaming
Health	Health OR Healthcare OR Clinical OR Medical OR Medicine OR Rehabilitation
Desired outcome	Rehabilitation OR Diagnosis OR Behavior OR Education OR Training OR Management
Condition	Cardiovascular OR Surgery OR Upper limp OR Diabetes OR Mental OR Depression OR Nutrition OR Dentistry OR Cancer OR Emotion

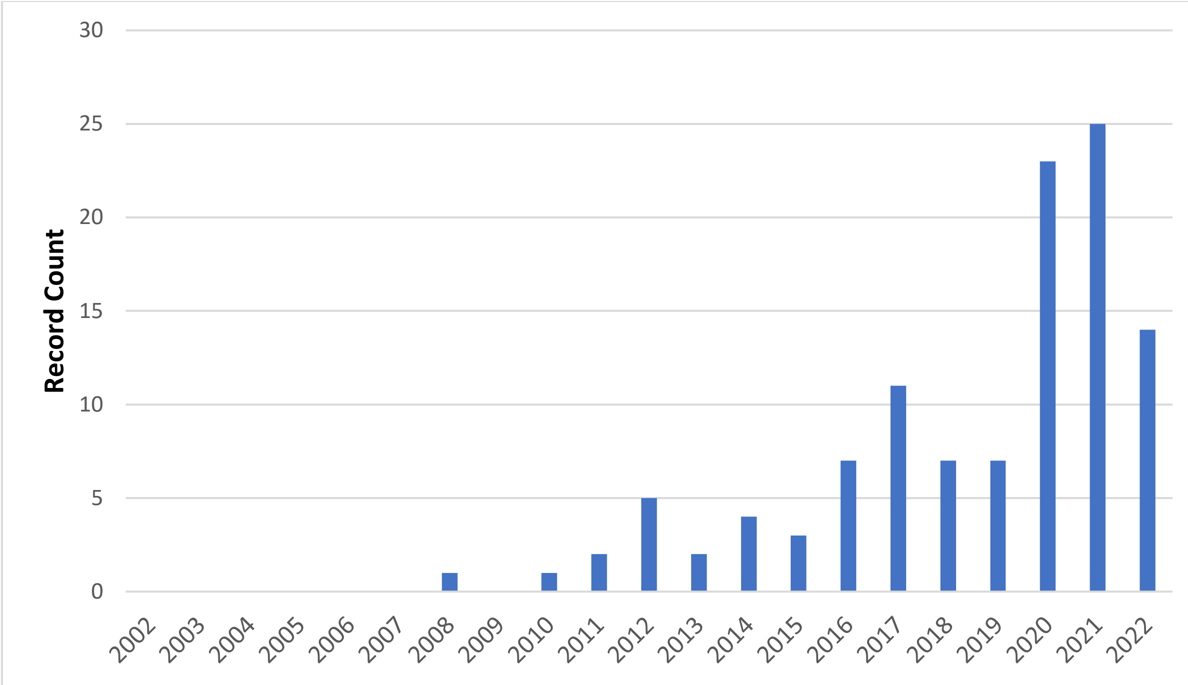


Figure 1-4: Number of publications of reviews about serious games for health per year.

After removing articles that did not constitute literature reviews and articles not associated with SGs for health based on their title or abstract, a total of 113 review papers, published between 2002 and 2022, remained. The search included articles that were published up until July of 2022, hence publications for 2022 cannot be considered complete. Figure 1-4 presents the number of identified review articles per year of publication. No relevant review articles were identified until 2008. This finding can be strengthened by findings reported in [43], where few SGs for health were identified before 2004 in research literature. In this earlier review paper, an increase in research articles regarding SGs for health has been identified for the period of 2004 to 2008. This justifies the publication of review articles investigating these studies from 2008 and onwards. The first identified review article in the field, published in 2008 [85], investigated twenty-seven research articles reporting results from design, development and validation of twenty-five SGs that promoted hearth-related behavior change. Starting from 2008 and until 2015 a relatively small number of review articles per year is depicted in the Figure 1-4, however a growing trend in publications per year begins to emerge. From 2016 and on, the volume of review articles per year increases rapidly and the growing trend in publication frequency becomes more evident. These findings are reinforced by the findings a recent review study [76], about the application of SGs in healthcare. The article's search strategy, from 1998 to 2022, illustrates vividly the increasing trend in research about SGs for health. This further substantiates the results of the search, conducted for the needs of the present thesis, in an attempt to map the field of SGs for health and search for limitations and opportunities.

An attempt to classify the identified articles according to their desired outcome is presented in Figure 1-5. The generated groups were based on the article title and abstract. SGs about medical expert training and mental health appear to gather significant research interest, followed by chronic disease management and rehabilitation. Rehabilitation, one of the first and most important fields of application for SGs for health appears to lose some ground in terms of research interest over the years, as SGs for training, mental health and chronic disease management become increasingly prevalent. Results from the identified review articles indicate that SGs for health are employed for a variety of purposes and have the capacity to provide meaningful interventions that gather increasingly higher research interest.

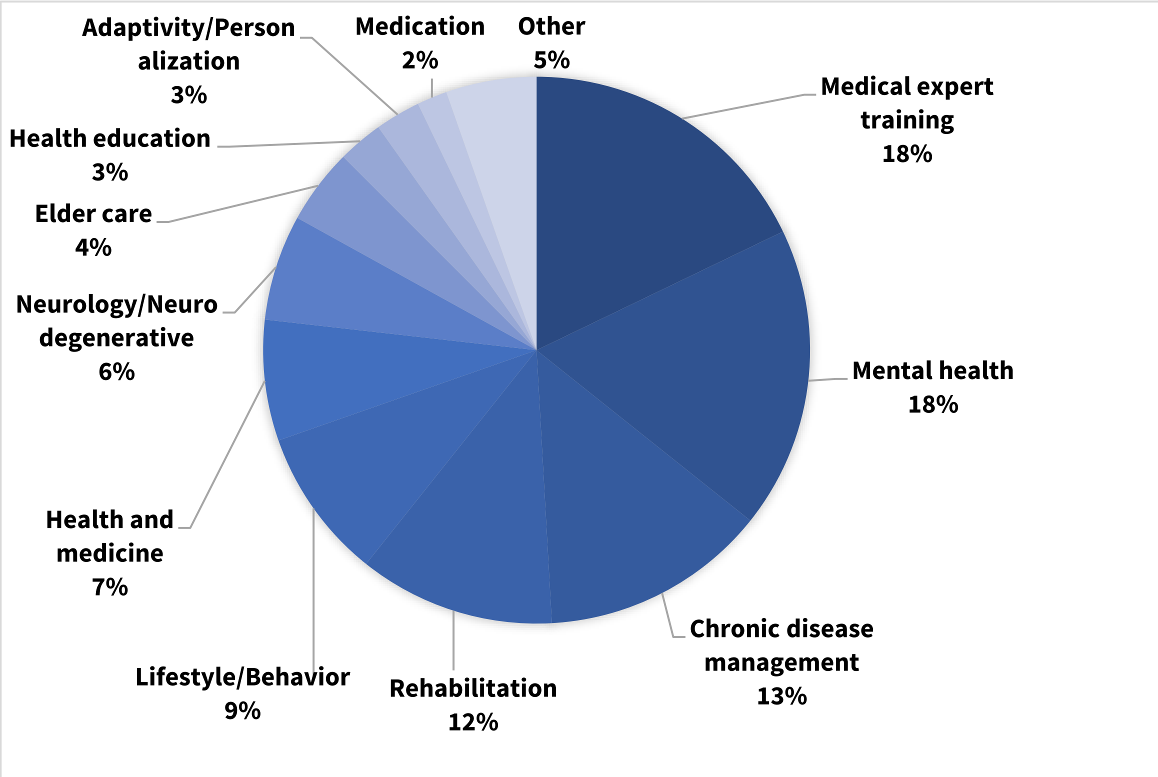


Figure 1-5: Classification of identified review articles.

This increase in interest for game interventions in health indicates that SGs have the potential to fit the needs of the evolution of healthcare towards predictive, preventive, personalized and participatory interventions and tackle modern problems. However, limited review articles that focus on adaptivity and personalization for SGs for health have been identified, indicating that research articles for smart SGs for health are yet not so prevalent. Personalization has proven to be of paramount importance in health interventions, focusing on individual patient needs, while providing a sense of agency to the user over their healthcare, enhancing their participation [32]. Digital games provide a multitude of tools and mechanics to be employed for personalization, and as such SGs can be designed to be adaptive and smart. This adaptivity not only enables the selection or generation of SG content that fits the user's clinical needs, but also induces a state of flow [86]. In addition, adaptive SGs can help maximize engagement during training and education, while evaluating user performance and weaknesses [43]. Finally, even with the incorporation of technologies that allow for adaptivity, SGs for health remain affordable and easy to use interventions that can engage remote and hard-to-reach groups in a point-of-care manner [87]. According to the findings presented above, adaptivity in SGs for health is still a nascent field of research compared to other more prevalent ones. The potential value of incorporation of such techniques, and their success in other fields of application, are certain to drive future research directions towards such interventions and present the motivation for the present Doctoral Thesis.

1.4 Adaptivity in serious games for health

Adaptation is an important field of research in games. A common definition for the term is the "ability to make appropriate responses to changed or changing circumstances" [88]. In computer science, adaptation refers to the property of a system to change its behavior by considering information collected about the user. This property often applies to digital games, as many of their mechanics are in essence adaptive techniques that are driven by player performance and interaction. These player-centered mechanics offer significant advantages when compared to static, player-agnostic mechanics, as they try to match the player's cognitive skill and preferences and deliver a better experience [89]. Adaptive techniques in games are continuously evolving, with new technologies mostly from the field of AI being employed. Modern entertainment games have the ability to create player models, procedurally generate content, dynamically adjust difficulty levels and present human-like adaptive non-player characters (NPCs). Some modern examples of successful digital games that employ such techniques are “Stardew Valley” (ConcernedApe and Chucklefish, London, England), a story driven resource gathering and crafting game that incorporates map generation, reward and dynamic difficulty adjustment, “No Man's Sky” (Hello Games Ltd, Guilford, England), a space exploration survival game that generates new planets and landscapes the moment the player discovers them and the Borderlands series (Gearbox Software L.L.C, Frisco, Texas and 2K Australia Pty Ltd, Canberra, Australia), a competitive shooter game that incorporates weapon generation.

The benefits of employing techniques for adaptivity in games are plenty. According to the model of flow [90], a construct describing an optimal experiential state that involves complete immersion in an activity and a deep sense of enjoyment [91], in a game if a player finds the game too difficult they will become frustrated, and if they find it too easy they will get bored. Adjusting game content according to player feedback allows for a maintenance of the state of flow [92]. This in turn promotes engagement and augments the quality of the game. Additionally, techniques that imbue games with adaptivity can reduce production costs. Procedurally generated content is often employed to achieve that goal, but also promote game re-playability. With the need to create new and novel games increasing by the year these approaches allow for small game studios to meet the challenge and promote the industry. Finally, adaptive games promote inclusivity, as they provide opportunity to tailor their content according to player needs, regardless of gender, age and experience [93].

Adaptive learning systems are defined as “learning programs capable of adapting themselves to the individual abilities of the learner, e.g., previous knowledge, interests, weaknesses or preferences with regard to forms of representation.” [94]. SGs as intervention tools can take advantage of the adaptive capabilities present in games and adapt their content according to the player’s individual needs. In this manner, SGs accommodate the player’s learning pace and style, enhancing their efficiency towards the target outcome. However, these adaptive techniques when incorporated in SGs not only facilitate learning or skill acquisition but also improve the overall gaming experience by engaging the player [92]. As a result, heterogeneous player groups can benefit from customized experiences adapted to their needs. Adaptivity in SGs for health aims to enhance their capabilities as health interventions by dynamically generating game content during player interaction. Despite recent advances in the field of SGs for health, limited research has been conducted on tailoring persuasive game design to specific players’ needs. The reported results display that differences in receptivity of learning and behavioral approaches in SGs for health, among multiple user types. As a result, one-size-fits-all design approaches are not expected to be always effective. Furthermore, difficult instructions and tutorials, disproportionate levels of difficulty, and inadequate reward mechanisms are game elements that can demotivate the player [95]. A recent review study [6] highlights the importance of delivering personalized content in SGs and employs the term “individualization” for this purpose. Individualization in SGs for health can not only enhance game experience, but also address specific user needs linked to clinical outcomes, like disease self-management or knowledge acquisition. Affective player modelling and PCG can be applied to a great extent to provide SGs for health these individualization properties.

1.4.1 Affective player modeling

Player modeling is the study of computational models of players in games [96]. These models involve the detection, modeling, prediction, and expression of characteristics describing the player and manifesting through cognitive, affective and behavioral patterns. The detection of these characteristics is based on data collected during the interaction with the game and describe the complex phenomena that occur during play. This type of data can be applied towards modeling player actions, tactics, and strategies. Player modeling, therefore, is the study of computational and AI techniques for the generation of models of cognition, behavior and emotion, as well as aspects beyond elements of game interaction, such as personality and background [96]. In the case of SGs for health this can be enhanced to include clinically relevant information that can be in turn translated in adaptive properties. Player modeling can be applied towards various objectives in games [96]. Applications of player modeling in game design and development include:

- Adaptive player experience and game balancing, which applies models in achieving a balance between game difficulty and player skills.
- Personalized game content generation, in which generated content is based on specific player needs identified through player models.
- Believable game agents, in which behavioral, cognitive and affective patterns recognized be player modeling are applied towards realistic human simulation.
- Playtesting analysis and game authoring, driven by relationships between game context, player state and in-game behavior yielded by player modeling.
- Monetization of free-to-play game, where microtransactions are optimized by modeling player purchasing behavior and its link to in-game behavior.

Player modeling is considered by some researches to be a rather loose concept [97]. It can involve everything from the expected player reactions in response to game events, to machine learning (ML) models that predict player actions based on large volumes of collected data. This lack of cohesion introduces difficulties for researchers that try to navigate the field. A player model can be easier defined based on four facets, namely, the scope of the application, the purpose of use, the domain of modelled detains, and the source of a model's motivation [97]. For player models applied to deliver adaptivity in a SG for health player's affect can provide valuable insight towards improving adherence to the intervention. Affective computing refers to processes that attempt to assign machines with human-like capabilities of observation, interpretation, and simulation of affect [98]. The field aims to enhance the quality of human-computer interaction while improving the intelligence of the machine. Early research attempts that investigate the nature of affect and emotion can be traced back to the 19^{th} century [99], however the concept of affect was rarely connected to machines and was traditionally studied by psychologists. Nowadays the field of affective computing is considered a multidisciplinary field of research that encompasses computer science, psychology and cognitive science. Affect, or emotion, often carries a stigma in science, believed to be inherently non-scientific, with little reason found to be brought into tools of science [100]. This has been traditionally particularly true regarding attempts to bring "affect" into applications of computers. However, given the essential role affect has in both human cognition and perception, the reasoning behind the field of affective computing is that machines with affective capabilities will improve the quality of the assistance they provide. Incorporating techniques for affective computing in games is often referred to as affective gaming [101] and is considered to be a field of research that attracts increasing attention during the last two decades. The gaming and game design community appears to understand the importance of affect in the development of engaging games. The main focus of affective computing techniques in game design is the recognition of player emotion and the dynamic adjustment of game content according to it. An additional aim of these techniques is the generation of human like emotions in NPCs, however less emphasis is placed on the modeling of such emotions in comparison to understanding how player affect responds to game content.

A recent study [102] that compared versions of the same game, one that incorporated affective computing techniques to interpret player emotions and enhance game experience, while the other did not, indicated statistically significant differences in the experience of the participants as it was reported in relevant questionnaires. In particular, most of the positive experience ratings were increased in the version that employed player affect as a means to adjust game content. However, literature suggests that it remains a difficult task to merge affective computing techniques and game design elements without reducing immersion and worsening the experience that is offered by the game itself. To this end, multiple approaches have been proposed [89] as means to incorporate affective computing successfully as part of game design, however dynamic difficulty adjustment seems to be the most exploited on in literature [103].

Multiple means of identifying human emotion during play have been investigated. The most common sources for affect information [102] during gameplay are:

- Facial expressions
- Body actions and posture
- Brain computer interfaces
- Haptics
- Sensors (EEG, PPG, EKG, respiratory rate devices, galvanic skin response, electrodermograph, EMG)

To this day, there not many commercial games that can be characterized as affective games. Some examples are, Left 4 Dead 2 (2008), in which player stress levels gathered via galvanic skin response determine the pace of the game. Brainball (2003), where the player moves a physical ball by trying to achieve a relaxed state that is measured through EEG. Oshiete Your Heart (1997), that simulates dating and monitors player heart rate and sweat levels. Missile Command (1980), where a player must destroy moving targets and the game environment changes according to their heartbeat rate. Bionic Breakthrough (1983), in which sensors pick up facial movement the player needs to control in order to move a bouncing ball into a brick wall. Several attempts have been identified in the literature regarding games that have been designed and developed in order to assess the potential of affective games to enhance game experience [103]. These games tend to feature PCG techniques that alter the game environment, game-controlled elements, player-controlled elements, NPCs and game objects. Findings from these studies indicate that despite the fact that these techniques seem to be enhancing game experience it is critical that they don't limit creativity in game design.

Focusing on SGs in the health sector, the rapid advancements in sensing technologies make feasible the implementation of patient-tailored interventions supported by properly designed SGs [3]. More specifically, sensor based adaptive SGs, including player modeling and profiling, in terms of health status and lifestyle habits, along with recognition of engagement in real time, have the capacity to address important challenges in chronic disease management such as the presence of inter- and intra-patient variability while offering low-cost services at the point of care. Research on the recognition and employment of engagement and other affective states to achieve individualization in SGs for health is still limited. A handful of relevant publications have been identified, with the most common case being biofeedback SGs for stress management. Features extracted from breathing signal and heart rate variability (HRV) analysis have been used to predict affective states, such as stress and engagement during game play, in a biofeedback context for stress management therapy [104], [105]. The use of ECG signal transmitted in real time to a therapist has also been reported in the context of a virtual reality SG for emotional regulation in adolescents [106]. Moreover, a methodology for multimodal affect recognition for SGs targeting the treatment of behavioural and mental disorders and chronic pain rehabilitation has been presented [107]. A SG for automated personalised exposure therapy that includes experience-driven PCG has also been proposed, employing ML techniques to predict stress from physiological signals [108]. Additionally, emotion recognition has been applied on speech components to support SGs aimed towards cognitive-based treatment for mental disorders, with results indicating successful recognition of interest, boredom, and anger [109].

Engagement has been argued to be of paramount importance in terms game experience [110], related with positive and negative affect through game activities and accomplishment of objectives

[3]. In particular, engagement is related to a range of emotions, such as satisfaction, fun and enjoyment, as well as a range of experiences, such as feeling immersed or in the state of flow. Different perspectives have been identified regarding the construct and measurement of engagement according to a recent review study [37]; the study adopts a three-part framework for engagement that includes the dimensions of behaviour, cognition, and affect. Recognition of the affective aspect of engagement and its employment in an "emotion-sensitive adaptive game approach" [5] can lead to heightened task persistence and improved learning process. There are various sources to collect informative data regarding player engagement during game play, such as self-report (e.g. game experience questionnaires), in-game metrics, wearable sensors (e.g. electrocardiogram (ECG), electroencephalogram (EEG) and electromyogram (EMG), electro-dermal activity), and posture recognition sensors [37], [111]. More specifically, pressure sensors, for posture and mobility monitoring, have been employed for this reason in learning environments [112], or during intense cognitive activity [113]. Features extracted from heart rate sensors have been identified as potential detectors of affective states, stress, and learning [114], [115]. Data collected from sensors can be augmented by in-game metrics and analytics that exhibit promising associations with learner's engagement [116], [117].

1.4.2 Procedural content generation

Player modeling can facilitate adaptivity in games by guiding PCG techniques. PCG is a term used to describe techniques incorporated in games to empower user engagement and increase replay value by generating new content based on user choices and interaction with the game automatically. A formal definition for PCG in games describes it as "applications of computers to generate game content, distinguish interesting instances among the ones generated, and select entertaining instances on behalf of the players" [118]. Besides the benefits associated with user experience, PCG techniques reduce game production cost and facilitate player-driven approaches and player generated content [119]. There are two types of games incorporating PCG in terms of whether they hide the process that generates content and assets from the player, or they provide access and control of it to the player [119]. In most games with PCG the player has some form of indirect control over the generated content through their in-game actions, or in more advanced approaches through data collected while playing from a multitude of sources. In most of these cases the player is not aware of the control they are given, or the effect their actions and decisions have on the generated content. However, there are cases where control over the generation of content is given to the player in a more direct manner. Perhaps the simpler examples of such approaches are story driven narrative games, where the player is given options that control the game narrative. The PCG method that controls the story is in this way driven by the player, who usually has a clear understanding of cause and reaction. In other, less simple, cases the player develop direct sense of control over the generated content, with the generator opening up to them and becoming an element of gameplay.

According to a recent review study about modern trends in PCG techniques for games there are numerous approaches, algorithms and methods employed. Researchers and developers apply them not in a systematic manner [120]. Some broad categories that were identified during the review process were:

- Answer Set Programming, a programming approach that is based on tree search based on known facts about the game space. This approach has been successfully employed for level generation.
- Artificial Neural Networks, ML models able to output PCG assets and game content, such as personalized quests and dynamic narrative.
- Coevolutionary genetic algorithms (GAs), an evolved approach that employs GAs and is based on subjective fitness functions that avoid early convergence states and deliver optimal game content.

- Dynamic difficulty adjustment, a broad category of simpler approaches that is quite popular due to their ability to tailor difficulty levels according to player interaction and improve engagement to the game.
- Experience driven PCG, another broad category that uses player experience, such as gameplay, skills, activity, or even questionnaires, in generic approaches through rule-based systems and are often the basis for other more advanced forms of PCG.
- GA, a concept of applying natural selection on chromosomes representing game content and are deemed fit according to input provided by the current game state and player data.
- Monte Carlo tree search, algorithms that select, expand, and simulate through backpropagation, measuring playing patterns in the process and altering the design space.

SGs, those that target health in particular, can benefit greatly from PCG techniques [3]. However, according to a recent review article on the use of PCG [120] in digital games, incorporation of PCG is SGs is far less documented when compared with its prevalence in entertainment games. Advancements in the fields of data analysis and AI, along with the development of low-cost, portable, and unobtrusive sensors, enable real-time individualization of PCG based on data collected from a multitude of sources [92]. PCG methods can be augmented through real-time recognition and employment of engagement in a constant feedback loop that adapts game content based on the player state [121], [122]. Novel technologies, from the fields of ML and deep learning, utilizing data from various sources have been applied to enhance PCG in entertainment games [123]. For instance, a recently proposed PCG framework employs intrinsically motivated reinforcement learning that builds knowledge about the player's preferences by searching for unexplored information and being rewarded for discoveries [124]. Intrinsically motivated reinforcement learning, thus, makes feasible the development of experience driven PCG that considers the impact of the generated content on the player's affective state. Such frameworks can be combined with novel techniques that procedurally generate individualized content specific to self-health management needs and preferences [125]. The procedural generation of SG content built to accommodate educational and behavioural objectives regarding the targeted condition, is thus controlled by an engagement feedback loop. These objectives include, amongst others, knowledge about the management of the condition and daily self-health management goals [126]. Maximizing player engagement not only promotes the SG's effectiveness towards these objectives, but also increases adherence to the intervention, leading to sustainable improvement in self-health management. Additionally, sensors employed for the recognition of engagement can produce clinically relevant data or lifestyle parameters [127], [128]. The integration of such data in the SG feedback loop, along with in-game metrics, can ultimately lead to PCG that individualizes SG content according to condition and player specific needs while promoting adherence through engagement. SGs for health benefit greatly from the incorporation of state-of-the-art sensing technology [129]. PCG techniques employing this type of sensing data can produce patient-tailored and clinically relevant content, resulting in a smart personalized health intervention. The availability of real-time sensing data providing information regarding the user's lifestyle, behavioral habits and health status, makes feasible the development of sensor-based adaptive SGs with increased capacity to address important challenges in self-health management [3], [130].

1.5 Thesis structure

Motivated by the findings presented above the present Doctoral Thesis focuses on investigating the feasibility of a novel conceptual framework in terms of accurately and reliably recognising engagement in real time and delivering efficient personalized content in SGs for health. In the following chapters results from two experimental procedures and a pre-pilot study, aiming to assess the effectiveness of the developed tools, evaluate user acceptance, and evaluate the proposed framework are presented. These procedures employ novel SGs for health designed and developed to take advantage of the capabilities provided by the proposed conceptual framework.

In chapter two, a novel conceptual framework for real time recognition and PCG in SGs for health is presented. The conceptual framework employs heterogeneous data that allow for generation of personalized content. Additionally, a novel approach for PCG based on a GA is detailed.

In chapter three the design and development of two SGs that aim to raise awareness, educate, and empower self-disease management is presented. The design framework that formed the basis for the development of the games is described in detail. The purpose of the developed games is to evaluate the feasibility and efficiency of the proposed conceptual framework. To this end, the games are designed to incorporate mechanics for PCG. Additionally, a platform employed in pilot studies that incorporates a SG for health is also presented. In conclusion, a novel technique for PCG in SGs for health based on a GA is described.

In chapter four, a carefully designed experimental process that aims to collect a multitude of data during play with a SG for health is presented. The collected data are then analysed in an effort to highlight potential features with predictive power towards real time recognition of engagement. Engagement is annotated in real time through screen recordings of play by participants. Data are collected from heart rate and pressure sensors, as well as interaction with the SG.

In chapter five, an experimental process and results from a pre-pilot featuring SGs that employ the proposed GA technique for PCG are presented. Multiple versions of the GA technique, featuring different parameters and aiming towards generation of different type of content are evaluated in terms of acceptance and efficiency of delivered content.

In chapter six, a methodology for employing deep learning agents that automatically play a SG for health is presented. The aim of this approach is to provide a means to evaluate and calibrate PCG techniques in SGs that serve as health interventions, as well as evaluate the ability of agents to be trained faster according to the content presented by the SG.

In chapter seven a discussion regarding the main conclusions of the research conducted in the present Doctoral Thesis is conducted. The contributions made towards the field of creating adaptive SGs that can serve as efficient and enduring interventions for health are detailed. Finally, future research directions focussing on trends recognised in the present study are presented.

2. Conceptual framework for adaptivity in serious games for health

In the present chapter a novel framework for adaptivity in SGs for health is presented. This framework has been designed to employ novel technologies towards smart and personalized health interventions based on the research trends and healthcare needs that were identified in the previous chapter. To this end, the aim of the proposed conceptual framework is twofold. On one side, it aims to create player models that deal with the affective aspect of engagement and apply them in the domain of SGs for health in order to promote adherence to the intervention. On the other side, the aim of the proposed conceptual framework is to model the player's clinically relevant needs in a way that can be comprehensive to the SG's PCG technique and produce tailored content. In its most basic form, a structure that allows for dynamic adaptation in SGs for health requires three elements [11]. Firstly, data and information about the player, relevant to the desired outcome of the dynamically generated content, need to be collected. Data collection can be conducted through various means, including questionnaires, observation, interaction with the SG, and sensors. This data can be both dynamic, e.g., game interaction, physiological signals, and static, e.g., demographic data, medical file information. Secondly, a player model must be constructed, based on the collected data, which defines health relevant needs, as well as game related preferences of the individual player. In this way, the player model will incorporate the capability to facilitate the empowerment of engagement, but also reliably detect clinically relevant needs. Finally, an intelligent agent must be designed and implemented within the SG to generate or select the individualized content. In this manner the structure for dynamic adaptation receives useful information about the player, understands it and generates game content according to their need.

A few frameworks for adaptivity in SGs have been identified in the literature. The Square Dance framework [131] allows for the development of adaptive SGs by taking into account the learner's behavior and success in game tasks and adapting difficulty and game environment. The framework is based on two components, the Director and the Module, with the first controlling the employed modules that are responsible for high-level game structures, such as mini-games and levels. The Cognitive Adaptive Serious Game Framework [132] measures the learner's cognitive load through an adapted version of the detection-response task by monitoring game performance and evaluating the learner's knowledge and experience. In this way the framework has the potential to deliver real-time personalized training environments. This framework expands on an existing and widely used conceptual framework [133] for SGs by including the capacity for real-time adaptivity. The I-Mouse framework [134] employs a combination of eye and mouse tracking data to train ML models that predict the need for player assistance in education SGs. Another approach [135] features a non-linear methodology for adaptive SG design that provides space for macro-adaptation of the game environment and the learning objectives based on the player's performance.

Focusing on frameworks for adaptivity in SGs for health, RehaBot [136] is a framework for adaptive generation of personalized SGs for remote rehabilitation. This approach employs motion tracking sensors, virtual reality environments, and virtual assistants to guide rehabilitation exercises with personalized guidance and dynamically adjust their difficulty according to the patient's needs. Another approach for adaptivity in SGs for rehabilitation, EasyAffecta [95], features a recognition module for creation of an affective player model in real time through face mapping based on the Facial Action Coding System. Information about the patient's affective state is then used by the adaptation module that adapts game content accordingly. An architecture supporting integration of SGs aiming to reduce sedentary time, employing internet of things and AI technologies to dynamically adapt game content and satisfy specific player needs was presented in [137]. Finally, a framework [138] that employs data from body sensory network, smartphone applications, and social media to develop context-aware SGs for a healthier lifestyle was identified.

2.1 A proposal for adaptivity in SGs for health

A novel conceptual framework for adaptivity in SGs for health, focussing on self-management in chronic conditions is presented. Based on the presented theoretical background, the proposed conceptual framework features PCG techniques that take advantage of heterogeneous data and player modeling to achieve their objective. Real-time recognition of engagement is employed to create a constant affective feedback loop that empowers adherence to the intervention by maximizing engagement. This control system helps the player maintain the desired state of flow through tailored game content and dynamically adjusted game difficulty. This in turn increases the player's adherence to the intervention provided by the SG. In addition, challenges in health interventions, such as chronic condition management, diagnosis, and raising awareness have been taken into consideration towards designing the proposed framework. Since several lifestyle factors affect the onset and the progress of health conditions, the monitoring of lifestyle habits and promotion of effective behavioral lifestyle changes constitute important challenges. Furthermore, inter-, and intra- patient variability is present in chronic conditions, requiring thus the delivery of personalized approaches. The ability to generate SG content tailored to specific player needs is expected to enhance the intervention's educational efficiency and increase its potential towards achieving effective and sustainable behavioral change. Achievement of sustainable health status is also promoted through the increased adherence to the intervention promoted by the constant engagement feedback loop. SGs addressing the aforementioned challenges can significantly contribute to empowering patients with chronic conditions in self-health management. In this sense, adaptive SGs benefiting by the proposed conceptual framework could offer personalized experience adapted to the patient's health status, lifestyle, and preferences, while maintaining increased levels of engagement.

To achieve these goals, the proposed conceptual framework takes advantage of affordable and unobtrusive sensor technology to measure the affective state of player engagement and provide insight regarding their specific health related needs. As it is depicted in figure 2-1, the proposed conceptual framework consists of four layers: (i) the serious game layer, serving as a health intervention, (ii), the PCG layer, that is responsible for generating appropriate game content adapted to the player's needs and preferences (iii) the real time engagement recognition layer, which utilizes sensing technologies and in-game metrics, and (iv) the health related data layer, that is responsible for generating important information regarding the player's health status and behavioral lifestyle habits. Compared with similar frameworks identified in the literature, the proposed framework aims to provide SGs for health with real-time adaptivity capabilities, with the generated content focusing on maximizing engagement and targeting the player's health related needs according to the SG's purpose. Real-time adaptivity is fitting to the dynamic nature of game experience, with the potential to further enhance engagement and increase game retention. This is of great importance, especially with mobile game applications, as player retention is generally low due to the multitude of available options. The main application field of this framework is SGs that aim to empower self-health management in chronic health conditions. These games promote sustainable and effective behavioral change, while educating and raising awareness regarding their target condition. This type of SG based health interventions benefit the most from increased adherence due to the difficulty of maintaining recently adapted healthy behaviors. Chronic conditions such as obesity, diabetes, depression, obstructive sleep apnea (OSA), are directly linked to dietary habits, physical activity, and sleeping quality. Sensing technology applied towards monitoring of these signals that provide insight about the progression of the condition can be of crucial importance towards more positive outcomes. However, the proposed framework is designed to feature generalization capabilities towards other types of SGs for health, such as exergames, lifestyle improvement games, and rehabilitation games.

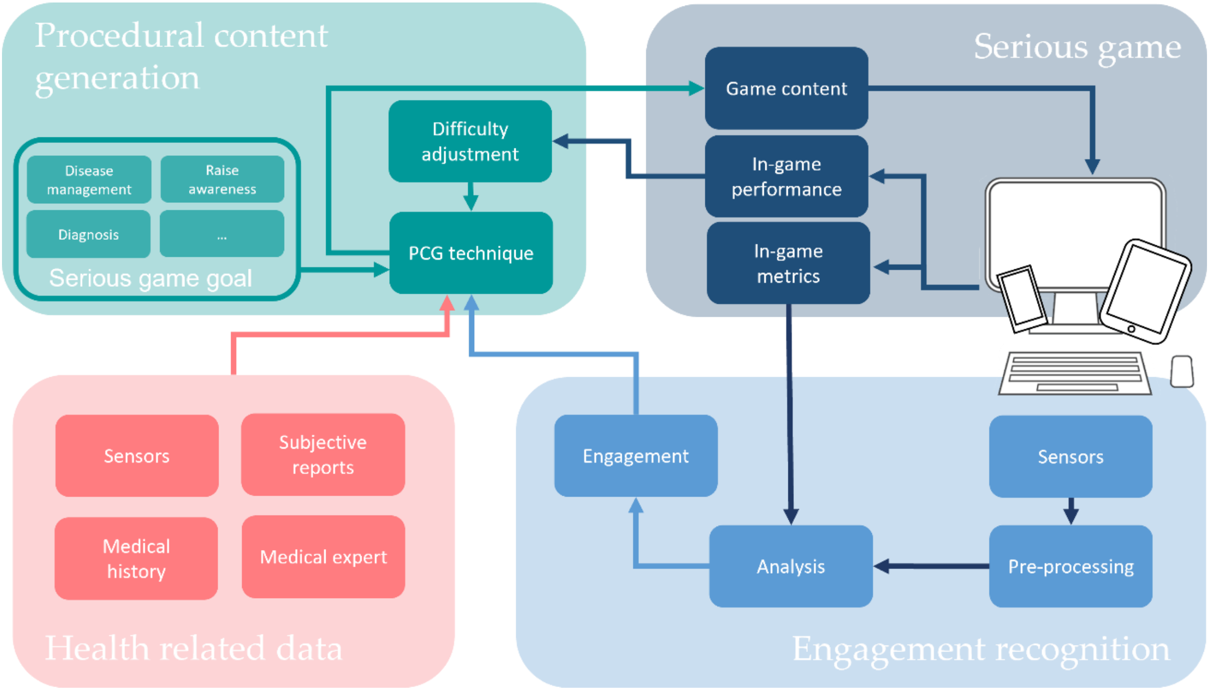


Figure 2-1: Conceptual framework for adaptivity in serious games for health

Each of the four layers of the proposed conceptual framework along with their components are presented below. In the following chapters the feasibility of the proposed framework is going to be assessed in a step-by-step approach, through SGs designed for adaptivity, and experimental processes.

2.1.1 Serious game layer

The first layer of the proposed framework includes the SG space. This space is composed of game content, in-game performance, and in-game metrics. Game content describes game mechanics, rules, aesthetics, in-game characters, plot, music, and everything else that can be considered a component of the SG. Some, or all, of these components can be controlled by a smart agent and be generated procedurally. The incorporated PCG technique takes up the role of this agent and controls game content. Game content can also be split in two main categories depending on its purpose, and whether it promotes the SG's objective or not. This is usually dictated by the conceptual framework and the theoretical basis that governs the SG design. This separation is not necessarily important in the application of the proposed framework; however, it facilitates the process and ensures that each procedurally generated game component is based on the intended aspects of the player model. Furthermore, this separation provides a proper design basis for ensuring that health related content is generated reliably and safely as part of the health intervention.

In-game performance refers to game progress, in-game rewards and achievements the player unlocks through interaction with the SG, as well as parameters such as number of victories and defeats and other elements that tied to specific SGs. Measurement of performance in a SG is crucial in the evaluation of its efficiency and its success in achieving the intended purpose. In terms of its involvement in the proposed conceptual framework, dynamic difficulty adjustment is usually based on elements related with performance besides affect. Maintaining appropriate game difficulty levels is proportionate to the player's ability to achieve the SG goals, as they are depicted through performance mechanics. Not all of performance elements are usually visible to the player, however, those that are can also be augmented through the adaptive content. In-game metrics refer to data extracted through the interaction of the player with the SG and is usually relevant to duration of play, interaction with particular in-game events, and other interaction data. Other examples of data can include the frequency of SG use and other patterns extracted from the application usage. In-game metrics provide an accurate description of player interaction with the SG and through the conceptual framework are employed in engagement recognition layer for the generation of player models of engagement. Besides engagement, these models can also provide information regarding player progress, learning rate and adherence to the intervention.

2.1.2 Procedural content generation layer

The PCG layer consists of the incorporated PCG technique that applies specific actions to generate game content (e.g., game mechanics, levels, NPCs, rewards, music, and story) and modify the game state. Generation of content can be translated in producing new game content or selection of appropriated content from an available pool. This layer contains all of these available actions that generate SG content and is designed to accommodate the educational and behavioral objectives regarding the targeted SG goal. These objectives include, amongst others, knowledge about the management of a condition and daily self-management goals. In addition, player engagement, recognized in real time, is provided to the PCG layer to facilitate the generation of appropriate content, in an effort to maintain the player in a state of flow and increase adherence to the intervention. To achieve both of these goals, information regarding the player’s game performance, level of engagement and degree of achieving health related goals, is fed to the PCG module in order for the latter to produce the game content that improves the intervention’s efficiency. The required information for this process is collected through the engagement recognition layer and the health related data layer. In the early stages of its operation, generation of content can also be employed to improve the model’s knowledge about the player’s preferences and needs. As mentioned in the theoretical background there is a large number of approaches and methods for PCG in games. In the present Doctoral Thesis and novel approach for PCG, based on a GA, is presented and incorporated in the SGs employed to evaluate the feasibility of the proposed conceptual framework.

2.1.3 Engagement recognition layer

The engagement recognition layer constitutes an important part of the conceptual framework since it delivers the level of identified affective engagement in real time to the PCG layer. The main objective in affective games is the modeling of various aspects of player emotion, such as engagement, stress, fear, and frustration. Multiple approaches have been attempted towards creating affective player models suitable for incorporation in games. In this approach, the use of sensing technologies that monitor physiological signals (e.g., heart rate, skin conductance, EEG signal) and identify postures and physical activity, is proposed as a means to recognise engagement in real time. Such approaches have been identified in the literature as potentially capable of discerning states of increased engagement during high cognitive learning interactions. Furthermore, data collected from these types of sensors can also be applied towards monitoring disease self-management, symptom progression, diagnosis, and adherence to doctor advice and medication in multiple chronic conditions. In addition, this layer takes into consideration in-game metrics such as mouse clicks per second and mouse idleness along with performance score and in-game decisions. This information has also been proven to be valuable towards recognising engagement. Acquisition of this type of useful information can be facilitated through the design process of the SG by incorporating specific in-game events estimated to trigger responses in engagement. The proposed conceptual framework focusses on the application of player engagement towards continuously generating game content through a real-time feedback loop. To assess the feasibility of this approach during interaction with SGs for health an approach for real-time identification of engagement has been investigated. This approach features annotated engagement in a continuous manner as perceived by the player to establish a ground truth, and investigate the predictive capabilities of selected features based on data collected from sensors and interaction with the SG.

2.1.4 Health related data layer

The health-related data layer includes advanced data analytics for the enhanced integration of heterogeneous data collected from different data sources such as sensors monitoring clinically relevant (e.g., glucose) and lifestyle (e.g., physical activity, quality of sleep) parameters, electronic health records, internet of thing devices and mobile apps. This facilitates the creation of unique and complete player's profiles in terms of health status and lifestyle. These profiles can be employed to facilitate the generation of SG content tailored to the player's needs. Data included in this layer can be both dynamic and static. The generated content can accommodate for the adaptive SG's diverse objectives in chronic disease management through appropriate game mechanics and systems. The intended empowered adherence to the intervention, facilitated by improving player engagement, allows for sustainable accomplishment of the SGs objectives. Recent advancements in sensor technology and the abundance of affordable and unobtrusive wearable devices, with the capacity to send data to the SG in real time, advocate towards the feasibility of the health related layer in the proposed conceptual framework.

2.2 A genetic algorithm approach for procedural content generation

A GA approach towards PCG in SGs for health has been designed and developed, focusing on producing content for interventions that target chronic conditions. This approach was first presented in [125] can be employed as part of the proposed conceptual framework for adaptivity in SGs for health. The proposed technique features a modular design to facilitate its application in multiple SGs of different genres. In addition, it can receive heterogeneous data as input and include player models that provide insight regarding both the affective state of engagement and health related needs. This technique includes options for dynamic difficulty adjustment after the generation of procedural content, enhancing its adaptive capabilities. Finally, it provides the opportunity to control the generated content and ensure that it is delivered safely to the player with regard to their tailored needs.

AI methods often rely upon heuristic search functions to achieve their goals. Heuristics act as an evaluator, consulting the algorithm about which step is best to make. One example of such search heuristics is the evolutionary GAs. Evolutionary algorithms consist of a broad group of optimization and search algorithms that are based on the principle of biological evolution [139]. GAs, along with classifier systems and genetic programming are part of this group. Drawing inspiration from Darwin's theory of evolution, GAs constitute search methods based on the principles of natural selection and genetics [138]. GAs have been used in PCG techniques due to their ability to produce highly customized content for a game, by "evolving" it according to the progress of the user. In order to achieve these goals, SGs make use of heterogeneous and complex content and mechanics, which benefit greatly from PCG techniques [139].

The incorporation of GA techniques for PCG in games is widely documented in the relevant literature. GAs have been employed in game level generation [140], with the ability to control aspects such as difficulty and aesthetics, as well as employing rhythm for generating game level geometries [141]. Procedurally generating levels and stages in games appears to be a customary field of application for GAs, however some approaches go further, also altering game physics and rules by incorporating them in fitness functions [142]. Besides level, GAs have also been applied to the generation of games maps in strategy games, taking into consideration the allocation of resources and confrontations between opponents [143]. Furthermore, GAs have been used to procedurally generate game quests and missions, through an evolutionary search strategy, guided story arcs [144]. GAs have also been employed towards dynamic difficulty adjustment techniques by controlling game characteristics through fitness functions [145].

Attempts to employ GAs for PCG in SGs have been reported in research articles. GAs present certain advantages in their application to real-time PCG, as they are able to search for optimal solutions by updating weights included in their fitness functions during live play. This form of learning is then applied when selecting the fittest individuals that will produce the next generation and control game content. The time and manner of the production of a new generation can be dictated according to the specific needs presented by the SG. In addition, the content of chromosomes and genes can be formed to accommodate for diverse SG content in a modular and generalized manner, allowing for the application of such approaches in a broad selection of SGs. A data driven PCG approach based on a GA and support vector machines to recognize player capabilities automatically, that can be generalized in various SGs was presented in [146]. This approach proved to enhance player performance in comparison to a non-adaptive version of the SG, but also displayed superiority in producing game content against a simple heuristic approach. Procedural level generation was conducted through the combination of bidirectional long short-term memory and fuzzy analytic hierarchy process - GA based on player performance and preferences in an educational game in [147], with results from comparing the method with other learning methods advocating toward the superiority of the combined approach. Furthermore, GAs have employed in digital educational interventions that do not employ game mechanics, in order to generate courses and material in a dynamic manner that fits the student's needs [148].

A PCG framework (Fig. 2-2) has been designed for incorporation in SGs for health, aiming to improve the SG's intervention capabilities and enhance overall game experience. The proposed technique is based on a GA and is responsible for the automated generation and selection of game content. This selection can be guided by various aspects of player modeling, built upon dynamic data, such as player interaction, game progress, and sensor data, as well as static data, such as demographics, questionnaire responses, and medical files. By translating SG content into chromosomes of the GA, the evolutionary algorithm searches for the appropriate game content. The GA's weights are updated based on the player model and their specific health needs and the fitness functions select the chromosomes that will form the next generation. Depending on the characteristics of the SG, game difficulty of the generated content is adjusted automatically. The resulting adaptive SG possesses the ability to automatically adjust game difficulty and present content tailored to the user's health related needs. This approach mainly focusses on SGs that aim to empower disease self-management in chronic conditions, through mobile and electronic health interventions. These types of SGs benefit greatly from the deployment of sensors to monitor health-related behaviors and habits relevant to disease progression. However, it has the potential to be generalized in SGs for health, or otherwise, with various objectives and goals.

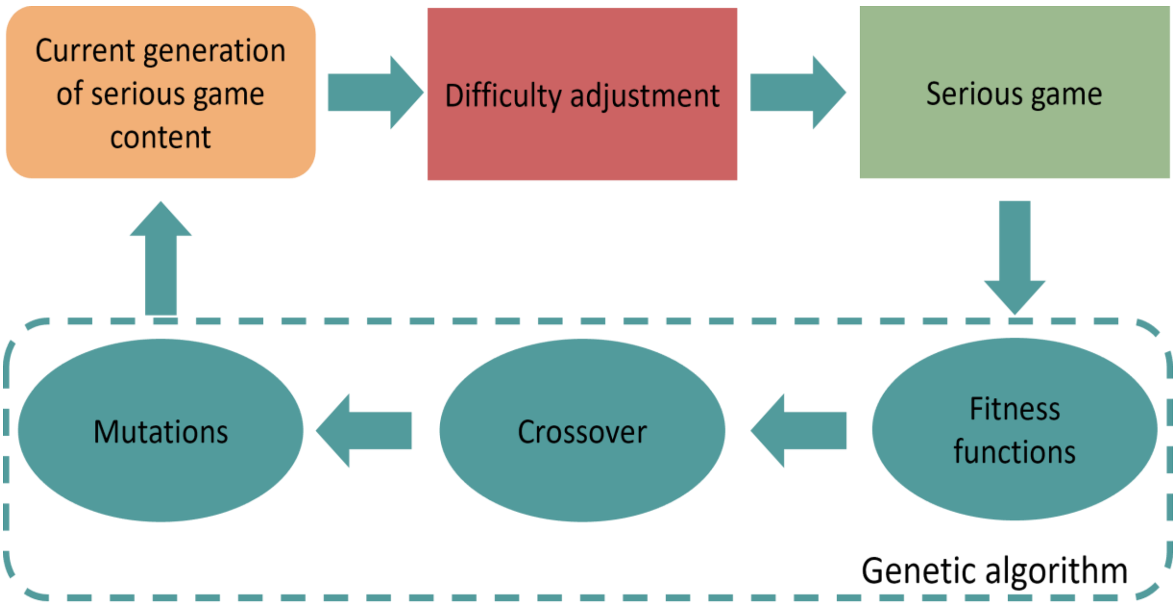


Figure 2-2: Framework based on a genetic algorithm for procedural content generation in serious games for health.

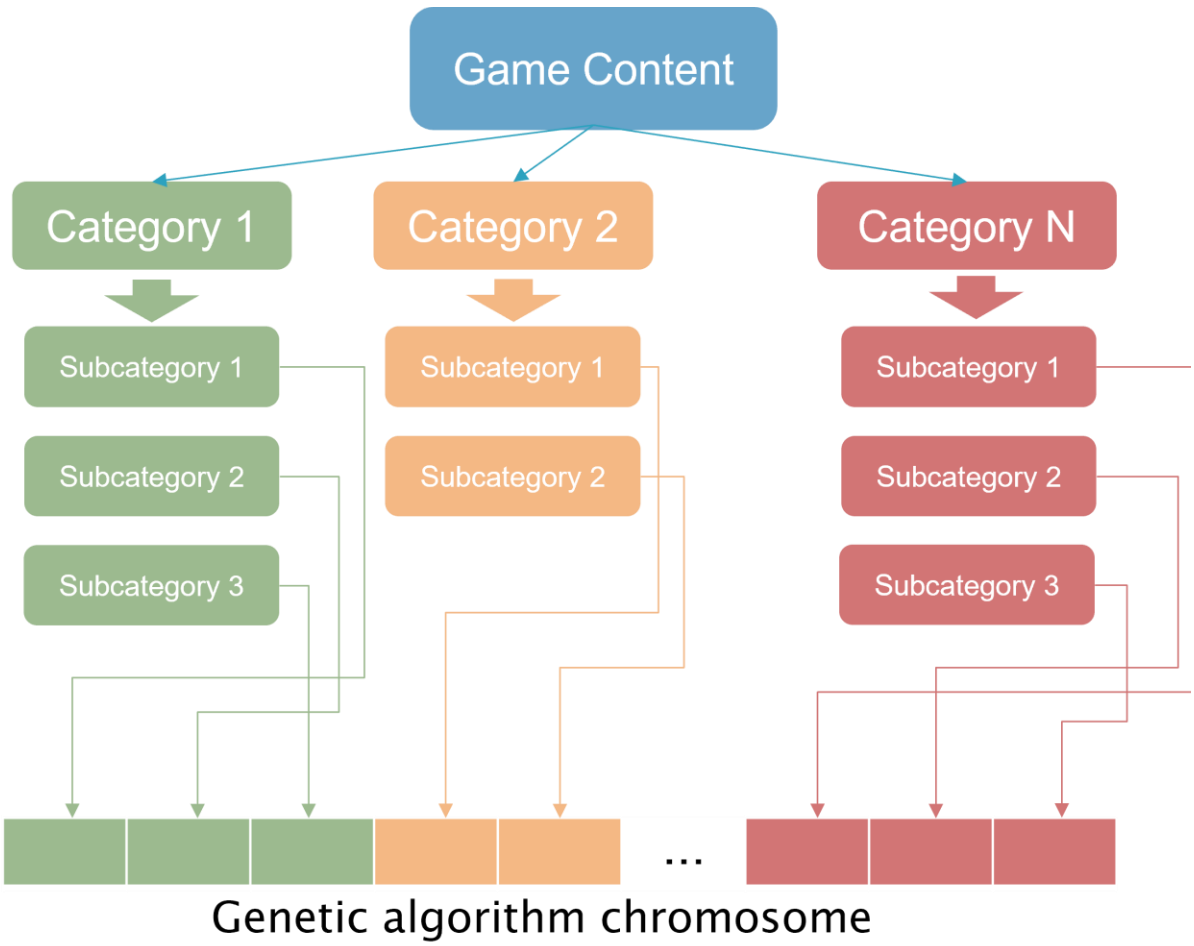


Figure 2-3: Translation of serious game content into the genetic algorithm's chromosome structure.

In order to effectively communicate game content to the GA controlling it, a novel approach of representing the adaptive aspects of the SG in elements of the algorithm was implemented. The proposed approach for PCG treats SG content and assets as a string of genes that form a chromosome (Fig. 2-3). Depending on the type of SG that incorporates the proposed PCG framework, different types of game content must be represented in genes. This approach is highly modular and can be generalized towards various types of game content. An example of the types of content it can accommodate, explained in detail in Chapter 5, can be found in the incorporation of the proposed technique in a SG for OSA. In this SG NPCs are described through attributes that are relevant to the intervention's educational and behavioral objectives. Each of these attributes is represented as a gene in the GA's chromosome. This way, in this example, a chromosome can be linked to an NPC by including the number and types of attributes that describe them. The genes take binary values, with "1" representing game elements that exist during the current iteration of the GA, while "0" signifies their absence. In the example mentioned above, the whole population of chromosomes in each generation describes potential NPCs that could exist in the SG. In each generation of the GA one or more of the fittest chromosomes are selected to determine the generated SG content according to their gene values. In the provided example, new NPCs would be generated and populate the SG as opponents for the player.

Weights are assigned to each gene and trained based on the data collected through various means, such as interaction with the SG and sensors. This data provides insight regarding player needs according to a player model. The weights can be updated whenever it fits the generation of new content, according to the needs of a specific game. Based on the updated weights, fitness functions calculate the fittest chromosomes for each generation. These chromosomes are going to be used to create the next generation, while mutations take place to decrease saturation in the produced SG content. The nature of the weight update and the fitness functions depend on the application in a specific game environment. A rule-based system then filters the generated content to dynamically adjust difficulty level according to player progress. The architecture of the proposed PCG technique is presented here in an abstract manner. Its generalization capabilities have been tested in the present Doctoral Thesis and two detailed examples from its application are provided in Chapter 5. The aim of this technique is to employ a multitude of heterogeneous data in the procedural generation of game content. By splitting SG content in different genes and employing weights that are updated in a different manner for each gene, this method allows for efficient control of the generated content by selected aspects of the player model. This way, modalities that can be applied towards enhancing engagement can work independently from modalities that generated personalized health related content. The proposed PCG technique has been incorporated into SGs for health that were used to evaluate the proposed conceptual framework for adaptivity. Details regarding the incorporation of the GA PCG technique in specific SG spaces are going to be described in Chapter 5.

3. Design and development of adaptive serious games for health

In this chapter the design and development process for two SGs for health that were created during the course of the present Doctoral Thesis are presented. The first SG was first presented in [149] and aims to promote food and nutrition literacy among adolescents and young adults and has the capacity to empower self-health management in chronic conditions with special dietary needs. The second SG was first presented in [125] and aims to raise awareness about OSA and induce healthy behavioral change in people suffering from it. In addition, a third SG that was employed in pilot studies with children suffering from obesity and type 1 diabetes mellitus, aiming to promote self-health management, and incorporating the proposed PCG technique is described. Preliminary results from a pre-pilot study and the description of this SG were first presented in [150]. Data collected during play with the presented SGs has been employed to investigate the feasibility of the proposed conceptual framework for adaptivity in SGs for health. All three games incorporate mechanics that facilitate PCG of their "serious" content. The design process of the SGs has been based on theoretical frameworks, and in the case of the third SG in collaboration with health experts.

3.1 A serious game for food and nutrition literacy

Express cooking train (ECT) [149] is a SG designed and developed to promote food and nutrition literacy skills in adolescents and young adults, by building knowledge regarding food preparation, portion sizes, healthy recipes, and nutritional values. Malnutrition is currently one of the greatest challenges faced by our generation, considered responsible for more adult deaths and disability than smoking habits and alcohol consumption [151]. The onset of multiple chronic conditions such as cardiovascular diseases and diabetes has been associated with unhealthy dietary habits and obesity [152]. Limited time spent in cooking along with insufficient skills hampers the ability of individuals to become healthier by assuming proper dietary habits [153]. Developing knowledge related to nutrition and cooking, maintaining healthy dietary behavior and achieving normal weight constitute important milestones towards health promotion and chronic disease prevention and management [130]. Food and nutrition literacy constitute two key concepts that are becoming increasingly important in the promotion of a healthy dietary lifestyle [154]. The ability to understand information relating to nutritional values, as well as the capacity to make proper dietary decisions is defined as nutrition literacy [155]. The ability to apply this knowledge during food preparation constitutes the concept of food literacy [156], [157]. It has been proven that effective dietary changes are facilitated and empowered by the development and enhancement of both nutrition and food literacy. To this end, ECT's game mechanics are linked to mediators aiming to facilitate behaviour change and knowledge acquisition, in an entertaining way.

The SG design is based on a conceptual framework that incorporates a recipe ontology and offers an immersive and engaging cooking simulation system, suitable for exploration through trial and error. ECT was developed using Game Maker Studio engine for Windows operating systems. The conceptual framework for the design of ECT is presented in Figure 3-1 and incorporates behavioral science theory concepts [158], along with features from a generalized SG framework [133]. The conceptual framework is split into three sections, representing the SG content (1), the mediators (2), and the desired outcome (3). All the elements of the SG content are included in ECT to facilitate and support the mediators, which intend to promote the SG's desired outcomes while entertaining the user. The desired outcome of ECT is not only knowledge acquisition, but also effective and sustainable behavioral change. The conceptual framework describes the way elements from the SG are employed as mediators to reach this outcome.

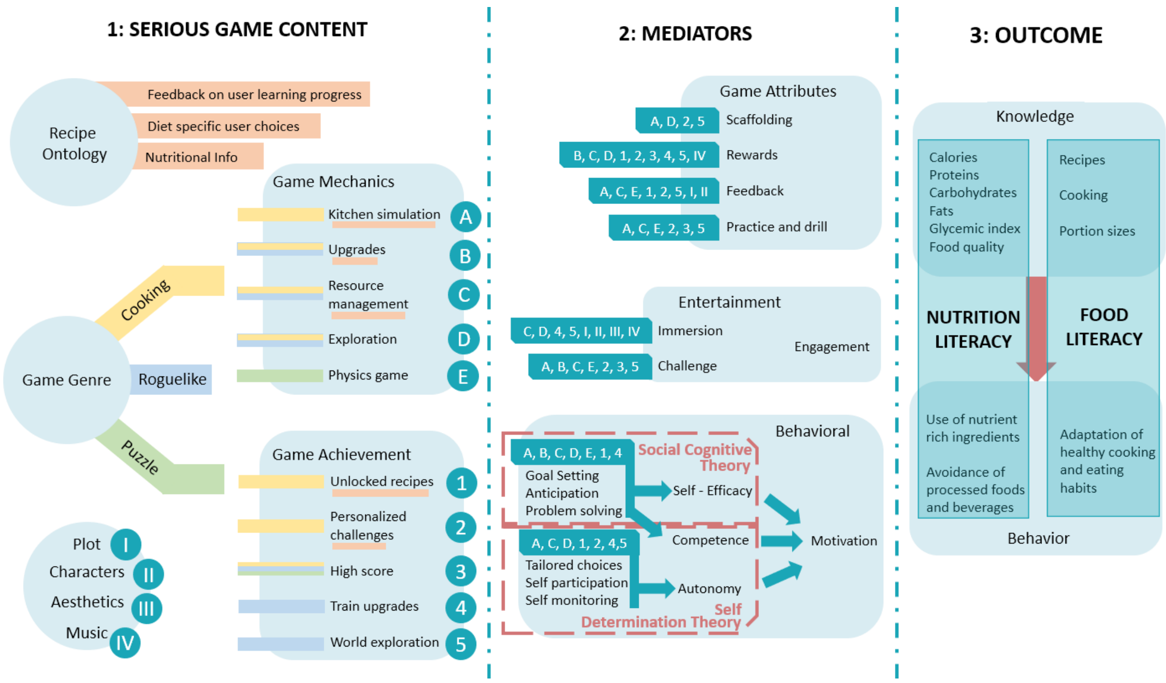
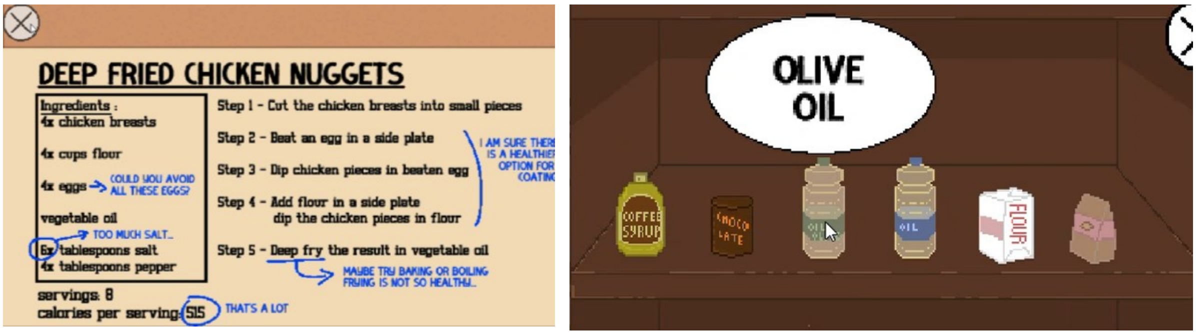


Figure 3-1: Express cooking train conceptual framework.

a) Recipe ontology

The ability to represent and handle information regarding food recipes (Fig 3-2a), ingredients (Fig 3-2b) and meal preparation is essential to the kitchen simulation mechanics employed by ECT. A realistic and fun simulation of meal preparation is critical to the educational and behavioral goal of the SG. To this end, ontological modeling has been selected to formally represent all the related concepts, features, and constraints linked to the kitchen simulation system. To this end, an extended model has been built on an existing ontology by BBC [159], incorporating knowledge representation fitted to the SG's needs, while ensuring extensibility, sustainability, and reusability. There are three main reasons for this approach in ECT. Firstly, this type of knowledge representation facilitates the handling of complex game concepts that will support a realistic kitchen simulation system in the SG. Secondly the ontology provides an easy way of expanding the described knowledge in a sustainable manner by including recipes and dietary options specific to tailored needs of self-management in various chronic diseases. Thirdly, ontological modeling facilitates the representation of game content, in terms of recipes and ingredients, in a way that can be employed by a PCG technique. The employed ontology includes food information relevant to game objectives and rewards, such as nutrition values and cooking recommendations. Additionally, representation regarding information collected from the interaction with the player has been included, to facilitate player modeling and delivery of personalized game content. This information contains in-game performance relevant to specific recipes or cooking habits monitored throughout the game. Finally, the SG ontology contributes to the calculation of game score and progress. Ontology concepts are employed by a rule-based system to evaluate the nutritional value of prepared meals within the kitchen simulation, as well as their similarity to in-game recipes. The nutritional value of each prepared meal is calculated by factoring the selected cooking technique, the amount of fats, calories, carbohydrates, proteins, and the glycemic index of the used ingredients.



(a) (b)

Figure 3-2: Express cooking train example recipe (a) and sample ingredients (b)

b) Game genre

Genre in digital games refers to an informal classification that is based mostly on playstyle elements of the game. Various attempts to produce a formal classification of game genres have been identified [160], however genres appear to be evolving continuously and organically with the development of new innovative games that combine multiple genres or introduce new ones. ECT consists of playstyles deriving from three different genres: cooking games; roguelike games; puzzle games.

At its core, ECT is a cooking game. Cooking games, as the name suggests, simulate the cooking process using various mechanics in a vast variety of manners. In ECT, an attempt to incorporate a realistic and fully explorable kitchen simulation environment (Fig. 3-3) was made. The player has a number of recipes and ingredients at their disposal, as well as a virtual kitchen equipped with appliances and various tools. Player interaction is conducted through the mouse by pointing and clicking. Every time the player selects an ingredient to use with a kitchen appliance a number of options appear, describing ways that the ingredient can be prepared by the particular appliance. The preparation time is represented by an in-game timer. Tips and advice appear if the player makes a mistake, loses an ingredient or makes an unhealthy decision. By employing the ontology model, the kitchen simulation mechanics of ECT provide an environment suitable for exploration, as the player is motivated to try new combinations of ingredients and decide whether they want to follow a recipe or try to discover new ones. In each playthrough a set of recipes is selected along with a few starting ingredients, provided in the ontology. According to the recipes provided and the available ingredients the game difficulty can be adjusted.

Figure 3-3: Express cooking train virtual kitchen environment.

Roguelike games are characterized by procedural generation of content and randomization mechanics. In this type of games, player progress is typically ephemeral in nature, as the player starts from the beginning with each playthrough. This playstyle encourages the player to explore the game, as in order to progress further they have to improve their skills and knowledge of the game by learning from their mistakes. In ECT if the player loses, they must begin a new journey, with new starting ingredients and recipes. In every new train trip, the stage's environment is generated procedurally along with the maximum travel time length. Rule based systems that control visual assets control stage generation, weather elements and day/night cycle. Some upgrades are retained between playthroughs, facilitating game progress and reducing repetition. In addition, between train travels, the player is able to make some selections that are going to affect the next travel in terms of rewards and achievements. Finally, puzzle games usually present the players with challenges that require strategic planning and problem solving. In ECT the presence of puzzle is quite subtle, aiming to break the pace of the cooking simulation and reduce its repetitiveness. A small number of physics games, point and click mechanics, as well as hidden plots and dialogues exist, to intrigue the player but also promote exploration. Initial design included a catapult mini game that was not included during ECT development as there were concerns that it could overshadow the "serious" purpose of the game.

c) Plot and characters

The game is played in missions, with the player taking control of a train (Fig. 3-4) trying to reach the next destination carrying valuable cargo. ECT is set in a post-apocalyptic setting and during these short trips, huge monsters chase the train. The ground level is a barren wasteland and towns are kept on the air by huge balloons. At the beginning of each trip the train is descended to the ground level and placed on tracks. The player must travel to the next town where air balloons await to lift the train to relative safety. The monsters roaming the land are extremely hungry but also very picky about what they want to eat. When they spot a train, they begin to hunt it, having grown accustomed to train drivers being some of the best cooks of healthy food available. The player, a rookie train driver, must cook healthy meals in the kitchen wagon and launch them with a catapult towards the monsters to satisfy their appetite and reach the next train station safely. Adding to the game difficulty, the train inventory contains mostly junk food recipes, forcing the player to explore ways of including healthier ingredients and applying better meal preparation techniques, as the monsters grow angry if junk food is thrown their way. The train engineer, Suzanna, is accompanying the player along their train trips offering advice in cooking and managing the train. The monsters, also accompanying the player in a way, have some interaction options that were initially designed to produce comic relief, with witty dialogues and comments. Eventually these dialogues were redesigned to also provide small pieces of plot that would eventually lead the player to understand what happened to the world and in this way provide players with incentive to explore and play. Additional characters, plot events and environments have been conceptualized but not yet implemented in the SG.

Figure 3-4: Express cooking train.

d) Game mechanics

Several game mechanics have been incorporated in ECT: (1) kitchen simulation, (2) upgrades, (3) resource management, (4) exploration, (5) physics game.

1) Kitchen simulation: A game environment providing a realistic game space to prepare meals. Cooking appliances and equipment along with a book of recipes and an inventory for ingredients are at the player's disposal to experiment without limitations. A variety of interface interactions are available, such as drag and drop, swipe, point and click, providing a user-friendly interface. Each generated meal is awarded with two scores, provided by the recipe ontology. As the game progresses the monsters become harder to satisfy, demanding healthier food options that are more difficult to produce.

2) Upgrades: An upgrade system allows users to make permanent improvements in the kitchen and the train. The ability to progress to later stages of the SG depends on the incorporated upgrades. This mechanism offers decisions for strategic planning and critical thinking and is directly involved with the process of better meal preparation in the kitchen simulation environment, as different cooking appliances and techniques are explored by the user. This mechanic has been designed but not yet implemented in the SG.

3) Resource management: Resource management is essential to game progress due to the finite nature of the ingredients. Recipe planning and ingredient procurement form an additional decision system promoting the acquisition of dietary and cooking knowledge and the empowerment of cooking skills. This mechanism offers self-participatory problem-solving interaction of a highly immersive nature.

4) Exploration: Exploration of the SG world provides an opportunity for the player to interact with NPCs, monsters and game locations, as well as experience other creative facets such as plot, aesthetics, and music.

5) Physics game: A physics puzzle mini game is employed every time the user launches a meal to the chasing monster with the catapult on the last train wagon. The difficulty of this mini puzzle game relies heavily on the similarity of the produced meal against reference "junk" food. Meals of higher nutritional value automatically reach the monster, encouraging the user to make healthier cooking choices. This mechanic has been designed but not yet implemented in the SG as it could shadow the core game mechanics.

Figure 3-5: Reward, progress and achievement screen of Express Cooking Train.

e) Game achievement

Game achievement (Fig. 3-5) offers ways to monitor player progress in the SG, as well as their understanding of the educational content and ability to perform better in game challenges. These mechanisms reward the player and provide an opportunity to evaluate the learning process.
1) Recipe unlocks: Through successful exploration in meal preparation, the user improves the existing recipes and discovers new ones. Furthermore, new recipes can be acquired through world exploration.
2) Personalized challenges: Based on the player's cooking profile, tailored challenges are generated to address identified unhealthy habits and weaknesses. This system allows for PCG based on player modeling; however, it is not implemented in the current version of ECT.
3) High Score: The high score is measured by allocating similarity and healthy points to food produced by the player. The scoring system is based on the incorporated ontology and is presented to the player during train trips and on the reward screen.
4) Train upgrades: Train upgrades make the train faster and sturdier. Besides beneficial upgrades there are also visual upgrades that allow the user to customize the train.
5) World exploration: The progress of world exploration is depicted on a review screen that shows the percentage of SG completion. It displays the number of recipes the player has discovered, the achievements, as well as budges awarded for exhibiting healthy habits. With the implementation of additional characters, locations, and plot points, they will also be included in this screen. This mechanic has been designed but not yet implemented in the SG.

f) Behavioral mediators

Behavioral mediators in ECT are based on the social cognitive theory and the self-determination theory [161]. According to social cognitive theory, motivation is empowered by the expectation of positive results in given course of action. This is achieved through the provision of a cooking simulation environment allowing the player to experiment with food preparation. This environment allows for a safe trial and error approach with short-term rewards. Moreover, goal setting, review and feedback are involved in game mechanisms in order to empower self-efficacy, which is a crucial motivational mechanism [48]. Self-determination theory dictates that autonomous motivation contributes to a sustainable change in dietary habits as opposed to controlled motivation. Autonomous motivation refers to the user's realization of the beneficial outcome of the intervention provided by the SG. ECT offers a self-participatory and personalized environment that enhances autonomy and competency which ultimately lead to self-determination [161].

3.2 A serious game for obstructive sleep apnea

Wake up for the Future (WuF) [125] is a SG aiming to promote disease self-management in adults suffering from OSA and raise awareness about the condition. OSA is currently the most prevalent sleep-related disordered breathing condition [162]. OSA manifests with recurrent episodes of upper airway collapse that result in a decline, or even interruption, of airflow with a duration of at least ten seconds. A variety of treatment options exist, however, proper self-management of the disease can also benefit patients suffering from OSA (e.g. healthy diet options, weight loss, proper sleeping routines and positions, limitation of alcohol and tobacco consumption) [162]. The condition is considered a major public issue, accounting for 936 million patients worldwide in 2019 [163], while most (80%) of the cases remain undiagnosed. If left untreated, OSA can increase the risk of health problems such as cardiovascular diseases, diabetes, and depression. Only 20% of OSA cases globally are estimated to have been diagnosed, indicating that raising awareness about OSA's symptoms can help reduce under-diagnosis and improve disease outcome.

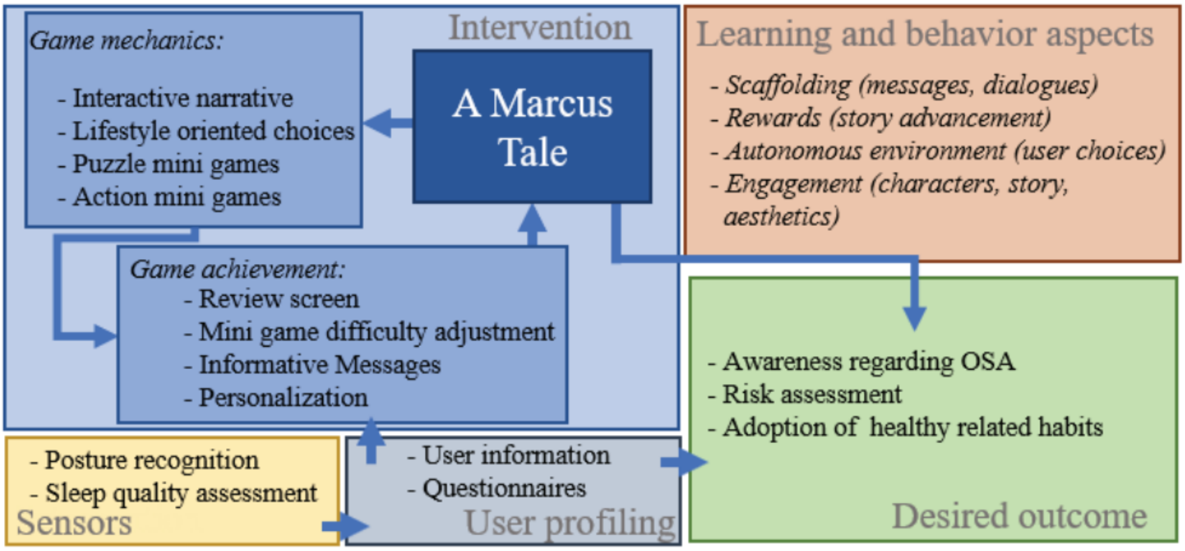


Figure 3-6: Conceptual framework for the design of A Marcus Tale [164].

Initially, another approach based on the conceptual framework used for ECT was attempted to create a SG for OSA. An narrative adventure SG, A Marcus Tale (AMT), was developed using the framework displayed in Fig. 3-6 [164]. AMT was implemented on the Unity platform to run on windows operating systems. Its conceptual framework links the game mechanics with raising awareness and behavioral motivation, targeting adults. In terms of game mechanics, it features a compelling plot and hand drawn aesthetics to engage the user and provide an immersive experience. The main character is Marcus, a former robber, that has been recently diagnosed with OSA. The user makes decisions that affect Marcus's health status and consequently the advancement of the story, offering the user a personalized experience. Throughout the game the user receives messages for healthy lifestyle while gaining knowledge about OSA. The user's ability to make correct behavioral decisions minimize the difficulty of puzzle and action mini games. Furthermore, periodical review screens of the user's decisions and the influence on Marcus's health are presented.

Based on insights gained from the design and development process of AMT the conceptual framework was revised to accommodate for PCG techniques. Mechanics and systems from digital card games were incorporated in the design and the desired outcome shifted away from OSA related risk assessment. This new approach aimed to raise awareness and promote self-health management, as these targets were deemed more fitting to the design approach. Monitoring disease and symptom progression in OSA using electronic and mobile health technologies is a subject that has been investigated extensively in the literature [165], [166], [167]. However, no SGs for OSA have been identified and a step-by-step approach was decided regarding the capabilities of WuF. Systems and mechanics capable of monitoring can be implemented and the future, along with options for medical decision support, however it is crucial to investigate the capabilities of the SG in terms of behavioral and educational outcomes before proceeding further. The resulting conceptual framework for WuF is presented in Fig. 3-7. Its design links the game mechanics with raising awareness and behavioral motivation, while targeting the adult population. As in the case of ECT, the framework incorporates behavioral science theory concepts [158], along with features from a generalized SG framework [133]. The SG features a novel and suspenseful plot with the user traveling from a dystopian future where all knowledge regarding OSA is lost, to the present. WuF is an open world digital game, featuring debate battles with the use of card decks. It has been implemented on the Unity platform and runs on Windows operating systems.

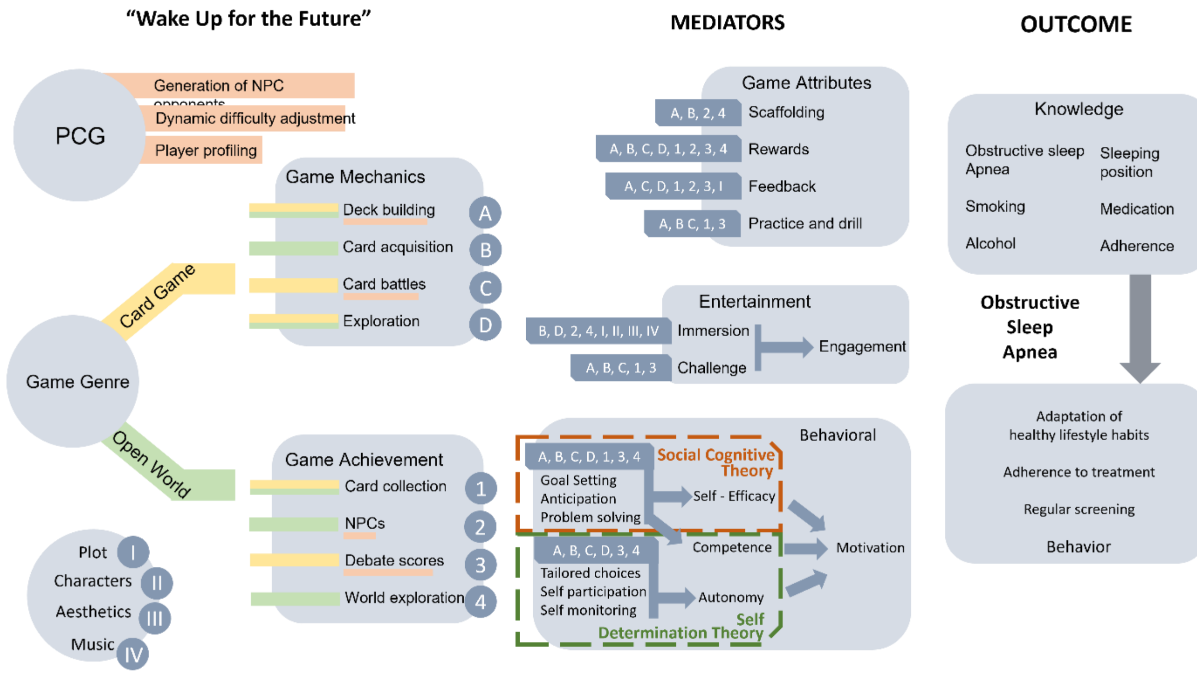


Figure 3-7: Conceptual framework for Wake Up for the Future.

The conceptual framework describes the way elements from the SG are employed as mediators to reach this outcome. In the following subsections, elements depicted in ECT's conceptual framework are explained in detail.

a) Procedural content generation

WuF incorporates mechanics that allow for procedural generation of NPCs and their short biographies. This approach aims to empower the SG's educational value, dynamically adjust game difficulty and empower player engagement (Fig. 3-8). The implemented system is modular and get employ a variety of techniques and methods. Every time the player enters a card battle, an NPC opponent is generated. According to the attributes of the generated NPC the player must apply their knowledge regarding OSA in order to defeat them in a simulated debate. The proposed PCG technique based on a GA, presented in Chapter 2, has been incorporated in WuF. In addition, a randomization rule-based system has been implemented to serve as a benchmark for the validation of the proposed PCG method. Details for the automated generation of game content in the form of NPCs in WuF are presented in Chapter 5.

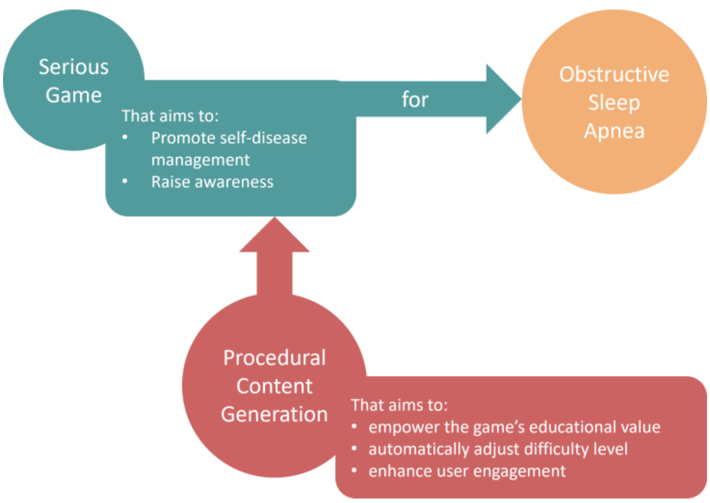


Figure 3-8: Scope of Wake Up for the Future.

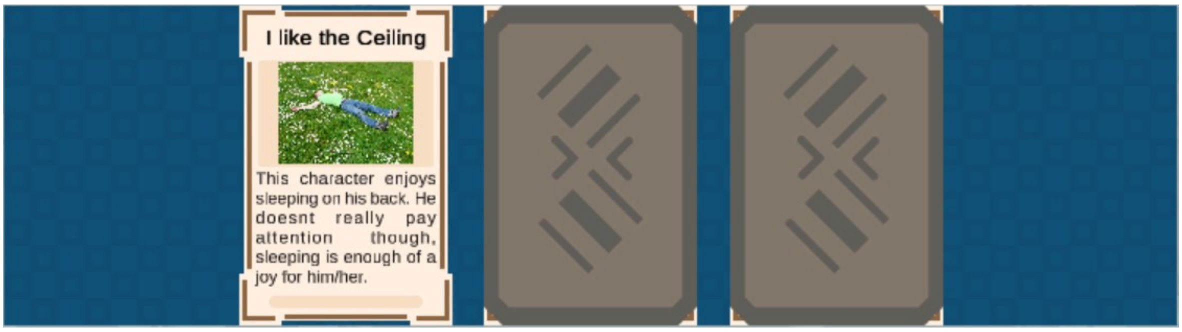


Figure 3-9: Cards in Wake Up for the Future.

b) Game genre

WuF borrows ideas and mechanics from two broad video game genres, card games and open exploration games. Card games have been frequently employed as educational and behavioral tools in the form of SGs in various subjects and application fields [168], [169], [170], [171]. These SGs usually adapt social and constructivist views of learning, such as the ones incorporated in the employed conceptual framework, and have been evaluated with positive results towards learning, attitude change and thinking skills. These SGs often include cards with educational content, the concept of formulating an appropriate deck to win, and the gradual collection of cards that are often increasingly strong in terms of how they affect the card game. In Wuf, the player collects cards (Fig. 3-9) that represent arguments about the prevalence of OSA, healthy lifestyle habits that are associated with it, and tips regarding adherence to proposed treatments and medications. Through this card collection they form decks to engage in card battles facing NPC cards with false arguments and perceptions. In addition, the SG features an open world where the player can explore and talk with NPCs, collecting more cards and learning about OSA in the process. This open world is presented through top-down classic pixel graphics (Fig. 3-10). The open world approach induces and sense of exploration and discovery and supports the game story while providing an immersive incentive for the player. The NPCs include opponents for the card battles, but also helpful characters that present the player with knowledge about OSA in the form of new cards.

Figure 3-10: Open world of Wake Up for the Future.

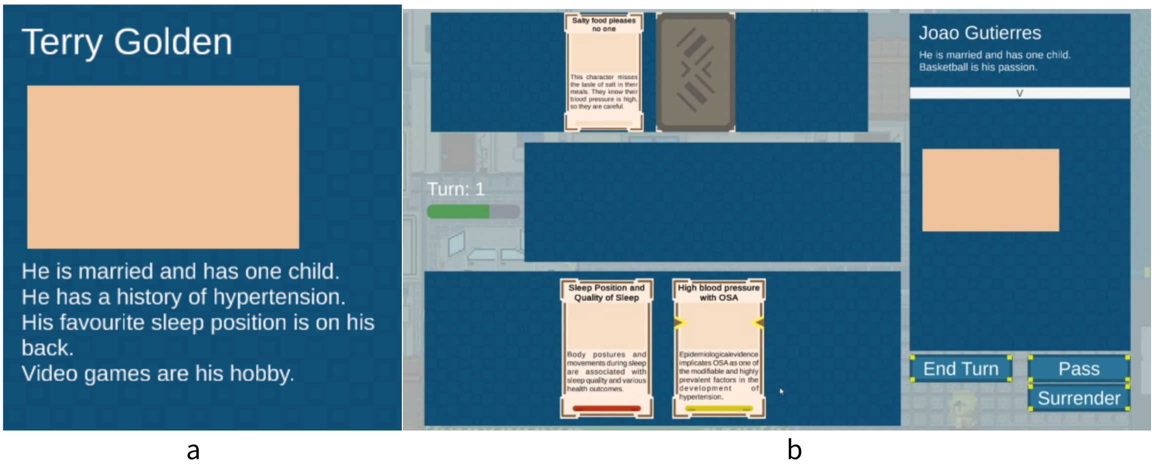


Figure 3-11: Non-player character biographies in Wake Up for the Future.

c) Plot and characters

WuF features a compelling story about time travel from a dystopian future to the present. In the future all knowledge regarding OSA has been erased by a corporation that controls the populace suffering from OSA. In this alternate reality the condition is affecting almost the entirety of the population and the evil corporation has found a way to exploit this. The main character is a person from this future that travels back in time to the world of today in an effort to collect knowledge about OSA and ways that can improve self-health management and treat the condition. With this knowledge the character intends to return to the future and save the world from the corporation by spreading knowledge and helping people suffering from OSA to manage their condition. With their arrival in the present the main character quickly finds out that OSA is under-diagnosed and widely spread. In an effort to learn and help people they meet and talk with people suffering from OSA in an effort to raise awareness and convince them about the seriousness of the condition and the need to manage it through healthy behavioral change. These debates are simulated by the SG's card mechanics.

d) Game mechanics

During the game, the user participates in debates with NPCs. These NPCs describe people with undiagnosed cases of OSA, with no awareness of their condition and without the necessary knowledge skills. The user's purpose in the game is to convince these NPCs and contradict their false beliefs and unhealthy habits. In this way, the user takes up the role of a mentor, trying to raise awareness and promote self-disease management for OSA. The debates are incorporated in the SG with a card game, where each card represents an argument, which is associated with an attribute. Attributes are habits (smoking, alcohol consumption, sleeping position, etc.) and chronic conditions (obesity, hypertension, diabetes, etc.) that are linked to the onset, diagnosis, and progress of OSA. Every NPC is characterized by some of the above attributes, creating a profile, which is presented to the user via a short biography before each card battle. Each of these attributes is linked to a false argument the NPC uses during the debate, which strengthens their lack of motivation towards healthy lifestyle change. The short biography (Fig. 3-11a) provides the user with some insight regarding the upcoming debate battle, enabling him to prepare his arguments properly, by selecting cards for their deck.

The debates are simulated by a card game system (Fig. 3-11b). There are two types of cards, NPC and player cards:

- NPC cards represent their false arguments, each linked to an attribute. Every NPC possesses one NPC card for each attribute in their biography.

- Player cards represent correct arguments about OSA, each linked to an attribute. Before every debate, the user must create a deck of five player cards. There is a pool of available cards, featuring two user cards per attribute.

The card game is then played in rounds. The player's maximum hand size is two cards. Before the start of the first round, the player can reshuffle their starting hand and draw a new one. At the start of every subsequent round, the player draws one new card, always respecting the maximum hand size limit. The NPC cards start the debate on the board facing down. If no NPC card is open at the beginning of every round, one is flipped open.
On every round the player has the following available actions:

1) Play up to two cards from their hand. A player card destroys a face up NPC card of the corresponding attribute. If no NPC card is facing up, then the player card reveals a face down NPC card of the corresponding attribute. If no appropriate NPC card exists, the player card has no effect. A played card is then discarded.
2) Pass the turn. In the player selects to do so, they have the option to reshuffle and draw up to two cards from their existing deck.
3) Surrender. The debate ends with a losing resolution.

The player wins the debate if they destroy all of the NPC cards within five rounds.

e) Game achievement

Game achievement, as mentioned in the case of ECT, offers ways to monitor the user's mastery of the game, their understanding of the educational content and ability to apply it on in-game challenges. These mechanisms deliver the pleasure of reward and provide a learning assessment opportunity.
1) Card collection: Through interaction with the game world, the player accumulates an increasing number of cards that represented compelling arguments that can be used in the debates. As the game progresses the player unlocks more powerful cards that are able to contradict more than one type of false arguments. This process provides the player with a sense of goal and fulfilment as their collection reaches completement.
2) NPCs: As the player meets NPCs in the open world, they unlock new dialogue options based on their progress in the SG. These dialogues further the knowledge of the player towards OSA, but also introduce narrative advancement and reward the player for their successes. This feature was not incorporated in the preliminary version of the SG.
3) Debate scores: Depending on the success of their deck formation and the ability to handle the cards they are dealt each turn, the player is able to complete debate battles faster and with less argument cards. This improves their understanding of the educational content and rewards them as they are able to defeat more difficult opponents produced by the PCG approach.
4) World exploration: The progress of world exploration is measured by the number of NPCs the player has managed to convince in the simulated debates and the number of NPCs they have befriended and acquired cards from.

f) Behavioral mediators

Behavioral mediators in WuF are based on the social cognitive theory and the self-determination theory [161]. Motivation, according to the social cognitive theory, is enhanced through the expectation of positive results in given courses of action. This is achieved through the provision of progress in simulated debates, with their difficulty governed by the PCG technique. In this way a smooth game progress that empowers motivation and engagement is achieved. The collection of new cards, formation of new decks and debate scores, along with the progress observed in the SG world, provides goal setting, review and feedback in order to empower self-efficacy, which is a crucial motivational mechanism [48]. Autonomous motivation can contribute, according to the self-determination theory, to a sustainable change in habits related to the onset and progress of OSA. Autonomous motivation is driven by the player's realization of the beneficial outcome of assuming the habits presented by WuF through card content and card mechanics.

3.3 A serious game for type 1 diabetes and obesity

The third SG that was employed towards the validation of the proposed conceptual framework is part of the ENDORSE platform. The ENDORSE project introduces a novel integrated platform to empower self-health management in type 1 diabetes mellitus (T1DM) and childhood obesity [172]. T1DM is a chronic metabolic disease characterized by elevated blood glucose levels. T1DM results from the autoimmune destruction of pancreatic beta cells, which causes the lack of insulin production. If not treated properly, the disease has serious short-term complications such as hypo- and hyperglycemic episodes and severe long-term complications such as micro-and macro-vascular diseases. According to the diabetes control and complications trial [173], diabetes complications can be prevented through intensive glycemic control. Obesity is an increasing public health issue as well. Obesity is associated with a range of comorbidities, such as, cardiovascular disease, OSA, diabetes, and cancer [174]. The platform is designed for children within the range of 6 to 14 years and their parents. The platform consists of a mobile SG for children, a mobile application for parents and one for healthcare professionals, enabling remote health monitoring. The platform leverages data collected from multiple sources, including sensors (i.e., physical activity sensors, continuous glucose measurement sensors), internet of things devices (i.e., smart insulin pens), as well as data collected during interaction with the platform. The platform is capable of delivering personalized content through the incorporation of a recommendation system. Content generated by the recommendation system is made up of two parts: missions for the SG and messages for the SG and the mobile applications.

The SG consists of several mini games that form educational and action game missions. Healthcare professionals have been involved in the SG design process, in order to ensure the validity of the game content and the efficiency of the SG intervention. Through the educational and action missions, the player collects in-game currency and food ingredients. The currency can be spent to customize their avatar, whereas food ingredients can be used in a meal preparation mini game to collect further rewards. Additionally, educational and progress messages appear daily in a predefined game space. The messages selected by the recommendation system come from a pool generated by the healthcare professionals. Educational messages include advice and tips about healthy lifestyle and disease self-management. Progress messages provide positive reinforcement based on the progress monitored by the platform's sensors. In order to minimize the SG's contribution in total daily screen time, only two game missions, an educational and an action, are available on a daily basis.

In the pre-pilot study of the platform with children suffering from obesity, four missions were included in the SG. "Cross the Swords" and "Dive and Rise" were the names of the educational missions and "Fruit Ninja" and "Balance Beam" the names of the action missions. In "Dive and Rise", the player collects healthy food ingredients and avoids unhealthy while swimming in a river. In "Fruit Ninja" the player to slices different kinds of food that follow random parabolic paths across the screen. With more slices, a bar fills gradually. The mission is successful if the sliced food ingredients are healthy. In "Balance Beam" the player has to cross a narrow bridge while maintaining balance by using a paddle. A random food appears on the side of the paddle and the player has to adjust its quantity to achieve a reach portion. If the right portions are chosen, the avatar maintains its balance and crosses the bridge. "Cross the Swords" places the avatar in a contest of reflexes. The player has to react quickly and pierce different kinds of healthy food ingredients. Across the SG, food classification in healthy and unhealthy has been conducted with the assistance of healthcare professionals. Additional educational and action missions were included in the during pilot testing of the platform. Finally, a mini game which allows the player to prepare meals is always available in the SG's main screen. The food ingredients that are collected throughout the game are available for use here, categorized in food groups according to their content. The player places these ingredients in a food basket and when they are satisfied with their choices, they can place the basket for scoring. According to their choices they are rewarded with coins based on the contents of the food basket. If the prepared meal is balanced, then the rewards are greater. As for the mobile application for the parents, it has been designed with four purposes in mind. Firstly, it presents the messages delivered by the recommendation system on a daily basis. Secondly, it provides the parents with the ability to communicate with the healthcare professionals with messages. Thirdly, it allows the parents to input the child's weekly weight and daily dietary options. Fourthly, it distributes questionnaires to the parents, such as the Post-Intervention Feasibility Study Questionnaire.

4. Modeling engagement in serious games for health

Player modeling, as introduced in previous chapters, is the process of creating a model that describes the player's traits and tendencies [175]. These described characteristics are relevant to the player's actions and intentions, style, and preferences, as well as goals and motivation. The generated model attempts to recognize, or even predict, this information through cognitive, affective and behavioral patterns that manifest during game playing [176]. Player modeling facilitates the procedural generation of tailored content; however, this approach is not considered a prerequisite to such content. Methodologies that generate game content procedurally generate content either automatically or with minimal guidance [177].The advantages of employing some type of player model to guide content generation are twofold [175]. Firstly, player modeling provides a data driven approach able to justify game adaptations and facilitate PCG. Secondly, PCG techniques based on player models are considered more scalable and able to generalize to other games. There are three main types of data that are typically employed for the creation of player models. Gameplay data, collected through player interaction, are the most common, subjective data that are collected mainly from questionnaires, and objective data, obtained mostly from biometrical observations through various types of sensors.

The conceptual framework for adaptivity in SGs for health proposed in the present Doctoral Thesis highlights and focusses on the modeling of player engagement in real time through a sensor-based approach, facilitated with game metrics produced during SG play. A SG for health able to generate game content that maximizes player engagement can benefit from increased adherence to the intervention it provides. In addition, data collected from sensors can also be applied to enhance the player model with clinically relevant information. This approach allows for procedural generation of SG content that is tailored to the players specific clinical needs, improving the overall efficiency of the SG as a health intervention tool. In this chapter, the design of an experimental process featuring interaction with ECT and the analysis of results obtained from sensors, towards real-time recognition of engagement are presented. The experimental process and results were first presented in [178] and [3].

To investigate the potential of the proposed framework (Fig. 4-1) for real-time recognition of engagement during SG playing a sensor-based approach was designed and evaluated through an experimental process. The experimental process involved the interaction of volunteers with ECT. The employed SG was selected for two reasons, firstly due to its capability to serve as an intervention in self-health management for chronic conditions, and secondly, due to its potential to incorporate procedural generation of SG content related to the targeted condition. During the experiment, heterogenous data from a multitude of affordable and unobtrusive sensors have been collected to provide insight regarding a multimodal approach in recognition of engagement. The predictive power of features extracted from the collected data in terms of real-time recognition of engagement has been assessed through a detailed analysis. An approximation of the ground truth has been produced through self-annotation of perceived engagement in a continuous manner by leveraging a state-of-the-art tool designed for affective recognition in video games.

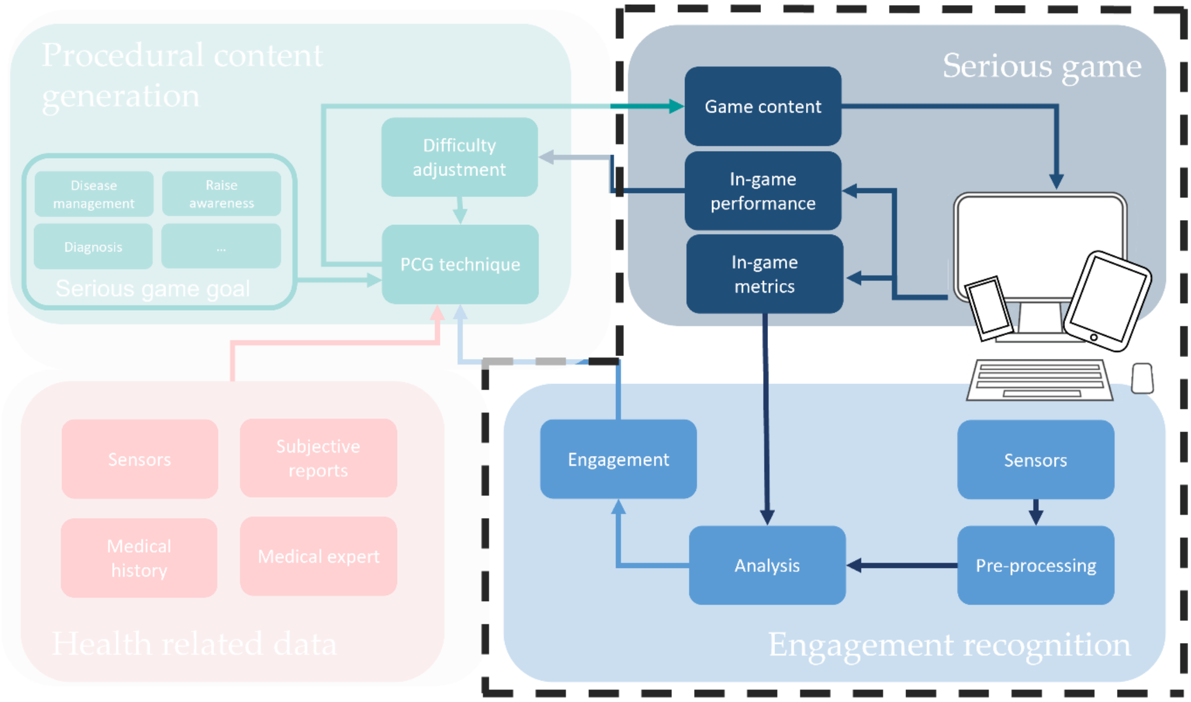


Figure 4-1: Conceptual framework for real-time recognition of engagement.

4.1 Experimental process

A sensor-based approach was designed to investigate the capability of recognising player engagement in real time during SG playing. The employed sensors (heart rate and pressure sensors) were selected due to their potential to be applied in the collection of clinically relevant data. In this way the application of the investigated methodology could be generalized within the proposed conceptual framework to incorporate PCG techniques that benefit both from real-time player engagement and recognized health-related needs. Data were also collected from the player interaction with the SG. The setup of the experimental process is presented in Figure 4-2. The process employs three interconnected spaces, which enable the necessary data acquisition and analysis. The SG space includes the in-game metrics that are gathered though player's interaction during playing. Self-reported annotation traces to be considered as the ground truth levels of engagement as perceived by the player during game play are generated through an annotation tool. Sensing data are collected by means of: (i) pressure sensors placed on a chair towards identifying postures and mobility, and (ii) a heart rate sensor providing the heart rate (beats per minute) and inter beat intervals. Two microcontrollers have been employed for the acquisition of sensor data. The collected heterogeneous data feed the data analysis space in order for the latter to apply thorough statistical analysis investigating the potential of sensor and in-game metrics data towards real-time engagement recognition. A summary of the investigated features is presented in Table 2. Each space is explained in detail in the following sections. Additionally, ECT's educational value was evaluated through a quasi-experimental study which involved a control group that received a traditional intervention. This alternative intervention was created based on text material. Finally, the effect of user interface changes based on player feedback provided after their interaction with ECT was investigated. The game experience was measured through the application of validated questionnaires after the participant's interaction with the SG.

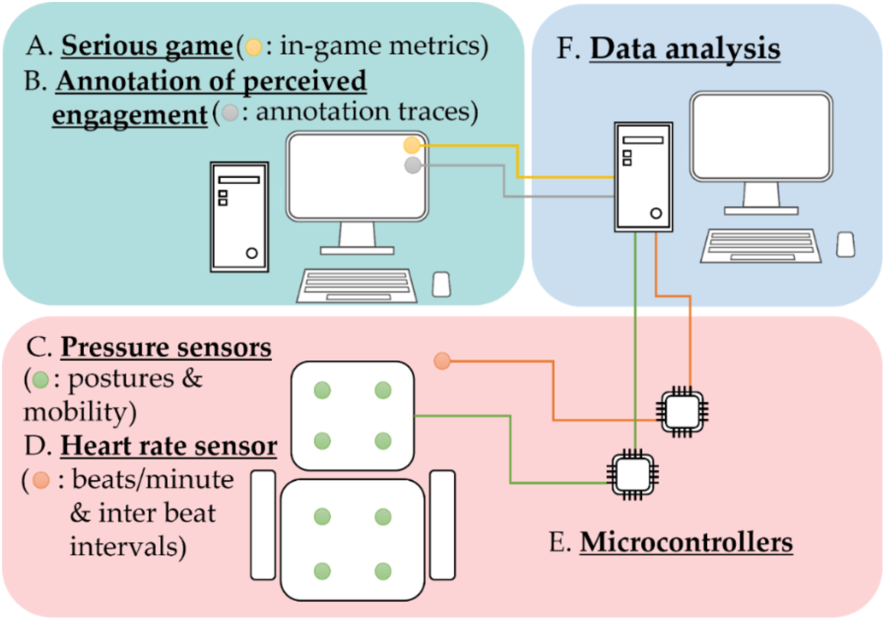


Figure 4-2: Setup of the experimental process.

Table 2: Summary of extracted features.

Feature source	Description	Symbol
In-game metrics (A)	Average number of mouse clicks per second	μMc
	Average distance of mouse movement per second in pixels	μMm
Annotation traces (B)	Average value of annotation trace	μA
	Area of annotation trace, calculated by the composite trapezoidal interval	$\int A$
	Amplitude of trace	$\hat{A}$
	Average gradient of trace	ΔA
Pressure sensors (C)	Average number of transitions (Postures)	μT
	Average gradient of pressure output (Mobility)	ΔT
Heart rate sensor (D)	Amplitude of heart beats per minute	$\hat{H}$
	Standard deviation of inter beat intervals	ΔH

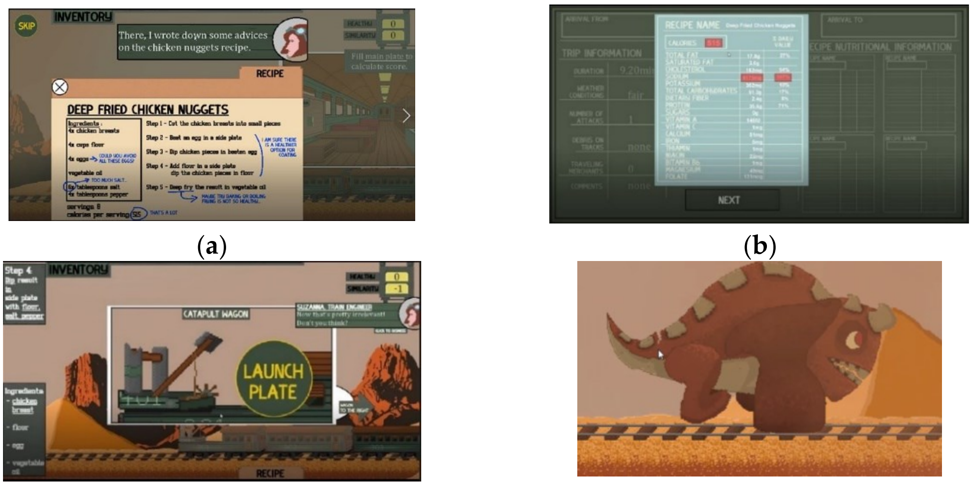
4.1.1 Serious game version

During the experimental process, a version of ECT that contains three game missions (Fig. 4-3) has been employed. Mission-1 includes a tutorial phase, a gameplay phase, and a review phase. During the tutorial phase, players are given instructions about game interface and mechanics, as they are introduced to the game world and objectives. Interaction during the tutorial is quite minimal as players are presented with explanatory text boxes and experiment with game functionalities under guidance. As tutorial ends, a hungry monster appears and chases the train, signifying the beginning of the gameplay phase. During this phase, players apply knowledge acquired in the tutorial to prepare healthy meals for the chasing monster and avoid it until the train reaches its destination. Finally, the review phase launches with a small cinematic of the train escaping the monster in case of a successful mission, or the monster catching up with the train in case of defeat. Following the cinematic, a review screen appears, containing game statistics, nutritional facts of recipes used, new discoveries and unlocked achievements. After Mission-1 is complete, players can continue playing for a maximum of two additional missions. The additional missions feature small train trips and thus include shorter gameplay and review phases. During playthroughs, the SG monitors in-game metrics, including the number of mouse clicks, mouse click duration and mouse idleness. Additionally, game score, mission progress and in-game decisions, cooking simulation parameters and game events are collected. Two features based on mouse interaction have been extracted, namely average clicks per second (μMc) and average mouse movement (μMm). Mouse movement measures the cursor distance travelled per second, in pixels.

4.1.2 Measurement of the intervention's efficiency

To evaluate the ECT's effectiveness in terms of its educational value and game experience, a two-part study was designed and implemented (Fig. 4-4). Part I of the study aimed at the evaluation of the game's educational value through a quasi-experimental approach. In Part II, the user's experience was assessed via a validated questionnaire. A total number of 29 participants were recruited (Table 3); nineteen of them (Game Group A and Control Group) participated in both parts of the study while the remaining ten participants (Game Group B) enrolled solely in the second part of the study. Most participants were undergraduate and postgraduate students at the School of Electrical and Computer Engineering of the National Technical University of Athens (NTUA). Data from 26 participants was employed for the investigation of real-time recognition of engagement. No participant had any apparent mobility or visual impairment and most participants had normal BMI scores (BMI score: 18.5 - 25), except for two slightly overweight (BMI score: 25-27), and two slightly obese (BMI score: 30 - 35). As ECT is in English, all participants reported a good understanding of the English language. All participants signed a consent form, and the study was approved by the Ethics Committee of the National Technical University of Athens (NTUA).

Upon arrival, participants were given a brief description of the experimental process, the aim of the study and the potential outcomes, while being encouraged to engage in conversation about possible concerns. Subsequently, they were provided with consent forms that included a detailed description of the experimental process. Upon providing their consent, the participants sat on the smart chair in front of a desktop computer and were asked to assume a comfortable position. The heart rate sensor was then placed on the ear lobe of their choice, and they were asked to perform exploratory movements while seated, to confirm that the sensor is not hampering them in any way. After setup was completed and participants felt comfortable, they were instructed to fill out some digital questionnaires including information about demographics, exposure to gaming and cooking habits . Participants were then instructed to start Mission-1 at their leisure while given the option to play the two additional game missions. Once the playthrough session was completed, the participants were given instructions on how to use the annotation tool. After a short time to familiarize with the mouse wheel control, participants annotated their perceived engagement on the video playback of their game session and concluded their participation. Mission-1 was played by all participants, while Mission-2 and Mission-3 by 14 and 3 participants, respectively.



(a) (b) (c) (d)

Figure 4-3: Screenshots from the three phases of Mission-1 and the monster: (a) Tutorial; (b) Gameplay; (c) Review; (d Monster

The SG's educational value was investigated against a control intervention based on the study of text-based material on dietary knowledge and food safety. The duration of both interventions lasted less than 20 minutes. Knowledge acquisition and retention were measured by administering a questionnaire a week prior to, immediately after, and a week after intervention. The Knowledge Questionnaire was created based on the combination of a General Nutrition Knowledge Questionnaire [179] and a Food Safety Knowledge Questionnaire [180]. The Knowledge Questionnaire consisted of a total of 25 questions in the form of multiple-choice and multiple-choice grid, with a maximum score of 52 points. The questions were selected to evaluate knowledge on food measurement, food safety and label food groups,, salt consumption and nutrition related chronic-diseases. These topics are widely used in tools measuring food and nutrition literature [181]. The questions were translated in Greek, with slight modifications applied in particular questions and answers for localization purposes.

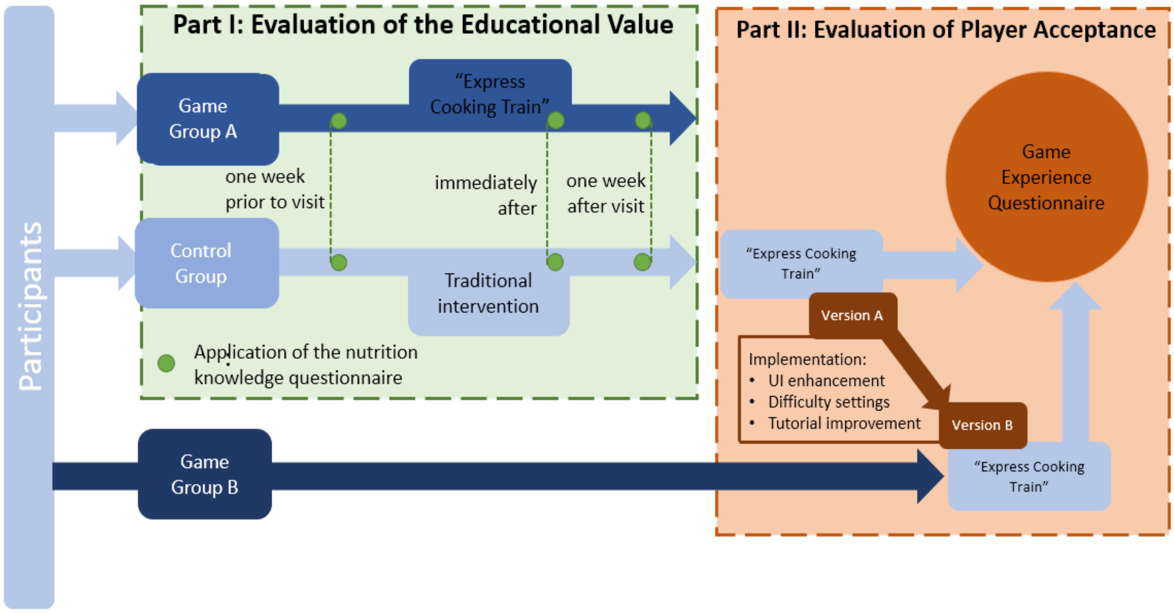


Figure 4-4: Design of the evaluation study for educational value and game experience of Express Cooking Train.

Table 3: Participants of the experimental process

Information about the participants		**Part I**		
		Part II		
	All (N=29)	**Game Group A (N=9)**	**Control Group (N=10)**	**Game Group B (N=10)**
Gender	male (19), female (10)	male (7), female (2)	male (4), female (6)	male (8), female (2)
Age	26 ± 4.08	28.67 ± 4.56	26.52 ± 4.70	22.81 ± 3.82
Gaming habits	**All (N=29)**	**Game Group A (N=9)**	**Control Group (N=10)**	**Game Group B (N=10)**
Never	10	1	5	4
Once or less per week	7	2	2	3
More than once per week	12	6	3	3

The evaluation of user experience was calculated by employing two modules of the widely used and validated Game Experience Questionnaire (GEQ) [182]. The Core module of the questionnaire comprises of multiple dimensions of game experience, that is, competence, sensory and imaginative immersion, flow, tension, challenge, negative affect, and positive affect. Likewise, the Post-game module deals with positive experience, negative experience, tiredness and returning to reality. Each dimension consists of a number of questions and receives a score based on the average value of its items, on a five-point Likert scale (from 0 to 4).

After concluding Part I, participants of the Control Group played ECT and responded to the GEQ. The data they provided, along with feedback data collected from Game Group A, were employed to apply specific enhancements to version A of ECT. This resulted in the updated Version B, which was used by Game Group B. Particularly, the following changes were implemented:

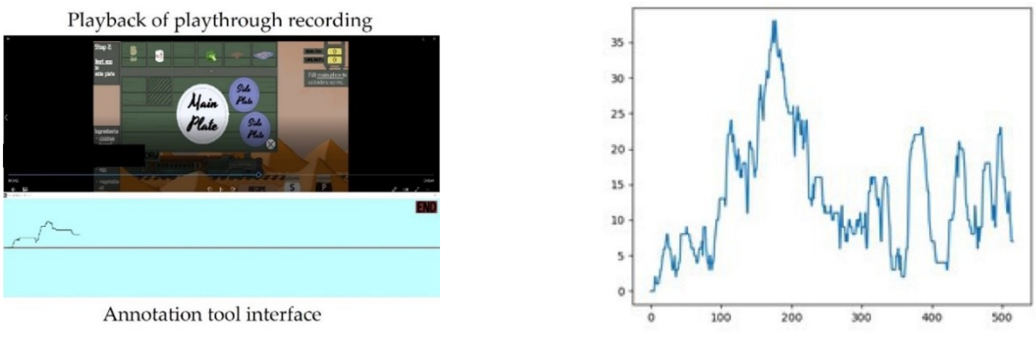
1) The overall difficulty was reduced by lowering the starting speed of the chasing monster.
2) User interface was augmented resulting in a more consistent and attractive user interface.
3) Tutorial was improved to facilitate user onboarding process.

4.1.3 Annotation of perceived engagement

Various annotation methodologies exist for the approximation of the ground truth regarding player engagement, varying greatly in terms of who the annotator is, when the annotation takes place, and how perceived engagement is annotated. During the present thesis, a tool based on RankTrace [183] was developed in GameMaker Studio (Fig. 4-5a). This design allows for continuous and unbounded annotation of affect by the player, by presenting them with screen recordings of the playthrough, right after the game session is completed. Using the tool, the player generates a continuous annotation trace (Fig. 4-5b), with an annotation sample recorded per second. The annotation trace was selected as reference in the harmonization process of the various heterogeneous data sources, and the sampling rate was set to facilitate it. Annotation data were normalized in the range [0,1] using minimum and maximum values of each individual annotation trace. From each observation frame generated, four statistical features were extracted corresponding to the mean annotation value (μA), the area of the annotation trace ($\int A$), calculated by the composite trapezoidal integral and normalized by duration, the amplitude ($\breve{A}$), calculated by the difference between maximum and minimum value, and the average gradient of the annotation trace (ΔA).

4.1.4 In-game metrics

During playthroughs, the SG monitors in-game metrics, including the number of mouse clicks, mouse click duration and mouse idleness. Additionally, game score, mission progress and in-game decisions, cooking simulation parameters and game events are collected. In the present study, two features based on mouse interaction have been extracted, namely average clicks per second (μMc) and average mouse movement (μMm). Mouse movement measures the cursor distance travelled per second, in pixels.



(a) (b)

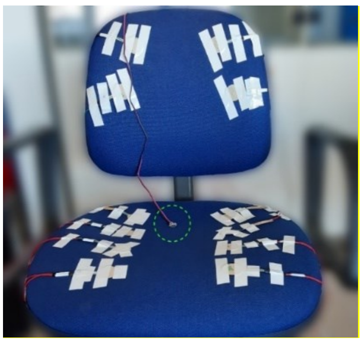
Figure 4-5: (a) The interface of the annotation tool along with the playthrough recording as shown in

4.1.5 Sensors

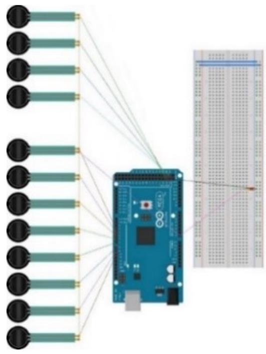
Sensor measurements were acquired through two Arduino Mega 2560 R3 [184] microcontrollers, one for the pressure sensors (Figure 4-6b) and one for the heart rate sensor, following the setup used in [185]. The microcontrollers transferred data to a desktop computer through a USB interface. PC port control was provided by the Python 3.6.5. The user interface to control the sensors was also developed in Python. A case was crafted and affixed to the back of the smart chair to hold the microcontrollers and breadboards in place and facilitate cable management.

A smart chair was employed to identify sitting postures and monitor their variations. A set of pressure sensors, FSR101 Shuntmode from Sensitronics [186], were placed on the seat and back of an office chair to measure pressure exerted by body weight during playthroughs of the SG on a desktop computer. The sensors were strapped on the chair, along with the cables linked to them. Afterwards, two cloth covers were fastened on the chair, one on the seat and one on the back, to secure sensor placement while reducing the risk of bias by making sure that sensors' location is not visible to the participants. Measurements were recorded from all sensors, with no load on the chair, to ensure that pressure from the tape or cloth cover did not affect sensor output. Eight sensors were placed to monitor pressure distribution on the seat of the chair (four under each thigh), while four sensors were placed on the back of the chair (two sensors on each side) to detect sitting back postures. Sensor arrangement along with the employed smart office chair are shown in Figure 4-6a.

Data collected from the smart chair were used to identify sitting postures during playthroughs based on a sensor activation methodology [113]. Postures were identified by detecting different sensor activation patterns and matching them to predefined sitting positions. A sensor placed on the smart chair was considered active when its output value exceeded a certain threshold. A set of six sitting postures (Table 4) were identified during the experimental process. Participants were observed to assume mainly upright postures, always activating most of the four front seat sensors due to the placement of monitor, mouse, and keyboard on the office desk. Additionally, no postures including leg crossing were observed during playthroughs. Data collected supported these observations, with sensors situated in the middle of the seat being always active. Based on these observations and preliminary analysis, data acquired from sensors 1,2,5 and 6 (Figure 4-6a, Table 4) were excluded from posture identification and the activation patterns shown in table 4 were selected. Postures P1 and P4-P6 include activated sensors on the back of the chair, whereas postures P2 and P3 do not. Two features were extracted from the observation frames. First, the total number of posture transitions (μT), normalized by duration was extracted to acquire a macroscopic measure of participant mobility [113]. Secondly, a feature of relative change (ΔT), calculated as the average gradient in sensor output, was extracted from the sensors included in posture detection, to provide insight regarding mobility observed in pressure distribution [112].



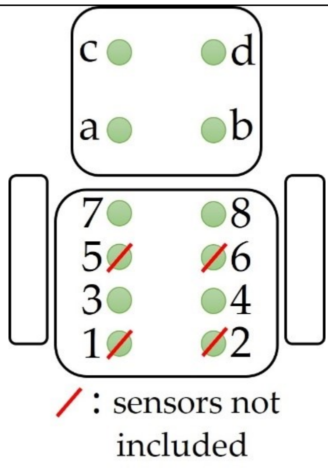
(a)



(b)

Figure 4-6: (a) Smart chair with initial arrangement of 12 pressure sensors. PulseSensor (placed in the user's ear lobe) is marked by the green circle.; (b) Microcontroller setup for acquisition of data from the smart chair.

Table 4: Set of sitting postures and their activation patterns.

Posture	Description	Activated sensors	Sensor location on chair
P1	Upright position with backrest	(3 or 7) and (4 or 8) and (a or d) and (b or c)	
P2	Upright position without backrest	(3 or 7) and (2 or d)	
P3	Front sitting position	(3 or 4)	
P4	Front sitting position with backrest	(3 or 4) and (a or b or c or d)	
P5	Upright position with right backrest	(3 or 7) and (2 or 8) and (a or c)	
P6	Upright position with left backrest	(3 or 7) and (2 or 8) and (b or d)	

The sensor activation threshold constitutes a vital component for the reliable identification of sitting postures. In this study, a separate small-scale experiment was conducted to estimate a general activation threshold. A group of six individuals within the healthy BMI range (BMI: 18.5 - 25) were recruited, yet these participants were excluded from the main experimental process to reduce the risk of bias. After an initial visual presentation of postures P1-P6, participants were told to test them, while sitting on the smart chair, with no time limit. Consequently, these postures were displayed through an application developed in GameMaker Studio, for 10 s each in random order, until all possible posture transitions had been presented. Participants were instructed to assume postures as they appeared on the screen.

To determine the appropriate threshold, a wide range of sensor output values (1 – 300 mV) were considered to identify sitting postures. The obtained average accuracy across all participants for all activation thresholds is depicted in Figure 4-7. The maximum average accuracy (0.96) was achieved for activation threshold values ranging from 86 mV to 93 mV. Multiple ANOVA single factor tests were applied on batches of 50 consequent activation threshold values, to investigate statistically significant differences in mean accuracy values across participants with respect to different threshold values. No significant differences were observed for activation threshold values higher than 30 mV ($P > 0.05$). Consequently, the general activation threshold selected for sitting posture identification was 90 mV.

The heart rate sensor employed within the frame of this study is the PulseSensor from World Famous Electronics llc. [187]. PulseSensor is an affordable and non-obtrusive sensor that can be placed around the finger, or on the ear lobe. For this study, the sensor was placed on the ear lobe to avoid obstruction during game play. The sensor detects pulses through a light-emitting diode generating a photoplethysmography (PPG). Inter beat intervals (ms) along with beats per minute were obtained in real time from PPG. Data from PulseSensor was collected at a rate of approximately 25 Hz. Inter-beat intervals were isolated and ranked in time order. The intervals were then pre-processed for the removal of ectopic beats and outliers in Python 3.6.5. From the resulting values, two features were extracted: the amplitude of heart beats per minute ($\widehat{H}$), measured as the difference between maximum and minimum heart rate value, and the standard deviation (σH) of inter-beat intervals [188].

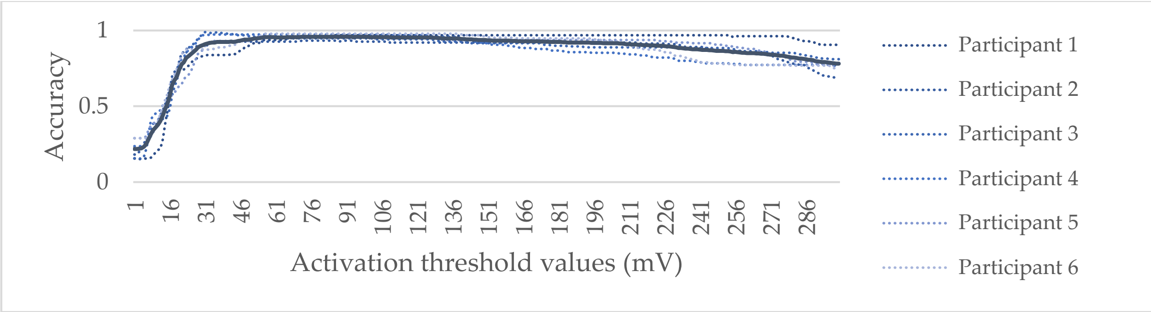


Figure 4-7: Accuracy results towards the determination of sensor activation threshold for posture identification.

4.1.6 Data analysis

Regarding the validation of ECT, for the first part of the study, Knowledge Questionnaire scores were acquired by Game Group A and Control Group, one-week prior to, immediately after and one-week after the intervention. Paired sample Student's t-test was employed between pre- and post-intervention scores to investigate the educational effectiveness of both interventions. Two-sample t-test was used to examine for significant differences between the scores of Game Group A and Control Group. For the second part of the validation study, in order to investigate potential improvement of version B over version A, in terms of user experience, GEQ scores were compared by applying two-sample Student's t-tests. Additionally, data collected from the in-game metrics regarding interaction through the computer mouse was analyzed using the Pearson correlation coefficient. One-way ANOVA measurements were employed to investigate possible association between mouse interaction and self-reported user experience.

Regarding the recognition of engagement in real-time, data harmonization was conducted to ensure the synchronization of heterogeneous data collected from different sources during the experimental procedure. The considered data sampling rate for in-game metrics and pressure sensing data was 1 Hz. Annotation traces and recordings of playthroughs were employed as synchronization reference. A moving average filter with a cut-off frequency of 1 Hz was applied on the signals obtained from the pressure sensors for noise removal and synchronization. Data from the heart sensor were pre-processed and synchronized as described in Section 4.1.3.

Different types of observation frames (continuous and reactive) were considered [183] to link the ground truth, investigated features, and identified postures with different types of gameplay and specific game mechanics, as depicted in Figure 4-8. Initially, continuous, and non-overlapping observation frames, representing different game phases (Tutorial, Gameplay, Review, Mission-2, Mission-3) were generated. The average duration of all game phases for all participants is presented in Figure 4-8. Due to the limited number of participants advancing to the Mission-3, the frames corresponding to this mission were omitted from the analysis. Game phases correspond to the different types of gameplay, with engagement levels expected to vary according to user preferences. The Tutorial phase is a linear and educational phase, with rich text content and minimal interaction (Figure 4-3a). The Gameplay phase requires a higher degree of game interaction and provides an exploration experience, with the player being free to experiment with ingredients and cooking tools while switching between train wagons (Figure 4-3b). Furthermore, the Gameplay phase includes the danger imposed by the chasing monster and the possibility of defeat. The Review phase contains a lot of information yet allows the player to survey it freely (Figure 4-3c). Additionally, the Review phase features the element of reward in the form of score points, discoveries, and achievements unlocked. The Mission-2 phase provides a similar gameplay experience to the one provided by the Gameplay phase.

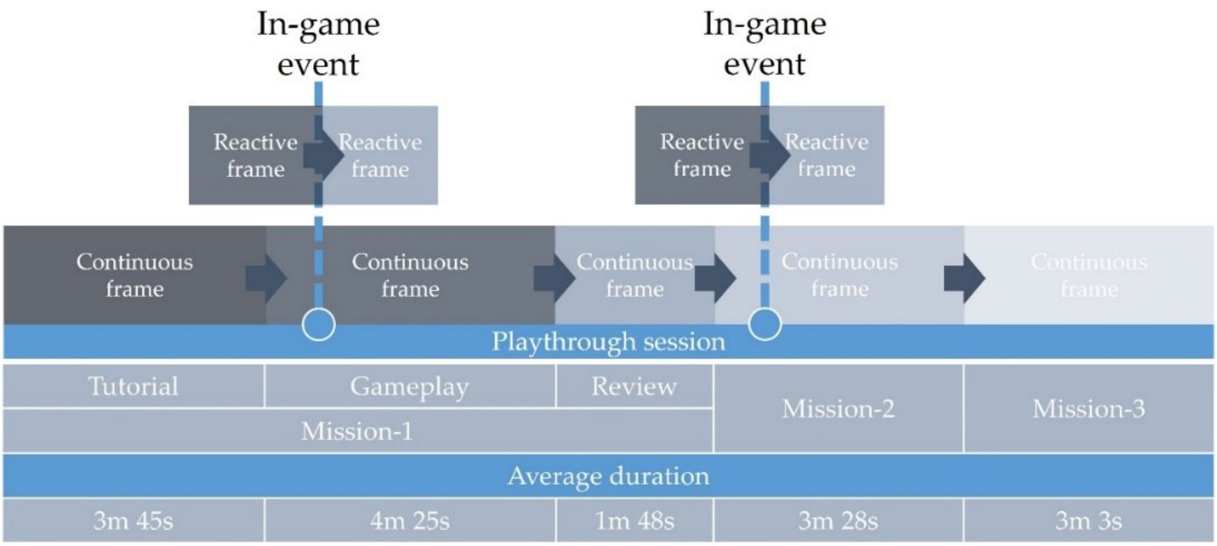


Figure 4-8: Continuous and reactive frames

Reactive observation frames were specified as those triggered by specific in-game events. In-game events include a visual alarm indicating danger and triggered by close monster proximity to the train, and monster related player interaction such as launching a meal to the monster with the catapult or clicking on the monster. These events were selected to point towards game moments that favour changes in player engagement. The nature of these events is memorable, aiming to produce more accurate annotation traces of perceived engagement around them. Furthermore, the manifestation of these events is not scripted and is based on the player's actions and performance; hence players cannot expect or plan them, reducing thus risk of bias. Each event generates two reactive observation frames, prior to and after the event. A total of 89 in-game events were produced during the participants' playthroughs. Reactive frames of different duration, 10 s and 30 s, were investigated in accordance with current practice for ultra-short analysis of heart rate variability [189], [190]. As is evident in Fig. 4-8, the duration of reactive frames was quite shorter than the duration of continuous frames.

Features were extracted from all data sources and sitting postures were identified for all observation frames, continuous and reactive, across participants. The feasibility of real time recognition of engagement during play was investigated in two parts as shown in Figure 4-9. The first part relied on statistical analysis of sitting postures and features of perceived engagement as extracted from annotation trace. Contingency tables were generated for each observation frame, continuous and reactive, to perform transition analysis [113]. The element (P_x, P_y) of the contingency table represents the number of times a transition was identified from posture P_x to posture P_y. Distributions of identified postures were extracted from the contingency tables for each observation frame. Wilcoxon signed-rank tests were employed to search for statistically significant differences between observation frames in terms of identified postures and perceived engagement. Additionally, whisker boxplots were created from the annotation features to accurately present trends in perceived engagement.

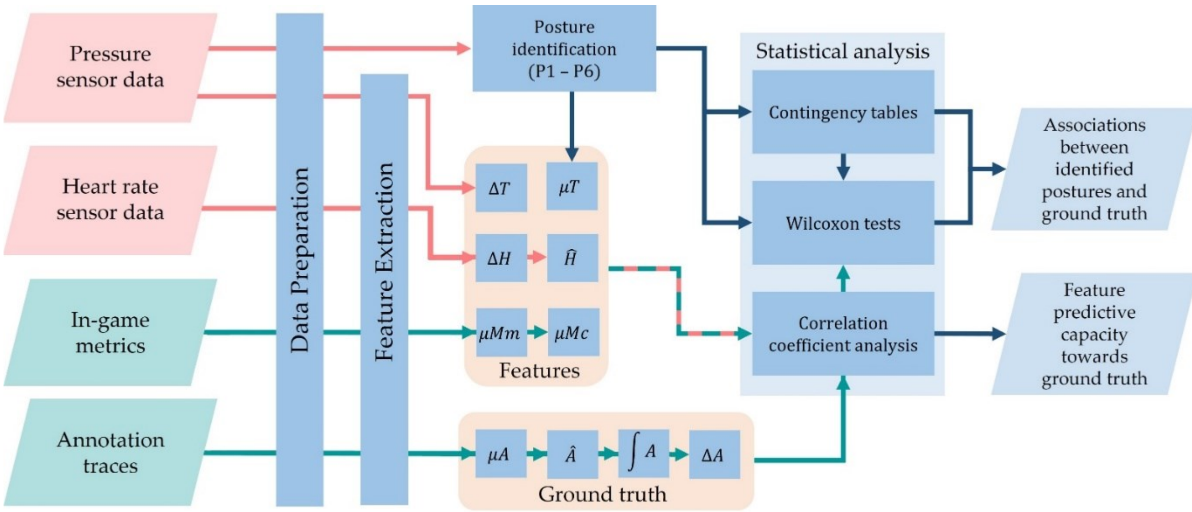


Figure 4-9: Flowchart for data analysis.

The second part of the analysis evaluates the predictive capability of features extracted from sensors and in-game metrics, based on relative changes observed between adjacent observation frames, towards perceived engagement. To this end, an analysis based on correlation coefficients [183], [191] was conducted for continuous observation frames and reactive observation frames, separately. More specifically, a correlation coefficient,

$$c_{i-j}(\boldsymbol{z}) = \sum_{k=1}^{N}\{{}^{z_{k,i,j}}/_{N}\}, \text{ with } i \in [\mu A, \int A, \hat{A}, \Delta A],\ j \in [\mu Mc, \mu Mm, \mu T, \Delta T, \widehat{H}, \sigma H] \quad (1)$$

was calculated for every possible combination of pairs between annotation features (i), and sensor and in-game metrics features (j). For each participant, the observation frames were ranked in order of time, with N representing the total number of adjacent frames across all participants. By measuring agreement in relative change of features i and j between the $k-th$ pair of adjacent frames, $z_k^{i,j}$ was calculated as,

$$\begin{aligned} z_k^{i,j} &= +1, if\ relative\ change \text{ of } i \text{ and } j \text{ match} \\ z_k^{i,j} &= -1, if\ relative\ change \text{ of } i \text{ and } j \text{ does not match} \end{aligned} \quad (2)$$

If clear relative change in any of the examined features was not present, the corresponding $z_k^{i,j}$ was not included in the calculation of $c_{i-j}(\boldsymbol{z})$. The average number of pairwise comparisons (N) of all investigated feature pairs for all participants, per type of observation frame, after the exclusion of pairs that did not display clear relative change, is shown in Table 5. Relative change in the average number of posture transitions (μT) was clear in very few comparisons (25.2 ± 13.6) and no statistically significant resulting values of $c_{i-\mu T}(\boldsymbol{z})$ were observed. As such, $c_{i-\mu T}(\boldsymbol{z})$ values were excluded from the corresponding analysis. The p-values of $c(\boldsymbol{z})$ were calculated through the binomial distribution, with correlation being highly significant for $P < 1\%$, and significant for $1\% < P < 5\%$.

Table 5: Number of pairwise comparisons included in different types of observation frames.

	Reactive frames (10 s)	Reactive frames (30 s)	Continuous frames
Primary features	76 ± 5.6	83.7 ± 3.7	56.4 ± 0.6
Multimodal feature	76.5 ± 1.1	82.7 ± 0.4	55.7 ± 0.4

Motivated by the superior performance that can be achieved through combining different modalities [192], a majority voting scheme was investigated towards the generation of a new multimodal feature (V). The choice of this particular combination scheme was based on its approved robustness in binary cases [193]. The majority voting scheme assumes the dominant relative change observed in all primary (sensors and in-game metrics) features. Voting includes only clear relative changes, and in case majority voting does not produce a clear result, the pair is excluded from the calculation (Table 5). Consequently, $c_{i-V}(\boldsymbol{z})$ is calculated for all features (i) of perceived engagement.

4.2 Results and discussion

4.2.1 Measurement of the intervention's efficiency

The results from the assessment of the educational value of the SG are presented below. Answers to the Knowledge Questionnaire were collected from all participants from Part I of the study one week prior to and immediately after the intervention session. Regarding answers gathered on the week after the session, there were two dropouts from the Game Group A and one from the Control Group. The participants' scores on the Knowledge Questionnaire are summarized in Fig. 4-10. Paired sample t-test demonstrated statistically significant increase between score values before and immediately after the intervention for both groups (p=0.002 and p=0.025 for Game Group A and Control Group). The corresponding increase in the score value was 3.44±2.60 and 2.45±3.67 for the Game Group A and Control Group, respectively.

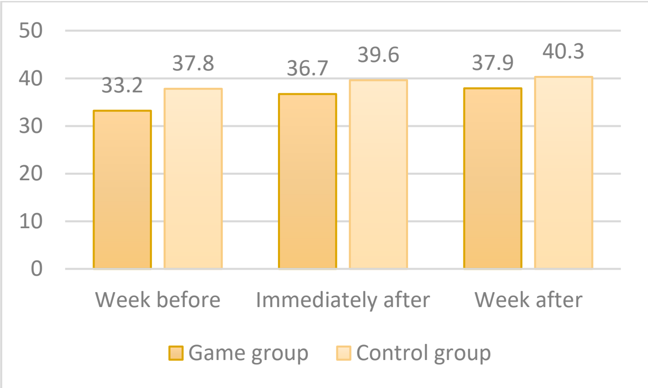


Figure 4-10: Average scores in Knowledge Questionnaire for Game Group A and Control Group.

A single participant from each group scored the same one-week prior to and immediately after the intervention. All of the participants from Game Group A scored higher immediately after the intervention, whereas two participants of the Control Group scored lower immediately after the intervention. No statistically significant differences were observed between Game Group A and Control Group in terms of score difference in the Knowledge Questionnaire one-week prior to and immediately after the intervention (p=0.25). Comparison between Knowledge Questionnaire scores administered immediately after and one-week after the intervention was not found to be statistically significant different for the two groups.

Knowledge Questionnaire score results demonstrated that both ECT and the traditional intervention are effective in terms of nutrition and food knowledge procurement. These preliminary results are in line with previous reports demonstrating SG effectiveness as an educational tool [194]. No statistically significant difference was observed, in terms of educational value, between the two interventions. However, it should be mentioned that all participants of the Game Group A achieved higher or equal score after intervention, while two participants of the Control Group scored lower, thus, implying the potential of ECT to demonstrate higher educational value. Administration of the Knowledge Questionnaires one-week after the intervention did not produce statistically significant results regarding knowledge retention, in both Game Group A and Control Group. Further investigation is necessary to determine proper time intervals between intervention for a better assessment of knowledge retention.

4.2.2 Game experience

Scores for the GEQ Core module dimensions are presented in Table 6 for both versions of the game that were employed. An increase across all positive dimensions except challenge was observed for version B over version A. Analysis of the results revealed statistically significant differences for competence and positive affect (p=0.006 and p=0.05 respectively). Comparison of versions A and B in terms of negative dimensions yielded no statistically significant differences. Scores for the GEQ Post-game module are presented in Table 6 for versions A and B of the ECT Comparison between the two versions for negative experience, tiredness and returning to reality produced no statistically significant differences. Positive experience demonstrated a significant increase of 0.75 points (p=0.007) after the game improvements were implemented.

Table 6: Results for GEQ for the two versions of Express Cooking Train.

	Mean ± Standard Deviation				
Core module (scale 1-4)	Version A	Version B	Post-game module (scale 0-4)	Version A	Version B
Competence	1.98 ± 0.62	**2.53** ± 0.44	**Positive experience**	1.53 ± 0.74	**2.28** ± 0.70
Immersion	2.79 ± 0.68	3.03 ± 0.44	Negative experience	0.30 ± 0.34	0.32 ± 0.44
Flow	2.49 ± 0.89	2.84 ± 0.65	Tiredness	0.24 ± 0.34	0.31 ± 0.95
Tension	0.51 ± 0.61	0.22 ± 0.44	Returning to reality	0.70 ± 0.56	0.77 ± 0.72
Challenge	2.31 ± 0.76	2.22 ± 0.73			
Negative Affect	0.42 ± 0.45	0.30 ± 0.37			
Positive Affect	2.87 ± 0.65	**3.22** ± 0.45			

In order to evaluate user experience in ECT against available SGs [37], GEQ scores of "Express Cooking Train" versions A and B were compared with corresponding results reported in the literature [195], [196], [197], [198], [199] for five SGs for the GEQ Core module and three SGs for the GEQ Post-game module (Table 7). The scores are presented in Fig. 4-11 and 4-12, for the Core and the Post-game module, respectively. Version B of ECT achieved the highest score in four positive dimensions (competence, immersion and positive affect and positive experience), and the lowest in two negative dimensions (tension and negative affect). However, comparison between SGs of different genres, designed for different target groups and purposes, is not expected to provide solid results regarding user acceptance. However, such a comparison could serve as a tool to investigate the effect of different choices in SG design. Differences in GEQ score results observed between game version A and B, indicate that the latter produces a better game experience amongst participants. The introduction of changes based on player feedback proved to have an immediate effect on user experience, thus, highlighting the importance of participatory user-driven game design for SGs.

Table 7: Information about serious games applying GEQ identified in the literature.

Serious Game	Rehabilitation game [195]	Adventures in Sophoria [196]	SimSe [197]	Ilha dos Requisitos [198]	Game Bridge [199]	Express Cooking Train
Participants	48	26	12	12	50	29
Age group	20-43	7-19	19-25	19-25	-	20-34
Game genre	Exergame	Massively multiplayer online role-playing game	Simulation	Role-playing game	Multiplayer puzzle	Roguelike, simulation, puzzle
Game purpose	Upper limb stroke rehabilitation	Facilitating intercommunication during cancer treatment	Company simulation	Survival skills	Problem solving in human sustainability	FL and NL skills

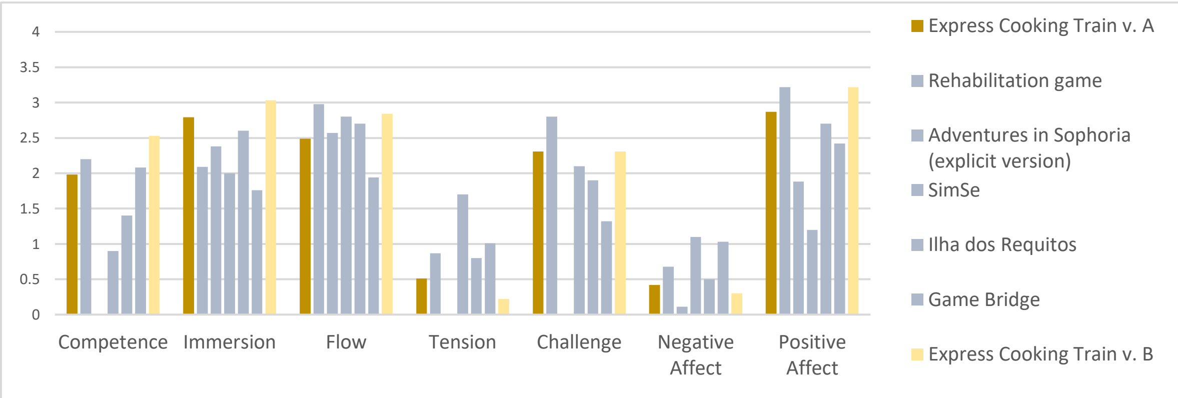


Figure 4-11: GEQ Core module dimensions for version A and B of Express Cooking Train, along with comparison with five serious games identified in the literature.

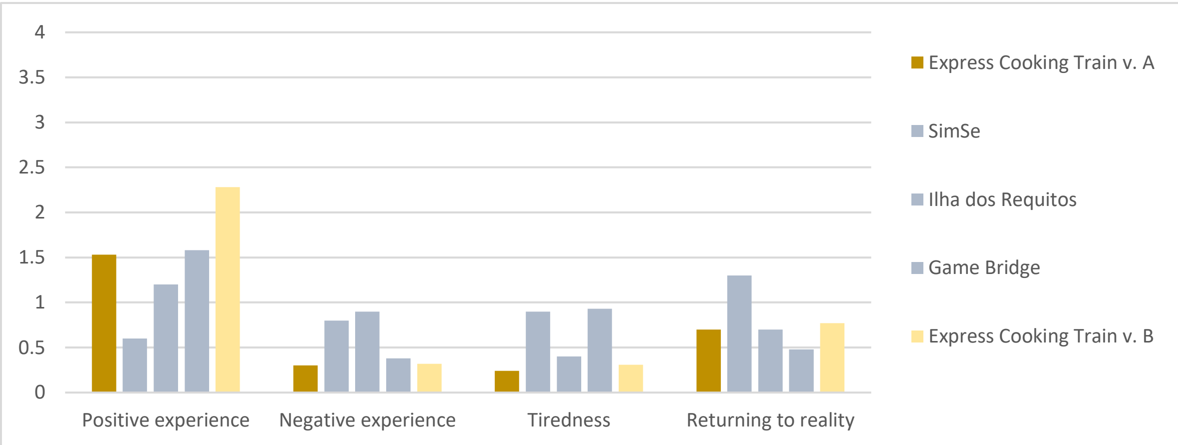


Figure 4-12: GEQ Post-game module dimensions for versions A and B of "Express Cooking Train", along with comparison with three serious games identified in the literature.

Preliminary analysis of data collected from player interaction focused on mouse clicks and mouse movement. Relevant data were collected from 25 participants. A strong positive correlation between average mouse clicks per second and average mouse movement per second was found (r=0.88, p=0.0001), as expected. Additionally, a weak negative correlation between average mouse clicks per second and average click duration was identified (r=-0.35, p=0.084). Based on the level of user's interaction in terms of clicks per second, three different clusters were identified:

- Cluster 1: low interaction with clicks/sec < 0.3 (N = 11)
- Cluster 2: intermediate interaction with 0.3 < clicks/sec < 0.4 (N = 9)
- Cluster 3: high interaction with clicks/sec > 0.4 (N = 5)

where N is the number of participants in the identified clusters. ANOVA test demonstrated statistical significance in four dimensions (positive experience, competence, tiredness, tension) of the GEQ (Fig. 4-13) for the second cluster. The clusters of participants with the highest and lowest mouse interaction levels scored significantly higher in terms of positive experience and competence, as well as significantly lower in terms of tiredness and tension dimensions in comparison with participants of cluster 2.

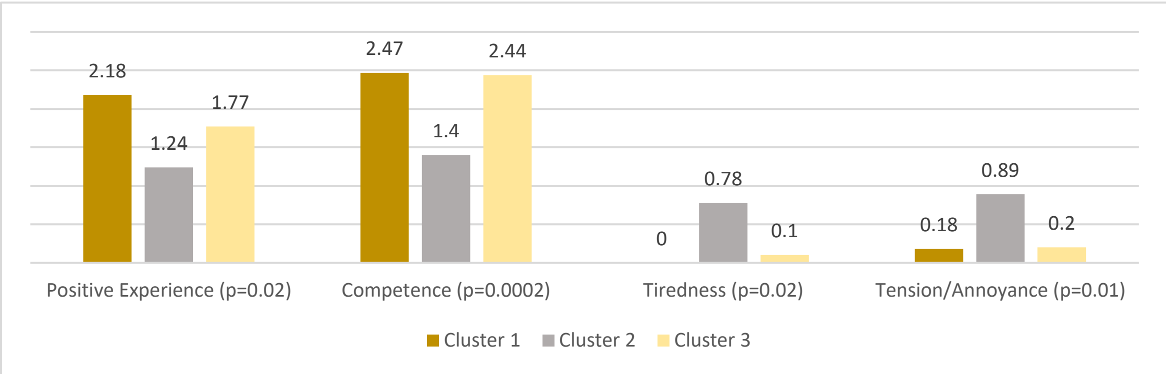


Figure 4-13: GEQ dimensions with significant differences among Clusters 1 (low interaction), 2 (intermediate interaction) and 3 (high interaction).

Analysis of data from player interaction with ECT against scores provided by the GEQ proved to include valuable information that can lead to improved personalization solutions. User characterization and clustering can provide unique insight to user specific needs [200]. Within this context, an important finding of the present study was that the users displaying intermediate levels of mouse interaction exhibited significantly lower satisfaction in terms of user experience. This could be considered as preliminary evidence that high levels of mouse interaction indicate engagement focusing on game mechanics, while low levels point towards an immersive experience with focus on game content.

4.2.3 Real time recognition of engagement

Results from both parts of the data analysis for real time recognition of engagement are presented for the investigated continuous and reactive observation frames. Data collected from two participants were excluded from both parts of the analysis for reactive observation frames, as in-game events were not recorded properly by the SG. Additionally, data collected from two more participants were excluded from the second part of the analysis due to movement of the heart rate sensor during play.

The distribution of identified postures (P1-P6) for continuous observation frames, Tutorial, Gameplay, Review, and Mission-2 is depicted in Figure 4-14. A statistically significant decrease in the percentage of postures including the back of the chair was observed from Tutorial to Gameplay (one-sided Wilcoxon: P = 0.03). Posture P3 (front sitting) was identified in very few occasions (≤ 0.1%) across all collected data. No other statistically significant changes in identified postures were identified between continuous observation frames.

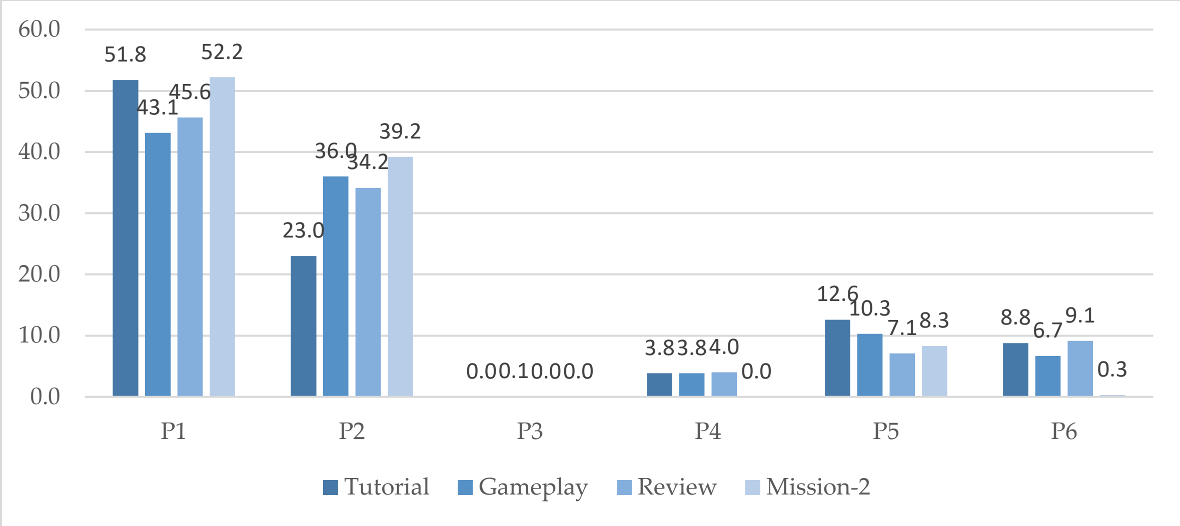


Figure 4-14: Distribution of postures identified for Tutorial, Gameplay, Review, and Mission-2.

Contingency tables showing the percentages of postures for continuous observation frames are presented in Figure 4-15a-d. The percentage of transitions is presented in Figure 4-16, with participants demonstrating the highest mobility in Gameplay. However, no statistically significant changes were present between any pairs of observation frames.

	P1	P2	P3	P4	P5	P6
P1	51.0	0.1	0.0	0.0	0.3	0.3
P2	0.0	22.3	0.0	0.0	0.4	0.3
P3	0.0	0.0	0.0	0.0	0.0	0.0
P4	0.0	0.0	0.0	3.8	0.0	0.0
P5	0.3	0.4	0.0	0.0	11.9	0.0
6	0.4	0.3	0.0	0.0	0.0	8.2

(**a**)

	P1	P2	P3	P4	P5	P6
P1	41.6	0.1	0.0	0.0	1.0	0.3
P2	0.1	35.1	0.0	0.0	0.5	0.3
P3	0.0	0.0	0.0	0.0	0.0	0.0
P4	0.0	0.0	0.0	3.8	0.0	0.0
P5	1.1	0.5	0.0	0.0	8.7	0.0
P6	0.3	0.3	0.0	0.0	0.0	6.1

(**b**)

	P1	P2	P3	P4	P5	P6
P1	44.3	0.1	0.0	0.0	0.7	0.5
P2	0.1	33.2	0.0	0.0	0.4	0.5
P3	0.0	0.0	0.0	0.0	0.0	0.0
P4	0.0	0.0	0.0	4.0	0.0	0.0
P5	0.7	0.4	0.0	0.0	6.0	0.0
P6	0.4	0.5	0.0	0.0	0.0	8.2

(c)

	P1	P2	P3	P4	P5	P6
P1	53.4	0.1	0.0	0.0	0.8	0.1
P2	0.0	32.9	0.0	0.0	0.5	0.0
P3	0.0	0.0	0.0	0.0	0.0	0.0
P4	0.0	0.0	0.0	0.0	0.0	0.0
P5	0.9	0.4	0.0	0.0	10.0	0.0
P6	0.1	0.0	0.0	0.0	0.0	0.5

(d)

Figure 4-15: Contingency tables for continuous frames: (a) Tutorial; (b) Gameplay; (c) Review; (d) Mission-2

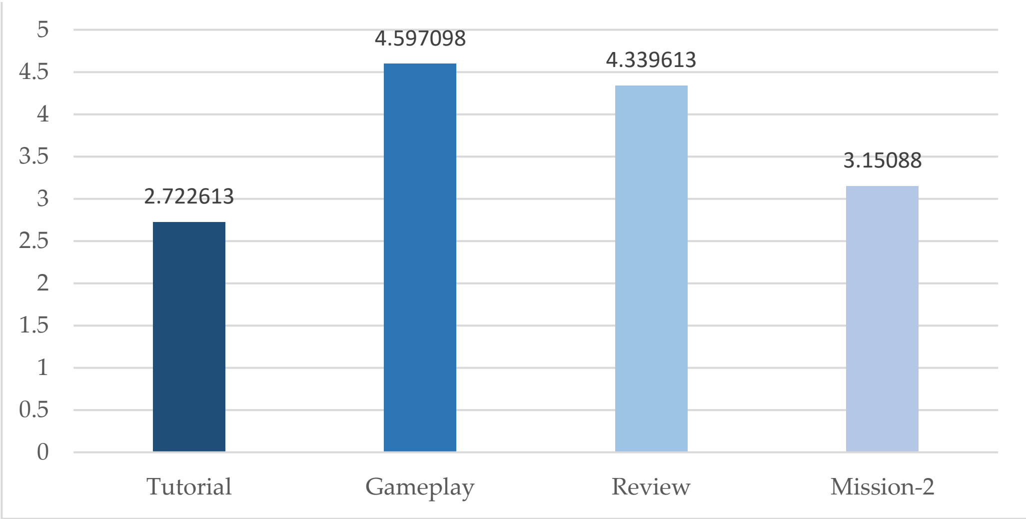


Figure 4-16: Percentage of posture transitions identified in each frame.

Whisker boxplots for features extracted from the annotation traces of perceived engagement are presented in Figure 4-17. The two-side Wilcoxon test revealed statistically significant (p-value<0.01) increase of 82.75 % and 79.31 % from Tutorial to Gameplay in mean value (μA) (Figure 4-17a) and area of the annotation trace ($\int A$) (Figure 4-17b), respectively. A statistically significant increase, of 34.09 % and 37.20 %, was also present from Review to Mission-2 for these two features, respectively. The decrease depicted for μA and $\int A$ from Gameplay to Review was not significant. Changes observed in amplitude ($\hat{A}$) (Figure 4-17c) were not statistically important. A decrease of 76.47 % and 428.79 % in the average gradient of the annotation trace (ΔA) (Figure 4-17d), from Tutorial to Gameplay and from Gameplay to Review respectively, were found to be statistically significant (two-side Wilcoxon, P < 0.01).

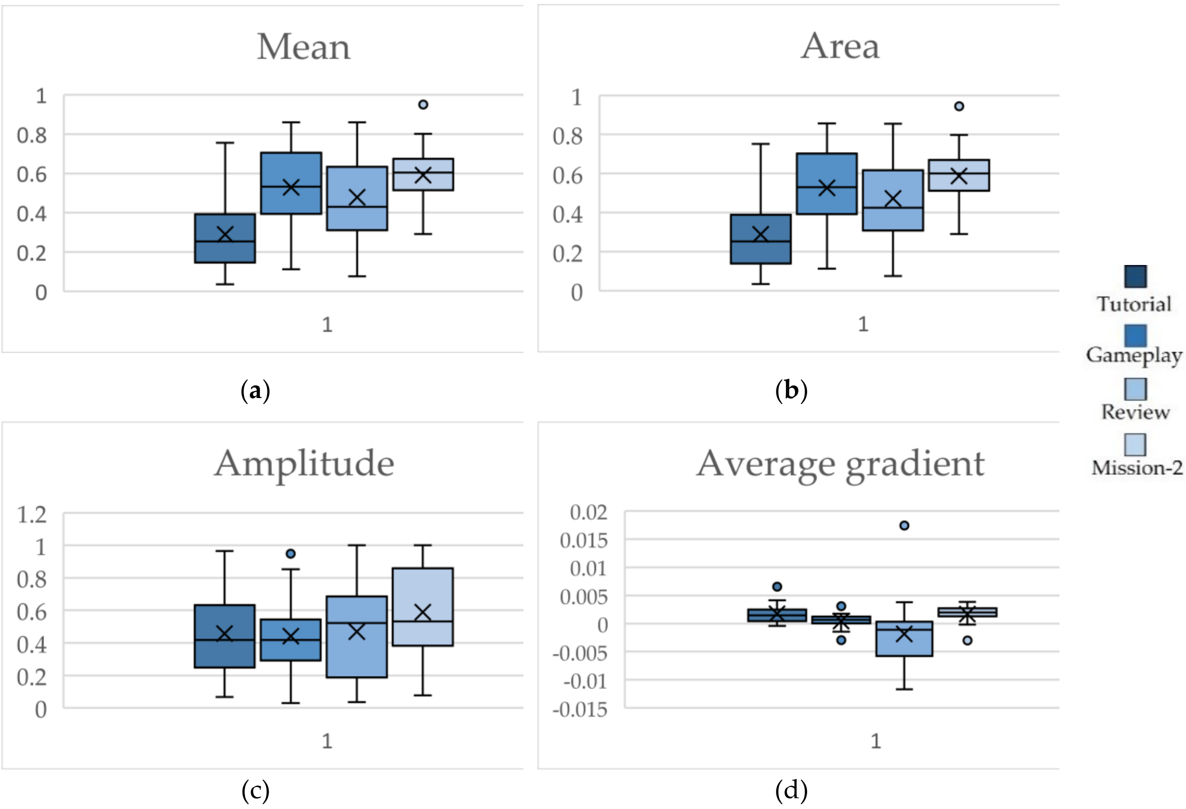


Figure 4-17: Whisker boxplots for Tutorial, Gameplay, Review and Mission-2, for (a) mean value of perceived engagement, (b) area of annotation trace, (c) amplitude of engagement, (d) average gradient of engagement.

The above presented analysis was also applied for reactive observation frames. The distribution of identified postures (P1-P6), for reactive frames of 10 and 30 s, is depicted in Figure 4-18. A statistically significant decrease in postures including the back of the chair was present in 30 s frames (two-side Wilcoxon: P = 0.03). No other statistically significant change in postures between reactive observation frames was observed.

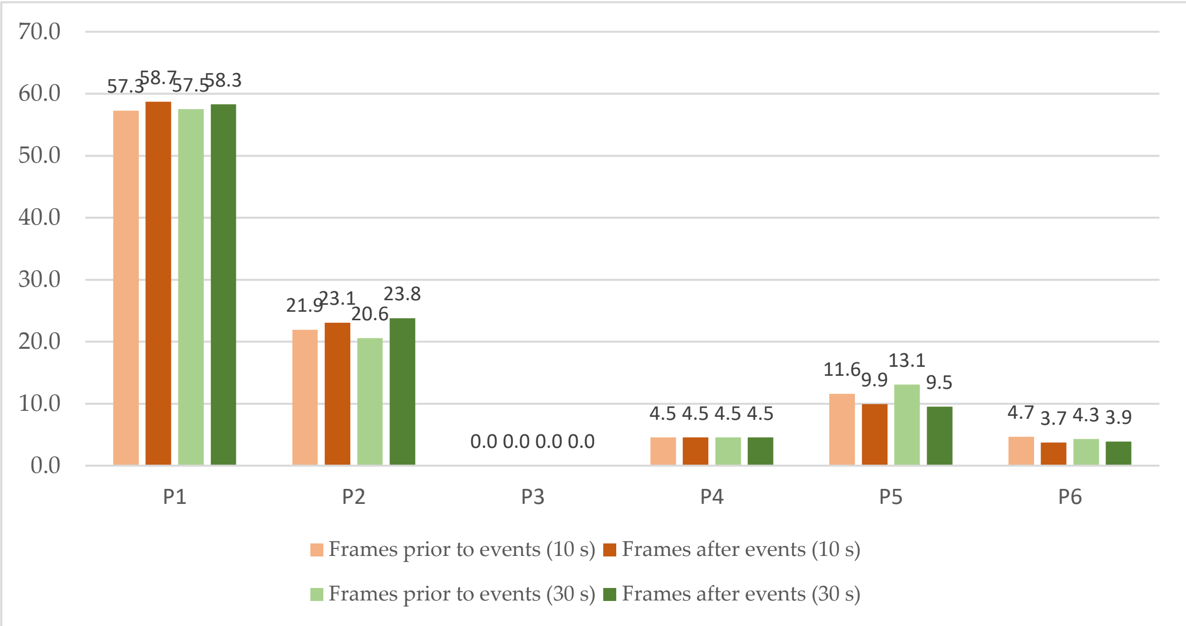


Figure 4-18: Distribution of postures identified during reactive frames.

	P1	P2	P3	P4	P5	P6
P1	56.1	0.1	0.0	0.0	0.1	0.6
P2	0.0	21.0	0.0	0.0	0.6	0.2
P3	0.0	0.0	0.0	0.0	0.0	0.0
P4	0.0	0.0	0.0	4.5	0.0	0.0
P5	0.1	0.4	0.0	0.0	10.6	0.2
P6	0.4	0.1	0.0	0.0	0.2	3.6

(a)

	P1	P2	P3	P4	P5	P6
P1	56.9	0.0	0.0	0.0	0.7	0.5
P2	0.3	22.3	0.0	0.0	0.1	0.4
P3	0.0	0.0	0.0	0.0	0.0	0.0
P4	0.0	0.0	0.0	4.5	0.0	0.0
P5	0.3	0.2	0.0	0.0	8.9	0.0
P6	0.6	0.3	0.0	0.0	0.1	2.8

(b)

	P1	P2	P3	P4	P5	P6
P1	56.1	0.0	0.0	0.0	0.4	0.3
P2	0.0	19.9	0.0	0.0	0.4	0.1
P3	0.0	0.0	0.0	0.0	0.0	0.0
P4	0.0	0.0	0.0	4.5	0.0	0.0
P5	0.4	0.3	0.0	0.0	12.0	0.1
P6	0.3	0.0	0.0	0.0	0.1	3.8

(c)

	P1	P2	P3	P4	P5	P6
P1	56.7	0.0	0.0	0.0	0.5	0.4
P2	0.3	22.9	0.0	0.0	0.1	0.2
P3	0.0	0.0	0.0	0.0	0.0	0.0
P4	0.0	0.0	0.0	4.5	0.0	0.0
P5	0.4	0.2	0.0	0.0	8.7	0.0
P6	0.2	0.4	0.0	0.0	0.1	3.2

(d)

Figure 4-19: Contingency tables for reactive frames: (a) Frame prior to event (10 s); (b) Frame after event (10 s); (c) Frame prior to event (30 s); (d) Frame after event (30 s)

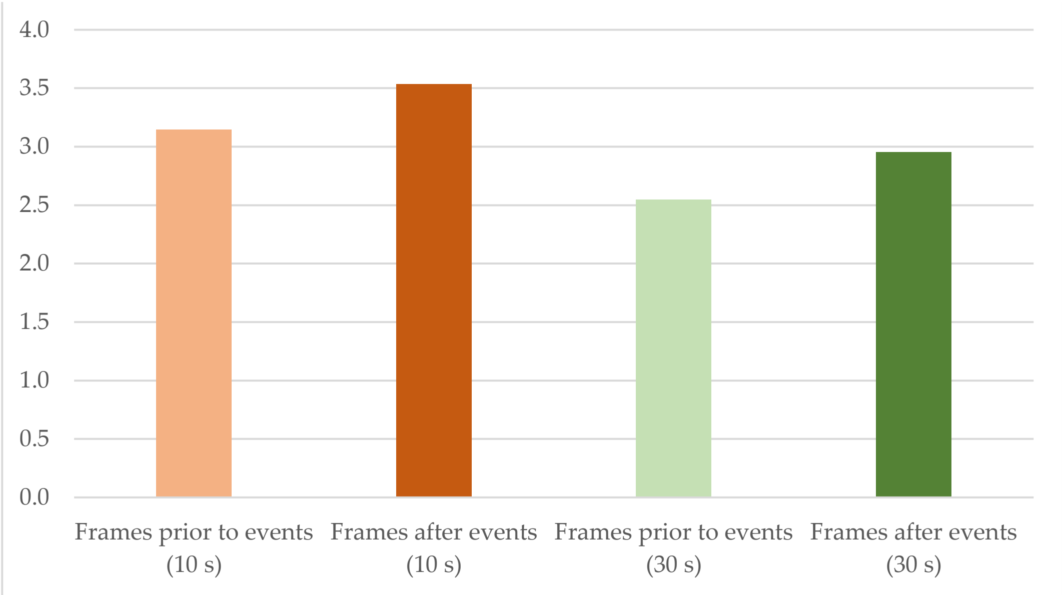


Figure 4-20: Percentage of posture transitions identified in each frame.

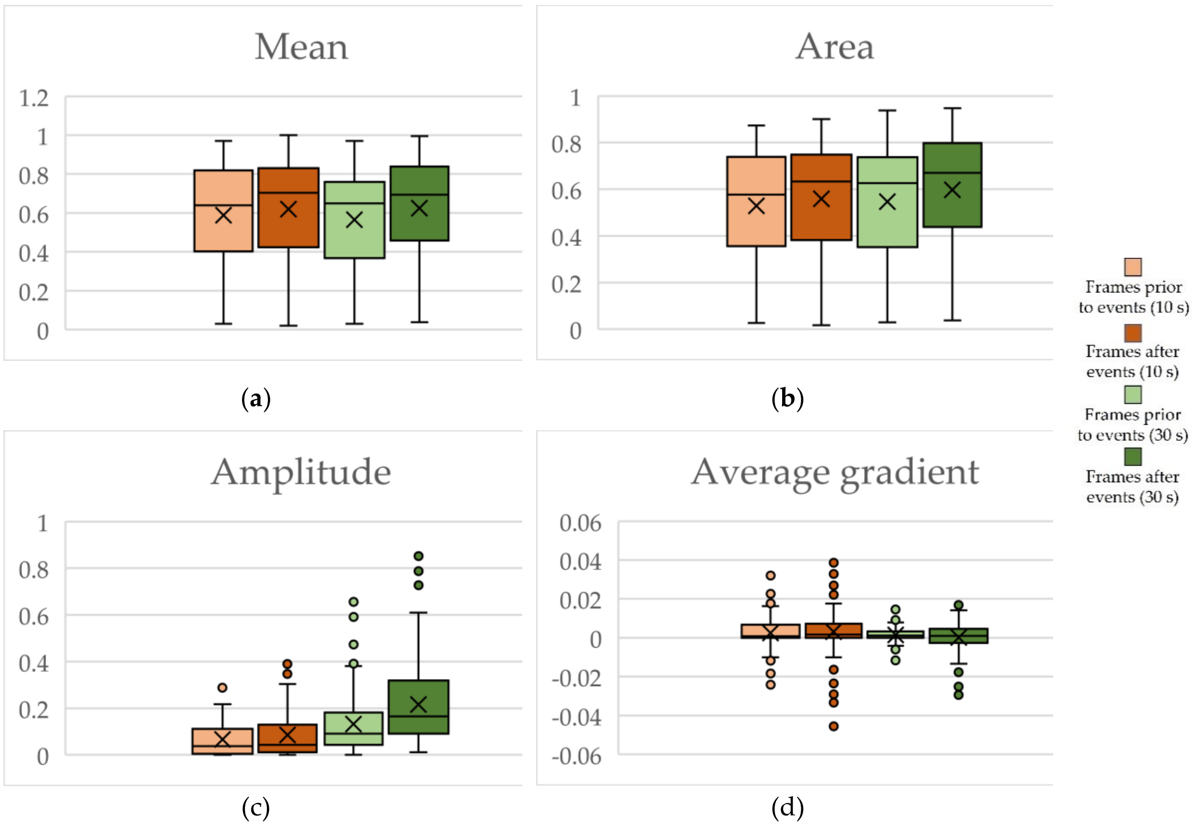


Figure 4-21: Whisker boxplots for reactive frames, for (a) mean value of perceived engagement, (b) area of annotation trace, (c) amplitude of engagement, (d) average gradient of engagement.

Contingency tables including the percentages of postures for reactive observation frames are presented in Figure 4-19a-d. The percentage of transitions is presented in Figure 4-20, with participants demonstrating higher seated mobility in frames after in-game events for both investigated frame durations. However, no statistically significant changes were identified in both cases. Whisker boxplots for features extracted from the annotation traces of perceived engagement for reactive frames are presented in Figure 4-21.

A statistically significant increase of +5.6% and +5.5% was observed in μA (Figure 4-21a) and $\int A$ (Figure 4-21b), respectively (two-sided Wilcoxon: P < 0.01), for 10 s observation reactive frames. The corresponding increases for 30 s frames were up to +10.4% and +9.5%, respectively (two-sided Wilcoxon: P < 0.01). A significant increase in $\hat{A}$ (Figure 4-21c) was also present in 30 s observation reactive frames (two-sided Wilcoxon: P < 0.01). Changes observed in average gradient (Figure 4-21d) were not found to be statistically significant.

For the second part of the analysis, the predictive capability of sensor and in-game features (Table 3), along with the multimodal feature V, towards features of perceived engagement are presented in Tables 8, 9, 10. The amplitude $(\hat{A})$ of the annotation trace presents the most cases of statistically significant correlation with sensor and in-game metrics features, thus highlighting its capacity to represent the hypothesized ground truth independently of type and duration of the observation frame. In particular, a negative correlation was observed with average mouse clicks (μMc) and variability of inter-beat intervals (ΔH) for 10 s reactive frames. Additionally, significant, and highly significant positive correlations were observed with the amplitude of heart beats per minute ($\widehat{H}$), and voting (V) for both 30 s reactive and continuous frames. Finally, a highly significant positive correlation with mouse movement (μMm) and a significant correlation with the average gradient of pressure sensors (ΔT) were evident in 30 s reactive frames. The mean value (μA) and the area ($\int A$) of annotation trace were correlated significantly with V, in 10 s and 30 s reactive frames, with both in-game metrics features in 30 s reactive frames, and ΔT in 10 s reactive frames. The average gradient of the annotation trace (ΔA) did not display significant correlations with any sensor or in-game metrics feature.

Table 8: Correlation of annotation features and features extracted from posture sensors, heart rate sensor and in-game metrics for 10s reactive frames. Significant values (P<0.05) are depicted in bold. Highly significant values (P<0.01) are denoted by (*).

Annotation features	Reactive frames (10s duration)					
	μMc	μMm	ΔT	$\widehat{H}$	ΔH	V
μA	0.00	0.20	**0.34***	0.00	0.12	**0.26**
$\int A$	-0.01	0.19	**0.33***	0.01	0.14	**0.27**
$\hat{A}$	**-0.29**	0.13	0.00	0.06	**-0.33***	-0.13
ΔA	-0.07	0.04	0.01	0.16	-0.11	0.15

Table 9: Correlation of annotation features and features extracted from posture sensors, heart rate sensor and in-game metrics for 30s reactive frames. Significant values (P<0.05) are depicted in bold. Highly significant values (P<0.01) are denoted by (*).

Annotation features	Reactive frames (30s duration)					
	μMc	μMm	ΔT	$\widehat{H}$	ΔH	V
μA	**0.28**	**0.51***	0.01	0.01	-0.12	**0.25**
$\int A$	**0.30***	**0.49***	0.03	0.04	-0.10	**0.28**
$\hat{A}$	-0.05	**0.31***	**0.26**	**0.24**	0.08	**0.33***
ΔA	0.06	0.18	0.11	0.13	-0.21	0.07

Table 10: Correlation of annotation features and features extracted from posture sensors, heart rate sensor and in-game metrics for continuous frames. Significant values (P<0.05) are depicted in bold. Highly significant values (P<0.01) are denoted by (*).

Annotation features	Continuous frames					
	μMc	μMm	ΔT	$\widehat{H}$	ΔH	V
μA	-	-	-0.05	-0.11	-0.05	-0.25
$\int A$	-	-	-0.05	-0.11	-0.05	-0.25
$\hat{A}$	-	-	0.25	**0.42***	0.18	**0.42***
ΔA	-	-	-0.12	0.14	-0.26	-0.07

Results from the first part of the investigation for real-time recognition of engagement are in accordance with those reported in the literature [113], indicating that the assumed sitting postures, along with the transitions between them and the overall seated mobility are associated with engagement as perceived by the player. Associations pointing in that direction were identified in both investigated types of observation frames. In continuous frames, the significant increase observed in the mean value (μA) and area ($\int A$) of the annotation trace from Tutorial to Gameplay was accompanied by a significant shift in assumed positions. Percentage of postures that include laying on the back of the chair was significantly lower in Gameplay than in Tutorial, with many participants leaning forward as engagement increased and game interaction intensified. This shift was also present on overall mobility (μT) during continuous observation frames, but no statistical significance was observed. From Gameplay to Review, sitting postures activating the back of the chair appeared to increase in frequency, accompanied by a decrease in perceived engagement (μA and $\int A$). However, these changes in assumed postures were not found to be statistically important. In contrast, a significant increase in perceived engagement from Review to Mission-2 was not accompanied by a significant change in identified postures. This may be in part because Review has rather short duration, in comparison to Tutorial and Gameplay. Additionally, the observed increase in perceived engagement between Review and Mission-2 was smaller than the one from Tutorial to Gameplay. These findings were consistently present in reactive observation frames. The investigated in-game events, for both 10 s and 30 s frame duration, resulted in significant increase in three extracted features (μA, $\int A$ and $\hat{A}$) of the annotation trace of perceived engagement. Furthermore, the frequency of postures activating the back of the chair was lower in reactive frames following in-game events, being of statistical significance only for the case of 30 s reactive frames. Finally, an increase of overall player mobility (μT) was also evident after in-game events, but no statistical significance was observed.

These findings indicate that an increase in perceived engagement is accompanied by the tendency to leave the back of the chair and an increased overall seated mobility (μT). The significance of these observations appears to be affected by the duration of the observation frames, with increased time yielding more significant results in both continuous and reactive frames. These observations are in agreement with findings of other studies that have employed different types of interaction tools [113], [201]. The presented results can be a steppingstone for the development of systems for real time recognition of engagement during SG play in office and home settings. In this direction, the identification of a general sensor activation threshold for posture monitoring, as proposed in the present study, is important. To this end, further investigation is necessary regarding the robustness of the threshold activation threshold across the BMI spectrum. Additionally, postures observed and identified during the intervention have not been particularly relaxed. Even though the participants have been instructed to get comfortable, this appears to be hindered by their presence in a research setting. This is expected to affect the type and number of postures assumed in other settings. A single and quite standard office chair has been employed during the present study. Desks in home settings tend to include chairs that vary greatly in size and comfort. The importance of data collection from home settings is, thus, highlighted and expected to provide more reliable results.

The second part of the analysis for real time recognition of engagement has identified significant predictive capacity of both sensor-based sources and in-game metrics towards player perceived engagement, as reflected by their significant correlation with features from the annotation trace (amplitude ($\hat{A}$)), mean value (μA), and area ($\int A$)). No significant correlation with the average gradient (ΔA) of the annotation trace has been found, despite its reported efficiency and robustness [183]. This may be attributed to the different data sources employed in this study and the larger duration of the examined observation frames. Reactive frames appear to produce the majority of features with significant predictive value, with 30 s frames revealing the most significant correlations. Features from sensors and in-game metrics present a range of significant and highly significant correlations (absolute values in the range 0.26 – 0.51) with perceived engagement, across all types of observation frames. However, an overly superior, in terms of robustness and effectiveness, unimodal feature could not be identified. On the contrary, the generated multimodal feature (V) was consistently found to be significantly correlated with features of the annotation trace.

In summary, data collected from affordable and unobtrusive sensors, assisted by in-game metrics features, appear to hold predictive value towards the hypothesized ground truth. Nevertheless, the presented analysis has investigated these features' predictive capability in a linear fashion. Supervised ML techniques can be employed to assess the features' potential to accurately recognize engagement in a non-linear fashion. Such methods can be incorporated in PCG, as part of a constant feedback loop that enhances adherence to SG-based health interventions by maximizing player engagement through generated game content. The suitability of the multimodal feature V needs to be further validated through advanced feature fusion techniques via ML. Deep learning methodologies, along with larger datasets, can be employed towards this investigation, given their increasing popularity in the field of multimodal affective recognition. Issues related with gender representation in participants should also be investigated towards identifying potential impact of gender imbalance on our core findings.

5. Procedural content generation in serious games for health

As described in the second chapter, PCG is a term that involves techniques incorporated in games to empower user engagement and increase replay value by automatically generating new content based on user choices and interaction with the game [202]. Data gathered from a variety of sensors can also be integrated into PCG techniques. With the integration of such types of data in the PCG workflow, the content provided is usually relevant to the desired SG goal, while also maximizing user engagement and thus adherence to the intervention. In general, SGs for health benefit greatly from the incorporation of state-of-the-art sensing technologies [129]. To this end, PCG techniques employing sensing data can produce patient-tailored and clinically relevant content, resulting in a smart personalized health intervention that can be applied towards raising awareness, prevention, diagnosis, monitoring, and treatment. More specifically, the availability of real-time sensing data providing information regarding the user's lifestyle, behavioral habits and health status, makes feasible the development of sensor based adaptive SGs with increased capacity to address important challenges in self-health management [3], [130]. Moreover, current research on PCG in SGs for health highlights the potential of the technology to generate game content automatically and on-demand, thus reducing the time and effort needed for design purposes and increasing replayability [203], [204].

To investigate the potential of the proposed conceptual framework for adaptivity in SGs (Fig. 5-1) firstly, an experimental process with WuF was designed and implemented. WuF, as described in Chapter 3, is a SG that promotes disease self-management for OSA and raises awareness for the condition employing card game and open world mechanics. Results from this experimental process were presented in [125]. The process consisted of a comparative analysis between the game experience of participants playing three versions of WuF, each incorporating different systems for the generation of NPCs. Each participant would play two versions of the SG without being aware of any differences between them. Analysis was conducted on data produced by game experience questionnaires. Secondly, data were collected from a pre-pilot study that included children with obesity playing the SG that was part of the ENDORSE platform for a period of 12 weeks. Results from this pre-pilot study were presented in [150]. The SG aims to enhance disease self-management capabilities in children with type-1 diabetes and/or obesity through mini games and educational messages. Both procedures involved the interaction of players with two novel SGs that were designed to promote disease self-management in chronic conditions (OSA, type-1 diabetes mellitus, and childhood obesity). Both SGs incorporated the PCG technique based on the proposed GA approach, that generates and promotes game content relevant to the players personalized needs and weaknesses. In the first case the PCG technique produces game content based on player interaction. In the second case the PCG technique is generalised to include data from sensors monitoring lifestyle information such as daily steps and sleeping quality, along with data collected from objective reports. Finally, both SGs incorporated a dynamic difficulty adjustment system that worked in conjunction with the PCG technique. This system changed game difficulty according to player progress and efficiency and worked towards maintaining the player in a state of flow that would promote engagement. The experimental procedures aimed at evaluating the validity of the generated content and player acceptance in relevance to PCG technique. As shown in Fig. 5-1 these procedures involve three of the four spaces of the proposed conceptual framework, the SG, the PCG space, and the health related data space. In the first process the SG and PCG space are implicated, whereas all three spaces are involved in the pre-pilot study.

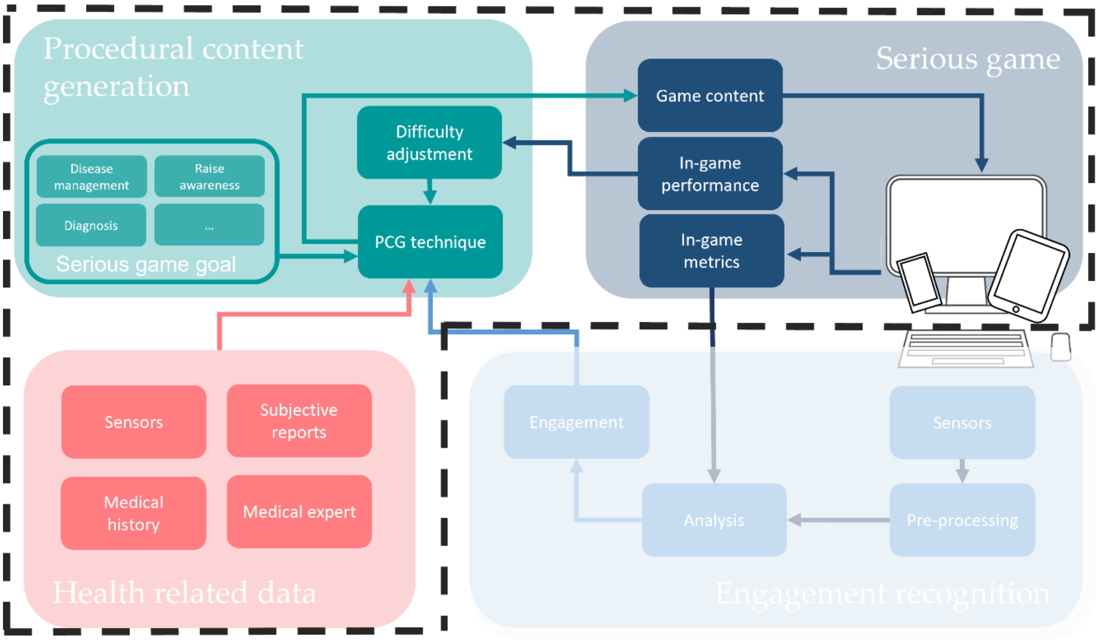


Figure 5-1: Conceptual framework for procedural content generation.

5.1 Incorporation of the genetic algorithm technique in a SG for obstructive sleep apnea

The proposed PCG technique was incorporated in WuF, to empower the game's educational value and improve overall game experience. The proposed technique is based on a GA and is responsible for the automated generation of new NPCs, based on user choices and game progress. The resulting adaptive SG possesses the ability to automatically adjust difficulty levels through a rule-based system and present educational content tailored to the user's needs. An initial population, with NPCs as individuals that are generated in a random manner, is set as the starting generation. After every debate card battle a new generation of NPCs is generated as offspring from a selection of the fittest individuals of the previous generation. For each card battle an NPC is chosen at random from the fittest chromosomes of the current generation to serve as an opponent.

For the representation of the NPCs in the GA, their available characterizing attributes are joined in a string as genes to form chromosomes. In this way, each chromosome features several genes with binary values, "1" represents presence of the particular attribute in the NPC's profile and "0" absence. An example of a chromosome characterizing an NPC with seven attributes is shown in Table 11. As a result, the number of expressed genes is proportional to the difficulty of the resulting debate, as every gene is translated in an enemy card that must be defeated. Furthermore, knowledge domains about OSA are linked to the chromosome genes in the form of attributes, constituting the core of the SG's educational content. In this manner, the GA adjusts game difficulty dynamically and presents tailored educational content by selecting the fittest individuals to produce offspring, based on fitness scores calculated after each debate.

Table 11: An example chromosome featuring seven non-player character attributes linked to obstructive sleep apnea.

Smoking	**Alcohol**	**Medication**	**Sleeping Position**	**Obesity**	**Hypertension**	**Depression**
S	A	M	SP	O	H	D
0	1	0	0	1	1	0

For the selection of the fittest chromosomes two fitness functions are employed, namely, the Winning Fitness Function (WFF) (Eq. 3) and the Losing Fitness Function (LFF) (Eq. 4):

$$WFS = W_a * a + W_b * b + \ldots + W_x * x \qquad (3)$$

$$LFS = L_a * a + L_a * b + \ldots + L_a * x \qquad (4)$$

After every debate, the fitness scores (WFS, LFS) are calculated for every chromosome of the current population. The parameters (a, b,…, x) represent the binary values of the genes in each of the generation's individuals. The parameters W_x and L_x constitute weights with initial values of zero, which are trained during each debate battle according to the following rules:
Training rules for W_x:

- $W_x = W_x - 1$, ∀ particular attribute of the NPC opponent.
- $W_x = W_x - 1$, ∀ x linked to player card used correctly.
- $W_x = W_x + 1$, ∀ x linked to player card used wrongly.
- $W_x = W_x + 1$, ∀ x linked to player card not played.

Training rules for L_x:

- $L_x = L_x + 1$, ∀ particular attribute of the NPC opponent.
- $L_x = L_x + 1$, ∀ x linked to user card used correctly.
- $L_x = L_x + 1$, ∀ x linked to user card used wrongly.
- $L_x = L_x - 1$, ∀ x linked to user card not played.

Weights linked to attributes that are subject to player interaction with corresponding cards are updated continuously during play in this manner, and the prevalence of corresponding attributes in future games is determined based on the need of the player to be exposed to the educational content. If the player wins a debate, the highest WFS scores are employed to select the fittest individuals of the current generation, benefiting mostly NPCs with never-before-seen attributes, or attributes associated with wrong player choices. If the player loses the current debate, the highest LFS scores are employed to determine the fittest individuals among the current generation. This way, the player who suffers losses will face opponents with higher likelihood of similar attributes, in an effort to revisit educational content that they do not seem to comprehend. In this incorporation of the PCG technique the GA creates new generations after each debate, providing a constantly changing game space. In addition, due to this constant flux in the game space, the weights of the fitness functions need to be re-initialized after each new generation in order to avoid rapid saturation of content and an abrupt increase in game difficulty by punishing of player mistakes. This leads to a somewhat memory-less PCG approach that bases decisions solely on the last game. This was found fitting for the short nature of the gaming sessions carried out during this experimental process. However, in longer and less frequent gaming sessions, an approach that would update weights after each debate and produce new generations at specified time periods would be more appropriate. This approach was adopted in the incorporation of the PCG in the SG for children with type 1 diabetes and/or obesity.

For the production of a new generation after every card battle, the fittest chromosomes form pairs. Each pair produces two offspring for the next generation, by exchanging genes based on a random crossover point. A number of fittest NPCs is selected so that every population retains its size across generations. In the resulting generation, every NPC has a small chance to be mutated, changing a random gene from "0" to "1". The mutation mechanism enhances the variability in the resulting populations. Afterwards, an NPC from the resulting generation is picked as an opponent for the next debate. The difficulty of the SG is automatically adjusted by a rule-based ELO ranking system that selects NPCs with a specific number of attributes. The number of the selected attributes is defined according to player losses and victories in the previous debates, that increase or decrease their ELO score respectively. The number of attributes is then selected according to preset thresholds.

5.1.1 Experimental process

To perform an initial evaluation of the proposed PCG technique, in terms of user experience, a blind experiment from the player perspective was implemented. Three versions of the SG were employed, version A, B and C. In each version, the user would play in five iterations of the debate card game, versus different NPCs. A pool of seven possible attributes was available for NPC generation (Table 11). A starting population of twenty NPCs, each characterized by exactly three attributes, was generated randomly for each of the three versions. After every card battle, the five fittest NPCs were selected to produce offspring for the next generation by pairing in all possible ways. In-game dialogues presented a basic tutorial to the player at the start of each version, explaining the rules of the SG and illustrating a simple presentation of the user interface. All versions were visually indistinguishable to the player and the participants were not informed of any differences between SG versions they played.

Version A incorporated the GA technique and was designed to deliver a punishing experience to the player. The generated content was affected greatly by in-game mistakes, and the player was presented constantly with enemy cards affiliated with these mistakes. To this end, the weights of WFF and WLF were not re-initialized after each debate to provide the algorithm a memory of the player's progress. In addition, the ELO system for dynamic difficulty adjustment was simplified to a more unforgiving version. In particular, NPCs with one additional attribute were selected from the resulting generation with each win and one less for each loss, to a minimum of two attributes and a maximum of five attributes.

Version B incorporated the GA technique and was designed to deliver a smoother learning experience. Weights of the WFF and the WLF were re-initialized after each battle, in order to provide the desired adaptivity within small gaming sessions. In this way, player mistakes were not exploited by the algorithm to entirely favor undesirable content. For difficulty adjustment, a starting rating value of "1" was assigned to the player. This value increased by one with each win and reduced by one with every loss. This approach delivered smoother difficulty adjustment without being transparent to the player. NPCs were chosen randomly among the fittest chromosomes from the current generation according to the following rules:

- Rating between 1 and 3: 3 attributes
- Rating above 3 (max 5): 4 attributes
- Rating below 1 (min -2): 2 attributes

Version C generated the attributes characterizing randomly the NPCs. The number of attributes of the NPCs was also chosen randomly, between 2 and 4, with the exception of the NPC of the first battle, who was characterized by exactly 3 attributes to provide a similar starting point to the other two versions. This version was designed to serve as a point of comparison between version A and B incorporating the GA-based PCG technique.

A total of 42 participants were recruited for this preliminary validation process. Participants were split in two groups, with group 1 interacting with versions A and C, and group 2 with versions B and C. Both groups played their corresponding versions, one on each of two consequent days. The participants were not informed of any differences between the two versions, and the order in which they played each version was chosen randomly. To measure user experience, two modules of the Game Experience Questionnaire [182] were deployed after each session. The core module provides insight about competence, sensory and imaginative immersion, flow, tension, challenge, negative affect and positive affect. The Post-game module asserts positive experience, negative experience, tiredness and returning to reality. Paired sample Student's t-tests were applied on the results to investigate for statistically significant differences between scores obtained from the questionnaires. Information about the participants is presented in the table below.

	All (N=42)	Group 1 (N=23)	Group 2 (N=19)
Gender	male (25), female (17)	male (12), female (11)	male (13), female (6)
Age	27.90 ± 4.93	29.56 ± 4.77	25.89 ± 4.35

5.1.2 Results and discussion

Scores of the GEQ are displayed in Table 12, for groups 1 and 2, respectively. Results display differences in terms of user experience between the adaptive and non-adaptive versions of the SG. Analysis of the results revealed a statistically significant decrease in competence ($t(22)=2.46$, $p=0.021$), and an increase in challenge ($t(22)=-2.48$, $p=0.020$) and negative experience ($t(22)=-3.42$, $p=0.002$), between versions C and A respectively, played by the first group. On the other hand, the feeling of competence increased significantly ($t(18)=-2.30$, $p=0.033$) and negative experience dropped ($t(18)=2.53$, $p=0.020$) between versions B and C respectively. The remaining dimensions of the GEQ did not display statistically significant differences ($p>0.05$) between game versions for both groups.

Table 12: Results in terms of game experience for groups 1 and 2.

GEQ (Scale 0-4)	Mean ± Standard Deviation								
	Group 1		Group 2			Group 1		Group 2	
	Version A	Version C	Version B	Version C		Version A	Version C	Version B	Version C
	Core Module					Post-game Module			
Competence	**2.03** ± 0.87	**2.46** ± 0.64	**2.45** ± 0.71	**2.13** ± 0.84	Positive experience	1.29 ± 0.86	1.50 ± 0.71	1.53 ± 0.86	1.49 ± 0.78
Immersion	2.21 ± 0.84	2.14 ± 0.66	2.28 ± 0.72	2.25 ± 0.78	Negative experience	**0.27** ± 0.30	**0.12** ± 0.24	**0.02** ± 0.08	**0.11** ± 0.17
Flow	1.88 ± 0.95	1.69 ± 0.86	1.60 ± 0.84	1.67 ± 0.95	Tiredness	0.06 ± 0.22	0.04 ± 0.20	0.13 ± 0.35	0.13 ± 0.31
Tension	0.47 ± 0.57	0.31 ± 0.45	0.15 ± 0.27	0.21 ± 0.43	Returning to reality	0.55 ± 0.56	0.46 ± 0.40	0.36 ± 0.56	0.47 ± 0.53
Challenge	**1.31** ± 0.50	**0.95** ± 0.42	0.84 ± 0.36	1.04 ± 0.44					
Negative Affect	0.84 ± 0.89	0.59 ± 0.50	0.39 ± 0.35	0.47 ± 0.42					
Positive Affect	2.34 ± 0.67	2.37 ± 0.75	2.45 ± 0.76	2.53 ± 0.77					

A small set of participants (N=10) were chosen randomly from both groups, to take part in semi-constructed interviews after their game sessions. They were asked if they became aware of differences between the versions of the SG they interacted with and provide feedback. All of the participants stated that while game difficulty was clearly different between the versions they played, they were not certain of other differences or the mechanics governing difficulty adjustment. Additional comments involved concerns about UI elements, the tutorial and the reporting of game bugs that took place during playing.

Initial validation of the proposed PCG technique incorporated in WuF indicated that game experience differed between the deployed SG versions. Version C, lacking any adaptive features, was used as a benchmark to test the effect of the other two versions on participant experience. Versions A and B were implemented to evaluate the potential of the proposed technique from different perspectives. Version A punished successful progress by increasing difficulty to very high levels. To achieve this, the weights of the fitness functions were not re-initialized, hence NPC attributes linked to user errors were promoted by the GA perpetually to the following generations. Additionally, the number of NPC attributes increased with every win. The effects of this harsh difficulty adjustment were evident in the responses of the GEQ questionnaire of group 1. Sense of competence was significantly lower in version A, while negative experience and challenge was significantly higher than those observed in version C. On the other side, version B was designed to provide a smoother experience which adapted to player progress. The weights of the fitness functions were re-initialized after each debate, due to the short length of the game session. This feature allowed for a wider variety of attributes to be promoted to next generations, resulting in more heterogeneous populations of NPCs and less content saturation. In this way, the user would face both NPCs that would share attributes linked to past mistakes and NPCs characterized by new attributes, resulting in more engaging experience. Furthermore, the number of attributes characterizing the NPCs scaled slower according to player wins and losses, avoiding thus sudden spikes in game difficulty. As a result, questionnaire answers obtained from group 2 are consistent with design considerations of version B, as sense of competence was significantly higher compared to version C, and negative experience was significantly lower. In contrast, sense of challenge was not found to be significantly different between versions B and C. This could point towards the fact that the proposed technique has the potential, if applied properly, to promote the perception of user competence without altering perception of difficulty. Results from this preliminary validation of the proposed PCG technique indicate its potential to enhance user experience while delivering tailored educational SG content. This observation was possible, despite the limited content the deployed versions of the SG and the relatively short duration of the session. Additionally, chromosomes in the proposed GA consisted of only seven genes, while only five generations of NPCs were produced in each playthrough.

5.2 Incorporation of the genetic algorithm in a SG for type 1 diabetes and obesity

The proposed PCG technique was also implemented in the ENDORSE platform in order to control game content based on interaction and sensor data. The platform's recommendation system employs a version of the proposed PCG technique that has been expanded from its initial application. In particular, for the incorporation of the PCG technique in the platform, SG missions and displayed messages have been translated into genes containing binary values, with the value of "1" signifying content presence. This collection of game content-oriented genes constitutes a chromosome, as was the case with the incorporation of the PCG technique in WuF, while each gene is represented in the chromosome in a designated position. As depicted in Figure 5-2, chromosomes in this case have been split into two sections. The first part (MG) is comprised of genes that designate the game missions that are available every day. On the first daily login in the SG, the application would send a pull request from the platform's recommendation system and receive the appropriate content. The second part (RG) selects appropriate messages to be displayed to the platform's end-users. During the start of the intervention, the GA is initialized with a population of chromosomes, each containing a selection of genes corresponding either to the availability of a mission or the appearance of a specific message. A total of 380 chromosomes are created for the starting population to provide sufficient diversity among the chromosome population. A five percent probability of mutation per gene has been applied to adjust for cases with limited diversity. Furthermore, the length of the chromosomes has been significantly enlarged in comparison with first incorporation of the PCG technique, containing 103 genes, 28 of them used for mission content and 75 for messages. Constraints have been applied to the chromosomes to limit the availability of daily game missions, based on expert recommendations for maximum daily screen time. In particular, there are no chromosomes allowed containing more than two missions per day, while each day has one educational and one action mission.

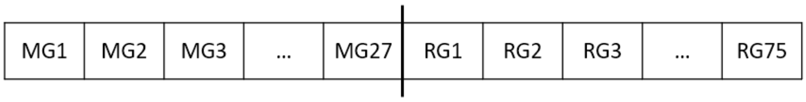


Figure 5-2: Chromosome formation, with genes controlling missions (MG) and genes controlling messages (RG).

On the first generation, a chromosome is selected randomly as a representative and controls the platform's content for the first week. From then on, based on data collected from sensors and user interaction, a fitness function (Eq. 5) is applied to determine the fittest chromosomes from the initial population.

$$FS = W_{MG1} * MG1 + \cdots + W_{MG28} * MG28 + W_{RG1} * RG1 + \cdots + W_{RG75} * RG75 \qquad (5)$$

The fitness function employs weights (W_{gene}), assigned to each gene. Weights are trained to be analogous to the desirability of the gene's presence in the next GA state. Weight training is based on relevant data collected by the platform. Weights for game missions (MG) are trained based on data collected through the interaction with the SG:

- W_{MG} = $W_{MG} + a_1$, if mission score is low
- W_{MG} = $W_{MG} - a_2$, if mission score is high
- W_{MG} = $W_{MG} - a_3$, if mission is played

Weights controlling the appearance of messages are trained based on the heterogeneous data collected from the platform's sensors. For physical activity and sleep monitoring, the Fitbit ACE 2 (for kids) was employed with the ability to track steps, calories and sleep (duration and quality of sleep). For diabetes self-management monitoring, the Freestyle Libre and Medtronic continuous glucose measurement sensors and the Insulclock smart insulin pen were employed. Specific thresholds have been applied on the obtained sensing data towards classifying every day as good, medium, or bad, in terms of recognition of healthy lifestyle habits and self-health management. For educational messages:

- W_{RG} = $W_{RG} - b_1$, for each good day
- W_{RG} = $W_{RG} - b_2$, for each medium day
- W_{RG} = $W_{RG} + b_3$, for each bad day

For progress messages:

- W_{RG} = $W_{RG} + b_1$, for each good day
- W_{RG} = $W_{RG} - b_2$, for each medium day
- W_{RG} = $W_{RG} + b_3$, for each bad day

This weight training ensures that the GA promotes content that is associated with identified player needs, replicating it further and passing it on to future generations. Weights were not re-initialized in this application of the GA. SG missions where the player scores low are more frequently promoted over missions that are either played frequently or are completed with high game scores. Additionally, educational messages linked to healthy behaviors monitored by the platform's sensors have a lower chance of appearing. Progress messages are more likely to appear to promote healthy habits or remind about identified unhealthy tendencies.

At the end of each GA iteration, the fittest chromosomes are paired to produce the new generation. The twenty highest scoring chromosomes are selected, based on the score provided by the fitness function, to create the next generation of chromosomes through the crossover. One of those is also chosen to be the representative, responsible for providing the system's content. Crossover (Fig. 5-3) occurs separately for mission game content and the messages (MGs and RGs), to ensure the structural integrity of the offspring and its compliance with the imposed constraints.

(a)

(b)

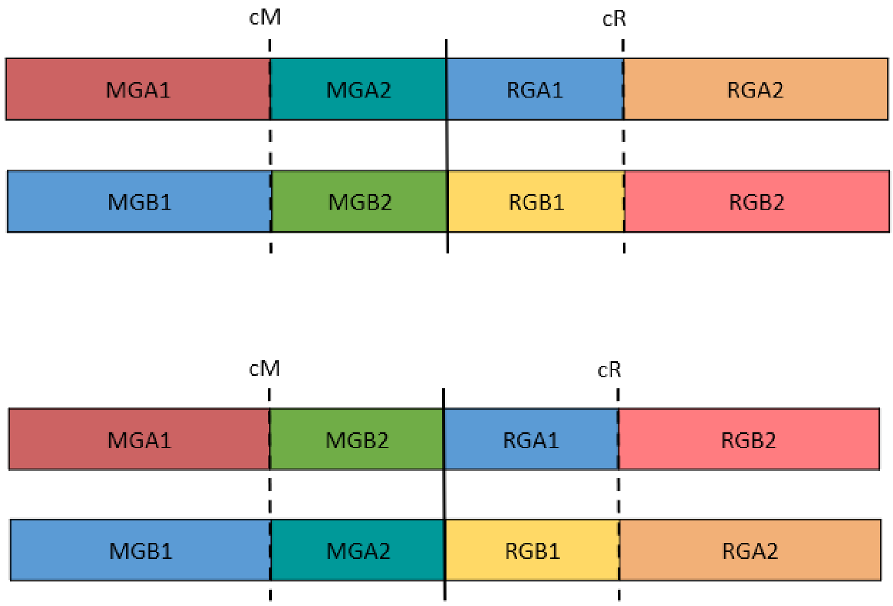


Figure 5-3: Chromosome crossover. (a) Two random points, cM and cR, split the chromosomes. (b) Two new gene sequences for MG and two for RG are generated.

Algorithm 1: Genetic Algorithm Update [150]

```
Data:    Previously selected chromosome CH
         Previous generation Gen, its difficulty D and its weights W,
         user interaction with the platform UI and data from activity
         tracker UD
Result:  Newly selected chromosome CH

function NewGeneration(CH, Gen, D, W, UI, UD)
  previous MG, previous RG ← split CH
  # region Serious Game
  D ← CalculateNewDifficultyMG(MG, D, UI)
  W ← CalculateNewWeightsMG(MG, W, UI, UD)
  Gen ← CrossoverMG(Gen, W)
  Gen ← MutationsMG(Gen)
  new MG ← SelectChromosomeMG(Gen, W)
  if  new MG same as previous MG
     Gen ← InitialisePopulationMG(W)
     new  MG ← SelectChromosomeMG(Gen, W)
 # endregion
 # region Messages
  W ← CalculateNewWeightsRG(RG, W, UI, UD)
  Gen ← CrossoverRG(Gen, W)
  Gen ← MutationsRG(Gen)
  new RG ← SelectChromosomeRG(Gen, W)
  if new RG same as previous RG
     Gen ← InitialisePopulationRG(W)
     new  RG ← SelectChromosomeRG(Gen, W)
  new RG ← ApplyRGMask(new RG)
  # endregion
  CH ← concatenate new MG and RG
  CreateNewSave(CH, Gen, D, W)
  return CH
```

For each pair of chromosomes, the MGs and RGs are split independently into two segments at a random point. Then each segment of the chromosomes is combined with that of another fit chromosome to create two sets of new MGs and RGs. Finally, the MGs and RGs are combined in four ways resulting in four total offspring.

Algorithm 1 summarizes of the update process the GA for the next generation. The algorithm receives, for every user, as input data from the previous GA state as well as relevant data from platform interaction and data collected from the activity tracker. The output is the generation's representative chromosome. As described, the crossover occurs separately for MG and RG parts of the chromosomes and takes into account the previous generation's fittest chromosomes. If the newly selected representative RG or MG is identical to that of the previous generation, the chromosome population is deemed saturated and is reinitialized while keeping the same weights. Then the representative is selected from the newly generated fittest chromosomes. This process facilitates the addition of genetic diversity to the population, in case the gene mutation isn't enough.

5.2.1 Experimental process

The proposed PCG technique incorporated in the platform's recommendation system was evaluated in terms of user acceptance, and accuracy of the generated tailored content during a pre-pilot study. Twenty children (aged 6 - 14) suffering from obesity were recruited in the pre-pilot study for a period of 12 weeks. The SG and a Fitbit Activity Tracker [205] were employed by the children, while the platform's mobile application was used by their parents. The pre-pilot study was approved by the national ethical committee. The intervention included 12 weekly iterations. During each iteration, data relating to the children's physical activity were collected and fed into the GA's fitness functions. Simultaneously, data from the children's interaction with the content and missions of the SG were collected. One of the GA's highest-scoring chromosomes was designated as representative and provided the weekly content. The report produced by the Fitbit activity tracker included the daily number of steps, sedentary time, and sleep time. For every participant, the average steps per day were calculated for days deemed as active. An active day was defined as a day with at least two thousand steps. This cut-off was set as a threshold for including data from the Fitbit Activity Tracker in the daily training of the relevant chromosome weights. Similarly, the average sleep duration was collected, taking into consideration days with more than zero recorded minutes of sleep. Thresholds for categorizing days as good, medium or bad were appointed based on relevant literature [206], [207].

The incorporation of the PCG technique was evaluated in two directions. In the first direction, a post-intervention questionnaire was created including questions relevant to the acceptance (Q1-.Q2.) and usefulness (Q3) of the personalized content [172]. The children's parents would answer each question with a score ranging from 1 to 5. The questionnaires were applied after the completion of the intervention plan. In the second direction, Pearson's Correlation test was applied on the data collected from the activity trackers and the messages provided by the GA in order to assess PCG's ability to provide content relevant to the child's lifestyle status. Messages controlled by the GA were split into the following categories: "Physical Activity", "Sedentary Time" and "Sleep Duration". "Physical Activity" and "Sedentary Time" were correlated with average daily steps per participant and "Sleep Duration" with average sleep duration. Each of these categories was represented in the GA by two genes, one designated for educational messages and the other for progress messages. Pearson's correlation was calculated for these genes both separately and in combination, against relevant data collected from the Fitbit sensors.

5.2.2 Results and discussion

The scores obtained from questions relevant to the acceptance of personalized content are presented in Table 13. Overall, a positive score was awarded to all three questions. The highest score was achieved in Q2, investigating the usefulness of the delivered messages towards the achievement of personal goals (Q2: score 4.13). The lowest score was reported in Q1, regarding the relevance of the displayed messages to the child's needs, (Q1: score 3.07). A statistically significant difference (score 1.04, p=0.0217) was identified by applying the Student's t-test between questions Q1 and Q2, regarding the usefulness and relevance of messages displayed within the SG.

Table 13: Post-intervention questions for user acceptance.

Questions	Scores	
	Average Value	Standard Deviation
Q1: How relevant to your child's needs did you find the messages shown through the game?	3.07	1,38
Q2: How useful did you find the daily messages you received to achieve your goals?	4.13	1,06
Q3: How useful did you find the game's ability to display educational messages?	3.40	1,50

Results for Pearson's correlation are depicted in Table 14. The average steps per day were 9840.6 and the average sleep duration was 458.6 minutes Among all participants. Out of the 20 participants, there were two dropouts and three cases where responses to the post-intervention questionnaires were not provided. Negative correlations were observed across all of the investigated combinations. Statistical significance was found between the number of "Physical Activity" messages sent, and the average steps (p = 0.042), as well as the number of combined "Physical Activity" and "Sedentary Time" messages and the average steps (p = 0.030).

Table 14: Pearson's correlation between delivered messages and sensor data.

Category	Correlation		
	Gene 1	Gene 2	Combination
Physical Activity	-0.20	-0.30	**-0.45**
Sedentary Time	-0.17	-0.25	-0.29
Physical Activity and Sedentary Time	-	-	**-0.48**
Sleep Duration	-0.23	-0.11	-0.19

Answers to the post-intervention questionnaires indicate an overall acceptance of the personalized content generated by the GA. A statistically significant difference was discovered between message relevance and usefulness. Based on the answers provided to Q1 and Q2, the messages sent by the GA were regarded as more useful than relevant. This may imply that due to the fact that the messages were designed with the help of healthcare professionals specifically for the intervention provided by the SG, they inherited general usefulness to the health-related needs of the recipients. On the other hand, the message's appearance in the SG content was defined by the GA. Despite its ability to identify and generate relevant messages, a comparatively lower score on Q1 is to be expected, as one of the GA's primary objectives is the diversification of content to avoid saturation. Both scores, though, advocate towards the fact that the messages had an overall positive effect in the intervention provided by the SG.

The negative correlations presented in Table 15 display the GA's sensitivity to trends regarding lifestyle and self-health management habits monitored by the platform's sensors. When high levels of physical activity were observed, the frequency of related messages declined and vice versa. The fact that this particular pattern was identified in all categories, displays the overall success of GA's incorporation. The fitness functions managed to translate the observed trends into proper weight training, thus enabling the GA to promote the most suitable genes for its chromosomes in the following generations. Furthermore, the GA avoided reaching high values of negative correlations, a sign that would point towards content saturation. The non-deterministic selection of content performed by the GA provided the recommendation system with the capability to sometimes omit chromosome genes with high weights. This behavior highlights the difference between a rule-based system and the GA, in terms of providing varying tailored content. These results indicate the generalisation capabilities of the proposed PCG technique in SGs of different genre, with the ability to control educational game content of different type. The inclusion of sensor data in the generation of game content appears to be successful. These findings are further substantiated by the preliminary analysis of data obtained from the following pilot study for children with obesity and pre-pilot and pilot studies for children suffering from type 1 diabetes. The ENDORSE platform incorporated the proposed PCG technique, while SG content was enhanced with more mini games and additional messages. The personalization capabilities, based on sensor data and interaction with the platform, afforded by the GA appear to produce an automated intervention space that can be as much effective as intervention tools that provide frequent communication with healthcare professionals. Further analysis on the data collected during these studies, that feature larger sets of participants and control groups, will provide insight towards the capabilities of the proposed conceptual framework.

6. Automated testing in serious games for health

In this chapter the process of developing deep reinforcement learning (DRL) agents that can navigate and play SGs for health automatically is presented. Automation denotes a machine with a self-contained principle of motion. Digital games, in contrast with physical games such as board games, always feature some level of basic automation, in the form of calculating game parameters, object collisions and distances, among others. However, AI-driven, non-human automated play is a different concept that gathers growing attention in the field of games. It includes a wide range of examples, such as autonomous NPCs, idling game worlds, non-human agents traversing multiplayer spaces, and smart self-learning agents [208]. Self-learning agents have been widely used to test digital games by automatically playing them [209]. Testing in video games is a crucial aspect that comes into play in almost the entirety of the production pipeline. It is very important in the ideation and design process, where it is usually conducted with the use of game prototypes. But is also of paramount importance during development and post-production to avoid the presence of bugs, exploits and broken game mechanics, which can prove to be detrimental to the game experience. Findings from a recent survey indicate that current testing processes rely mostly on manual playing and the tester's intrinsic knowledge [210]. However, this kind of testing is very expensive both in time and resources due to its lack of automation. Agents for automated playing can reduce the cost and time of applying manual testing. Various techniques towards automated playtesting have been investigated, such as active learning for game parameter tuning [211], deep player behavior models for balancing in multiplayer games [212], reinforcement learning for estimating player progress [213], and Monte-Carlo-Tree searches [214]. In addition, there are humanlike agents that are trained on data generated from actual players [209]. These agents are better suited towards analysing the difficulty of the game, generating playthroughs, becoming opponents for human players. Deep reinforcement learning (DRL) has also been applied to automate game testing [215]. DRL frameworks are capable of exploring game mechanics with increased test coverage, and can be used to find exploits, detect problems, and test game difficulty.

Automated testing can also be applied to SGs. The benefits mentioned above from this type of testing translate perfectly to the field of SGs. Moreover, automated play can provide valuable feedback about the SG's educational or behavioral purpose. However, few examples of automated testing based on some type of smart agent that explores the SG space and simulates playing and learning have been identified in the literature. In [216], a mathematical model helps track player game actions and present them by creatin a set of game states that simulate the gaming process in a SG about social engineering. In this approach player actions are not predicted, however the system records and structures them in a way that useful actions can be extracted and form context-specific questions about gaming sessions. In this manner the system can also identify problems in the game design that need to be addressed. In [217], a system that employs player data to evaluate a SG's educational efficiency automatically is proposed. The model uses game learning analytics, robotic process automation and statistical analysis as an automated post-test assessment tool. In [218], video analysis and data mining was employed to measure the implicit science learning during play with an educational SG. Specific game strategies were automatically detected in this manner and coded in a sample of 69 high school students. These attempts do not exactly constitute automated testing; however, they involve modeling of the game space and data driven approaches towards replicating playing.

The COVID-19 pandemic has also significantly impacted the application and validation of SGs. On one hand, SGs are tools that can be employed for remote and distance learning, and provide positive educational outcomes, enhanced engagement and motivation, through a safe and interactive asynchronous environment [219]. On the other hand, preliminary validation and testing of SGs in live settings, such as schools, hospitals, or research facilities, allows for hands-on observation and interaction with players. The widespread implementation of social distancing measures and lockdowns has made live testing difficult. As a result, there has been a shift towards remote testing methods, such as online surveys and virtual focus groups. While these methods hold some value, they also present their own challenges, such as difficulty in monitoring participant attention to the intervention and relying on subjective reports.

An approach to employ agents for automated testing, to gain insight about the incorporated PCG technique in WuF is presented below. To evaluate the impact of the PCG in the SG's educational efficiency, two versions of WuF that generate their content differently have been used. The first version is considered the control version, with randomly created opponents. The second version employs the proposed PCG technique to generate NPCs based on the player's interaction with the game. A conceptual framework has been developed, able to translate a SG for use in automated testing based on DRL. By using DRL to test WuF, it is possible to reduce the need for live testing and improve the efficiency of the testing process. One of the key advantages of using DRL for SG testing is the ability to simulate a wide range of scenarios and player behaviors. DRL allows for the creation of AI agents that can act as players and interact with the SG in a varied manner. This approach is motivated by the increasing demand for personalized and dynamic digital health interventions, such as SGs, and the need to efficiently design and evaluate these interventions. By training deep learning agents to play SGs, the aim is to automate the process of content generation and evaluation, and to assess the impact of dynamic game content on the speed and effectiveness of agent training. The goal of this methodology is to identify the key factors that influence the performance of deep learning agents as they play SGs, and to inform the development of more effective and personalized SG for health purposes. This can help provide a more comprehensive understanding of the game's effectiveness and identify potential issues that may not be evident in traditional testing methods.

6.1 Deep reinforcement learning in games

DRL is considered a huge step towards the creation of autonomous systems with a higher level of understanding of the environment they operate in [220]. In part this is one of the primary goals of AI techniques, to allow for fully autonomous agents that learn optimal behaviours and improve through trial and error. DRL is a subfield of ML that involves training AI agents to make decisions and take actions in a dynamic environment in order to maximize a reward signal. Reinforcement learning is a field of AI that is based on mathematical frameworks that operate on experience-driven autonomous learning [221]. The idea of these systems is based on the principle that interacting with an environment to achieve a goal, in turn leads to learning. In order for learning to occur the agent in question must be able to understand the state of the environment and take actions within it. These approaches however are often found to be lacking in scalability and are thus limited to low-dimensional problems. DRL, having the advantage of the powerful function approximation and representation learning properties of deep neural networks, is able to overcome these problems. DRL combines the use of neural networks with reinforcement learning algorithms to enable AI agents to learn through trial and error. DRL has been successfully applied to a wide range of tasks, including control systems, natural language processing, and robotics [222]. The key concept in DRL is the use of a reward signal to guide the learning process. An AI agent is trained to take actions in an environment in order to maximize this reward signal. The agent learns through trial and error, adjusting its actions based on the resulting rewards and attempting to find the optimal strategy for maximizing the reward. DRL algorithms can be classified into two main categories: value-based and policy-based. Value-based algorithms focus on estimating the value of each action at each state, whereas policy-based algorithms directly learn the optimal policy for selecting actions.

DRL plays an important role in game AI, by enabling the training of agents to automatically play complex single-agent and multi-agent games [223]. While there are still many challenges in the domain, there have been many cases where DRL agents have achieved super-human performance.

Games in general provide a safe and controllable virtual environment that generates an infinite amount of data, making them suitable for research in AI techniques. Due to these characteristics, DRL is gaining ground in its application in games. Deep Q-Networks have been combined with reinforcement learning, using video input, to play a multitude of old Atari 2600 games [224]. AlphaGo is the first automated play agent to beat a human champion of the game Go and combines tree search with deep neural networks [225]. The game of Go is considered highly complex, much more than chess, and applying automated play techniques has been an open problem for decades. Until 2016 the strongest available Go-playing AI could not defeat the best human Go player. DRL techniques have also been applied to play in games that include environments that are not 2D and fully observable to the agent. Architectures have been presented to tackle more demanding 3D environments in first-person shooters with results indicating that they are able to defeat build-in simpler AI agents [226]. DRL has also been applied to SGs, however, only applications that aim towards adaptivity have been identified. For example, an approach for interactive deep learning for adaptive gameplay that combines human player and trainer feedback in an effort to direct the learning process in SGs is presented in [227]. The incorporation of automated play through state of the art DRL agents in adaptive SGs for health is expected to facilitate testing of both the SG and the technique employed for adaptation. These approaches allow for testing in a repetitive and reproducible manner, without the need and effort included in human testing. Calibration of PCG and metrics that will provide insight regarding the effectiveness of the adaptive SG are going to be facilitated and in conjunction with human testing lead to an improvement in the health intervention provided.

6.2 Automated testing of a serious game for health

Evaluation of the proposed PCG technique based on a GA was also conducted by applying deep learning algorithms. Specifically, DRL agents were trained to play WuF autonomously. This enabled another means of evaluation of the incorporated PCG, fully automated and independent of the recruitment of human participants. To achieve automation and autonomy, WuF was generalized and translated in GDL (Game Description Language, Stanford University [228]), an environment that enables the DRL agent to interact with the SG through queries. This language allows for the representation of previously unknown games by giving information regarding their rules. To test the effectiveness of the two versions of the game, separate proximal policy optimization (PPO) agents have been trained on each version.

6.2.1 Conceptual framework for testing

To perform the testing of the SG in a standardized manner, a conceptual framework has been designed. It describes the process of translating a SG through adjustments that make it compatible with PPO agent training. The process of designing a SG is demanding, as it not only involves everything a regular game would require, but also demands careful planning and integration of the serious purpose. This is of paramount importance when dealing with mobile health applications, where the serious purpose and the limitations it applies often overshadows the rest of the game design process. To this end, it is crucial that the proposed automated testing pipeline is abstract in nature, to avoid interfering with the SG development. The framework comprises of three distinct parts (Fig. 6-1) that communicate with each other through interfaces. The first part is the SG and encapsulates the whole game space, including game logic, graphic user interface and serious purpose. The second part is the translation of the game mechanics to a format understandable by the DRL. The last part is the DRL itself. This architecture enables the whole process to be more versatile. In case the DRL model needs to be changed, only the translation needs modification. In the same manner, if a different SG needs testing, the translation is the only thing to be modified. For this modularity, the interfaces between the three parts play a crucial role. They act as means of communication and they standardize the way each part interacts with the others.

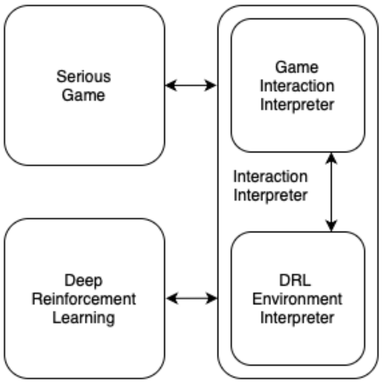


Figure 6-1: Conceptual framework for automated testing in serious games.

6.2.2 Serious game

The SG employed for automated testing, WuF, was presented in detail in Chapter 3, whereas the incorporation of the proposed PCG technique and its validation with human players was presented in Chapter 5. WuF is a SG aiming to raise awareness for OSA and improve self-health management of people suffering from it. The SG is mainly based on card game systems and mechanics but also features open world exploration. The incorporated PCG technique generates NPCs that are used as opponents in the card game battles. In WuF NPCs represent SG content with educational value that form the cornerstone towards achieving effective behavioral in the player. The player's ability to win in the simulated debate battles, as presented in WuF's conceptual framework, is directly linked to their knowledge about OSA and the desired SG outcome. In this manner, an assumption can be made, that an automated testing agent that learns to select appropriate cards and achieve victory approximates the human player learning process. According to this assumption the evaluation of the agent's capacity to navigate the SG space, learn to efficiently select and use game cards and achieve victory can be employed as a metric to evaluate the content that is produced through PCG.

In order to streamline the validation of PCG in the SG through automated play, a demo variant has been developed. This variant removes the open world exploration system from the game and presents the player with sequential debate battles against NPCs. These battles form the main system of the SG. As presented in WuF's conceptual framework, elements and game assets of the open world exploration also act as mediators towards the SG's desired outcome. However, almost the entirety of the educational content is present within the card game mechanics. Removing open-world content is expected to introduce bias regarding the validation through automated play agents. This bias, however, is expected to not impact the insights gained through this process greatly. The reason for the removal of the open world content is technical. The card mechanics are concise and simple allowing for an easy representation of the in-game state – action space. This is not the case with open world mechanics, as their representation in an understandable manner for the automated play agent would require extended design and vast resources, effectively defeating the purpose of conducting automated testing. Based on this variant, two SG versions have been created to be employed in comparative analysis in terms of automated agents' efficiency in playing them. The control version presents the player with randomly generated NPC profiles that feature a varying number of cards. This version is similar to Game Version C presented in Chapter 5. The test version employs the PCG technique to generate NPC tailored to the performance of the player, while simultaneously adjusting their difficulty. This version is similar to Game Version B presented in Chapter 5.

Figure 6-2 : Screenshot from the core game aspect of Wake Up for the Future.

The variant version of WuF can be effectively split in two major sub-games, the deterministic core game (Fig. 6-2) and the arbitrary meta game (Fig. 6-3). The core game consists of the card battles simulating debates between the player and NPCs. This sub-game is governed entirely by deterministic rules and its content is defined from start to finish in each iteration. Player cards have already been selected by the player, and the NPC card deck is formulated before the start of the game. Player actions are specific and finite, while the outcome is based solely on the current state of the game. In contrast, the meta game is governed by arbitrary rules that feature randomness and incomplete state of knowledge. The meta game consists of the part of the game which is outside the core game loop. It determines what manner of NPCs the player encounters, the attributes that describe these NPCs and player deck building. The player selects cards for their deck in a deterministic manner, with a large possible but finite number of possibilities available to them, however NPC deck formulation is governed either by the GA technique for PCG or by randomness.

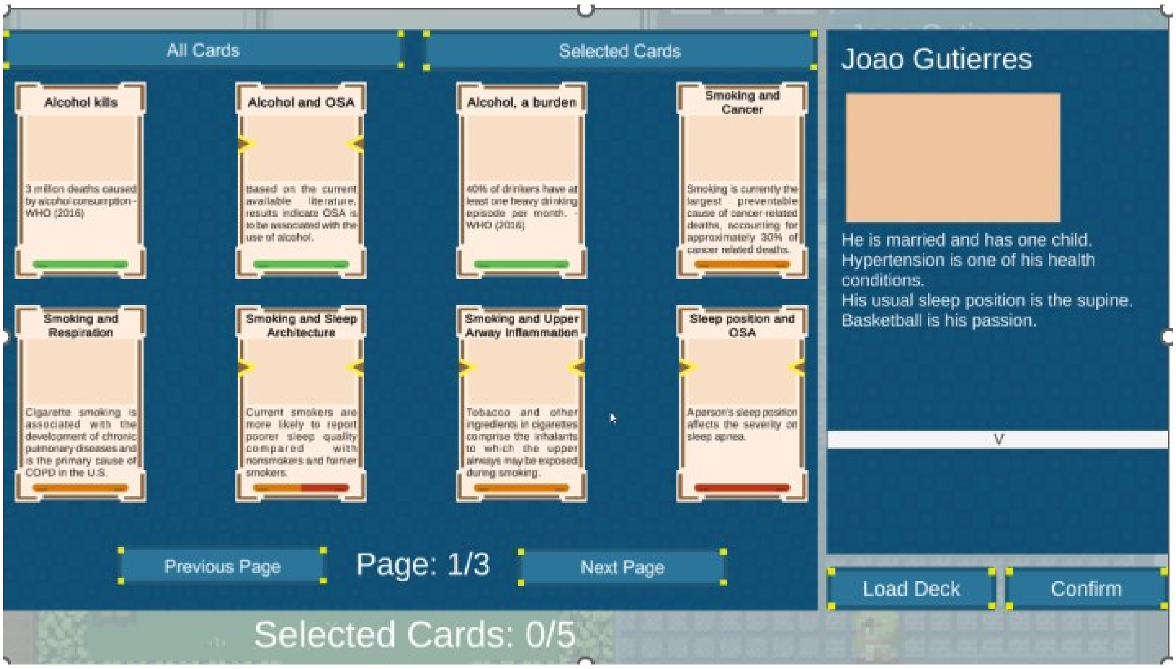


Figure 6-3 : Screenshot from the meta game aspect of Wake Up for the Future.

6.2.3 Interaction interpreter

The GDL [228], a language developed at Stanford University, was used to rewrite the SG in an interactive way to allow for interaction with the DRL algorithm. GDL is a rule-based language that is designed to represent the game mechanics of a wide range of games. The language is based on first-order logic, which allows for the representation of complex game rules and state transitions. GDL has been developed as a high-knowledge presentation formalism for axiomatizing the game rules. This approach applies a fundamental limitation to its representation capabilities, requiring the game to be governed only by deterministic rules. Due to this limitation only the core game of WuF was described by GDL in a more interactive way by representing the game mechanics in a formal language that can simplify the communication with the Interaction Interpreter.

An extension of the GDL has been developed to formalise rules of arbitrary games, however another approach was chosen to represent the meta game of WuF. A system was developed in Python to support different versions of the meta game, some of them incorporating the GA and some other enabling the generation of NPC profiles through randomness. This approach was deemed preferable, to be able to control the automated generation of content within the SG and to easily produce different versions of the SG for the agents to be trained on. By fusing the representations of the core and the meta game, GDL was integrated with the PCG technique to generate NPCs tailored to the player's performance while simultaneously adjusting their difficulty. This allows the represented game to provide a challenging and dynamic gameplay experience while still being able to interact with the DRL algorithm in a more efficient manner. This approach also makes the represented game more modular and easier to update in this setting, as changes to the game mechanics can be made in the GDL code, while changes in the PCG can be made to the Python based system, rather than the underlying game code. In this manner, both systems were merged in order to create an environment that would allow for deep reinforcement learning agents to automatically play the SG and be trained through each iteration.

6.2.4 Deep reinforcement learning

A family of policy gradient methods [229] for reinforcement learning have been selected to implement the agents. The DRL algorithm used is PPO, developed by OpenAI. PPO is a type of reinforcement learning algorithm that has gained widespread popularity in recent years due to its simplicity and stability. PPO algorithms are designed to learn optimal policies for decision-making in complex environments, such as game playing and robot control. One of the key features of PPO algorithms is their ability to update the policy in a more stable manner compared to other reinforcement learning algorithms. This is achieved through the use of a trust region constraint, which limits the size of policy updates and helps to prevent large, destabilizing changes. PPO algorithms also use an optimization objective that is simpler and easier to optimize than other reinforcement learning algorithms, further improving their stability and efficiency. These methods alternate between sampling data through interaction with the game environment and optimize a “surrogate” objective function using stochastic gradient ascent. In addition, these proximal policy optimization methods enable multiple epochs of minibatch updates, in contrast with standard policy gradient methods which perform one gradient update per data sample. OpenAI offers these methods in the stable baselines package, along with a variety of other DRL algorithms, developed in Python. The way to interact with PPO is through Gymnasium [230], formerly known as Gym by OpenAI. Although by using Gymnasium a variety of DRL algorithms becomes available for testing, PPO has been selected because of its simplicity to fine-tune and adequate performance in complex environments.

The performance of the PPO agents has then been studied to determine the effectiveness of the different way of generating NPCs in the employed SG. A recent study has shown the capability of PPO to present comparable attention to game elements to human players. This approach has allowed for the initial testing of the game in a controlled and efficient manner, without the need for in-person testing or human players. This approach is simpler to implement, more general, and has better sample complexity. During learning, an agent starts with no prior knowledge of the game, apart from the initial game state. Each epoch of training involves several interactions of the agent with the SG. The interactions include the agent providing the SG with a vector describing its current action in the state of the SG and receiving a response that includes the action's score and the next SG state. After each epoch the agent's weights are updated through backpropagation, with the ultimate goal of maximising the sum of individual scores collected by the agent's actions and thus its optimization to achieve victories against the NPCs. The results of the study provide insight into the potential of this novel approach for automated SG testing.

6.2.5 Analysis

For the evaluation the automated play agents' ability to learn in the game space, the two described versions of the WuF variant were employed. The SG version incorporating the PCG technique to produce content according to the agent's actions is denominated as "ga" in the results, while the SG version that produces game content using randomization rule-based mechanics is denominated as "simple" in the results. Through autonomous playing, the ability of the agent to optimize its actions in the simulated card battles was evaluated, and insight regarding the PCG's potential to increase the agent's training speed and capacity was attained.

To facilitate a systematic approach in comparison between efficiency in training on the employed SG versions, twenty different training scenarios were implemented, allowing the agents to train on the same initial conditions. This allowed for both an overall and a pairwise comparative analysis. In this manner, a total of 40 agents were trained, 20 on each version of the SG. Analysis was conducted by measuring the average win rate of each agent and the cumulative reward gained at regular intervals. The agent was rewarded for contradicting an enemy card and achieving victory in a card battle. Each interval included 500 time steps, with each time step indicating an action in the SG. The number of card battles included within each interval varied between 100 and 125, based the number of actions the agent would need to complete them. As the agents learned over time, they could achieve victory with fewer actions and the number of card battles included within an interval would increase. Each agent was trained for approximately one million steps. Finally, the percentage of the chosen attributes generated by the proposed PCG technique, and the randomization rule-base system were logged for every training interval.

6.3 Results and discussion

Results in terms of overall wining rate for the agents and cumulative awards gained over time are presented below. Training efficiency with the SG was grouped in three major categories according to three training patterns that were observed. The first category features cases of optimal training for the agent trained on the version of the SG featuring the GA approach, with agents displaying superior training capacity when exposed to the procedurally generated content. The second category features cases where the agents trained on the adaptive version of the SG displayed better performance in the metrics, however this superiority was observed after a certain number of training epochs and the overall improvement in metrics was not as significant as in category 1. Finally, category 3 features training cases where both agents, trained in the adaptive and simple version, performed with similar efficiency in the measured metrics. There were no cases observed in the twenty training scenarios that were implemented where the agent trained in adaptive content underperformed in comparison with the agent trained on the same environment in the simple version of the game. Out of the 20 environments, training on 7 resulted in patterns fitting to category 1, 8 in category 2, and 5 in category 3. An example for each of the three categories in terms of the observed metrics is presented in Fig. 6-4 through Fig. 6-9.

Example training of Category 1:

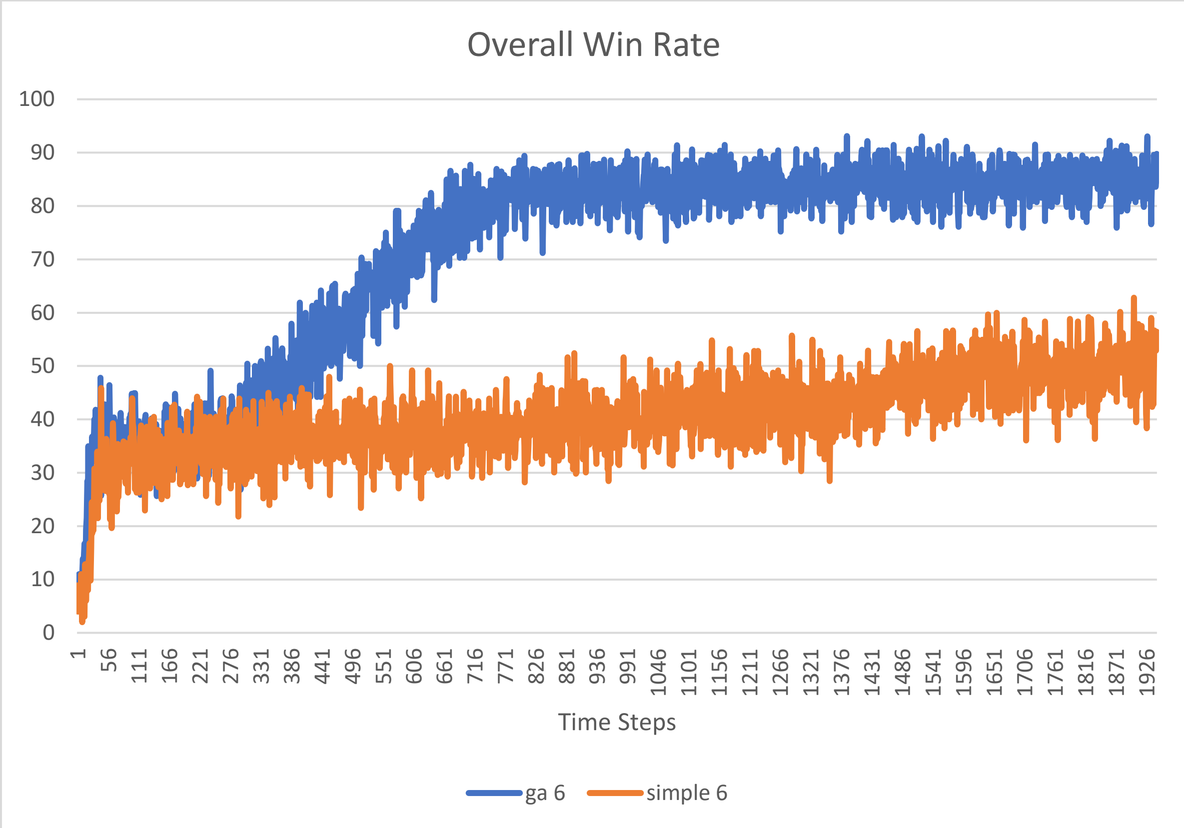


Figure 6-4: Overall win rate of an example training in category 1.

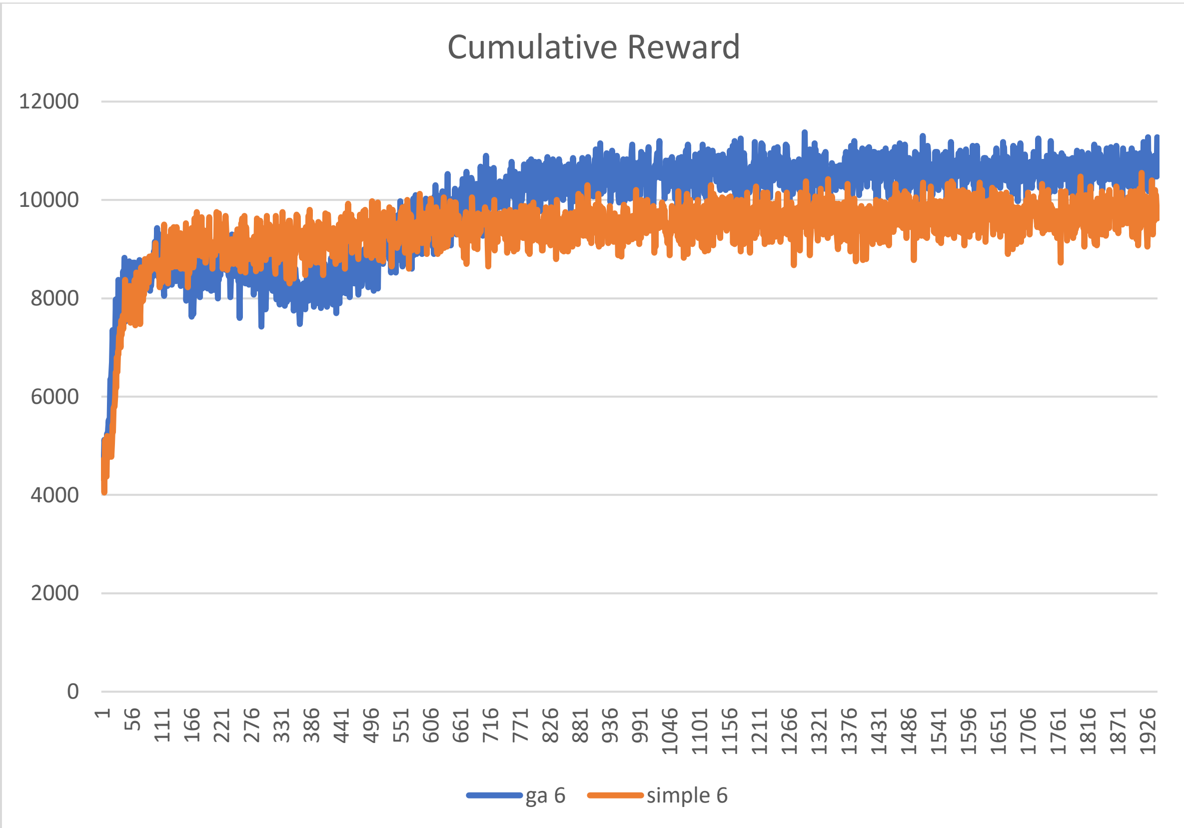


Figure 6-5: Cumulative reward of an example training in category 1.

Example training of category 2:

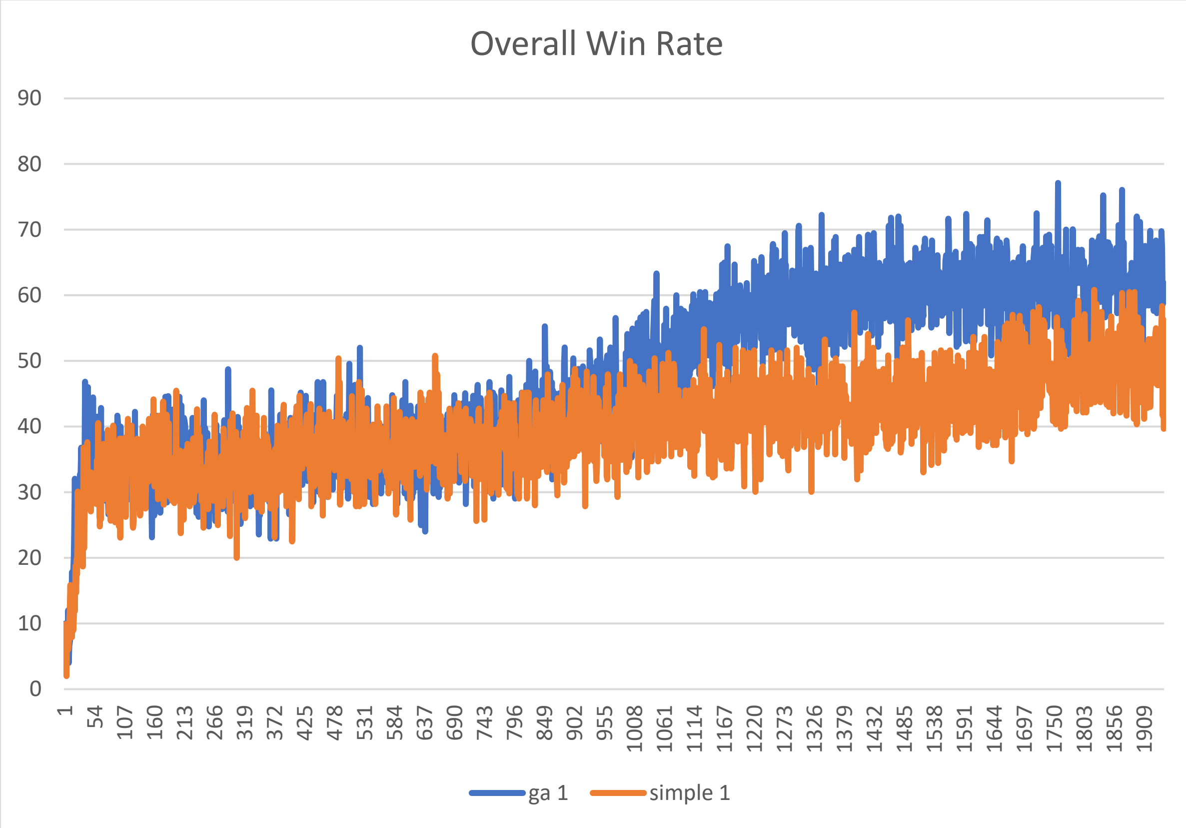


Figure 6-6: Overall win rate of an example training in category 2.

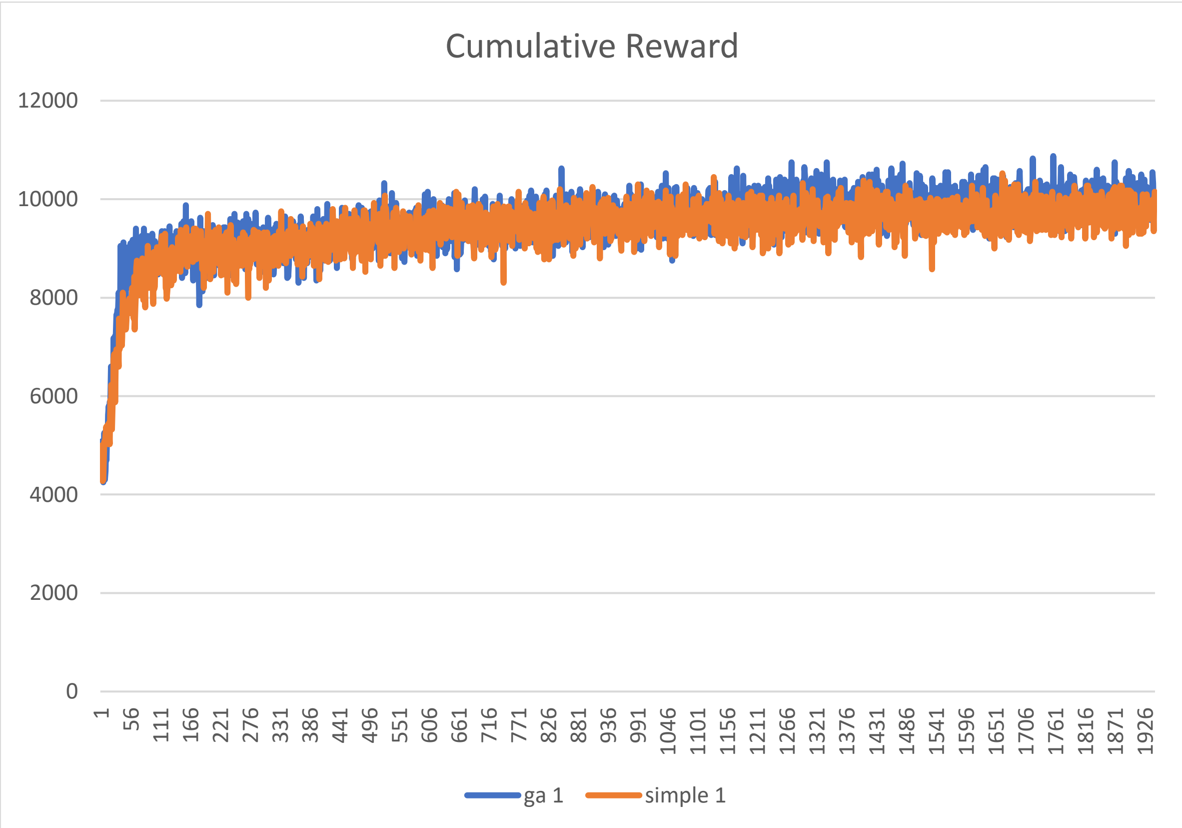


Figure 6-7: Cumulative reward of an example training in category 2.

Example training of category 3:

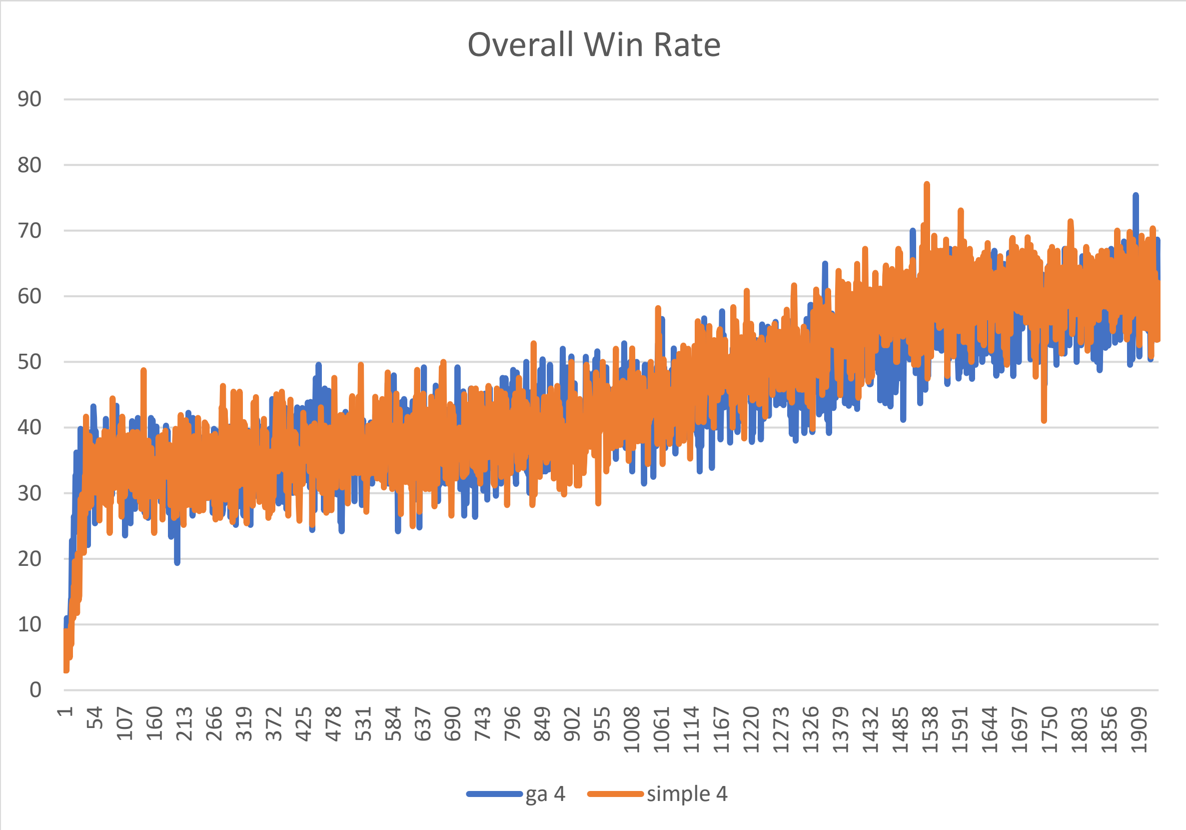


Figure 6-8: Overall win rate of an example training in category 3.

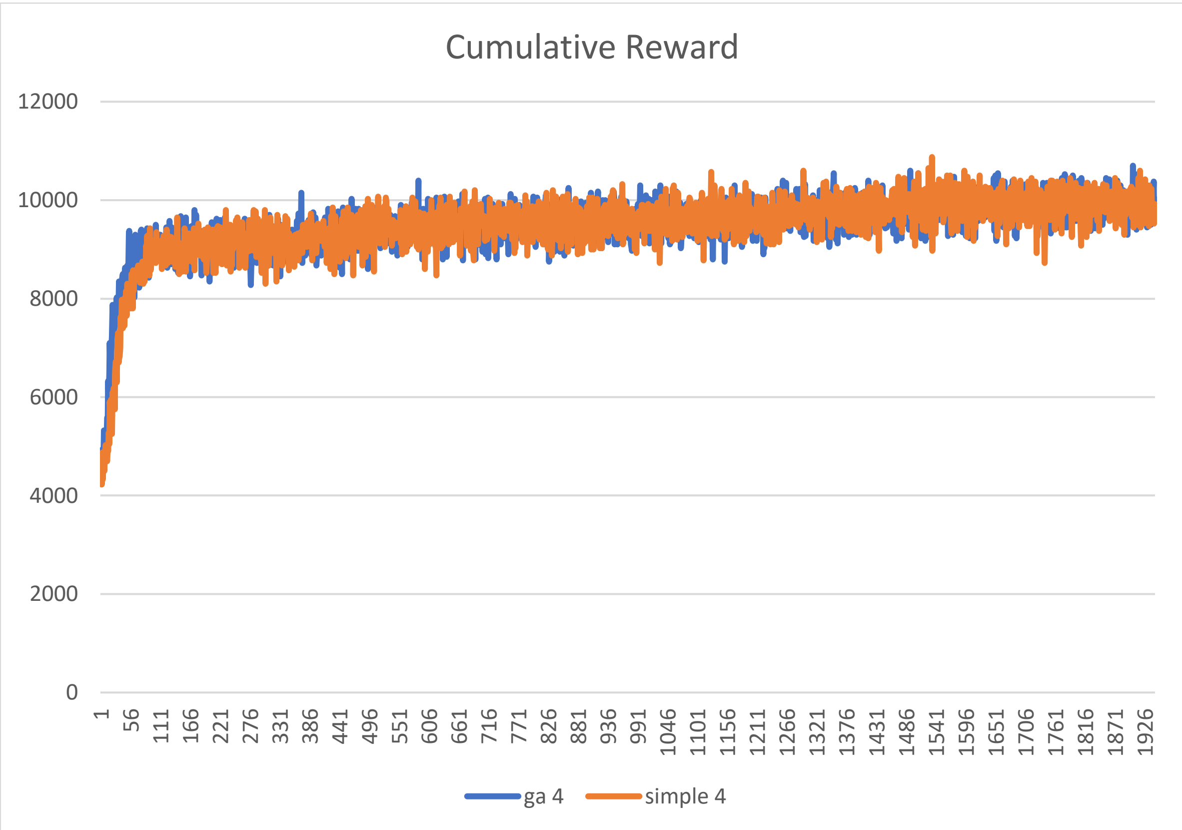


Figure 6-9: Cumulative reward of an example training in category 3.

Figure 6-10 displays the average performance of all agents in the twenty training scenarios, as depicted by average win ration in the game for every 10.000 training steps. A statistical significance in Student's t-test ($P<0.05$) was observed between agents been trained on the adaptive version of the SG (A) and agents trained on the simple version of the SG (B), after 30.000 training steps, with agents being trained on version A performing better. Figure 6-11 presents overall win rate for agents trained on versions A (blue) and B (orange) in every training scenario that was implemented.

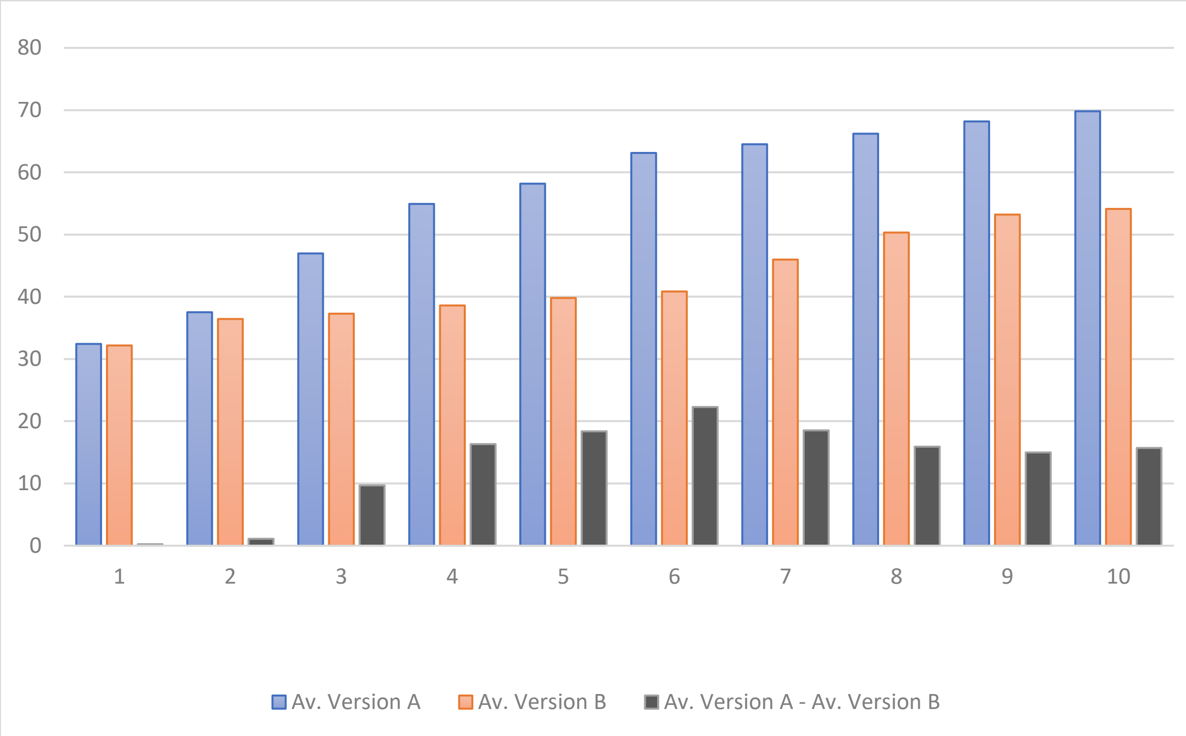


Figure 6-10: Average performance for all training scenarios for the adaptive version of the serious game (A) and the simple version of the serious game (B) every 10.000 training steps.

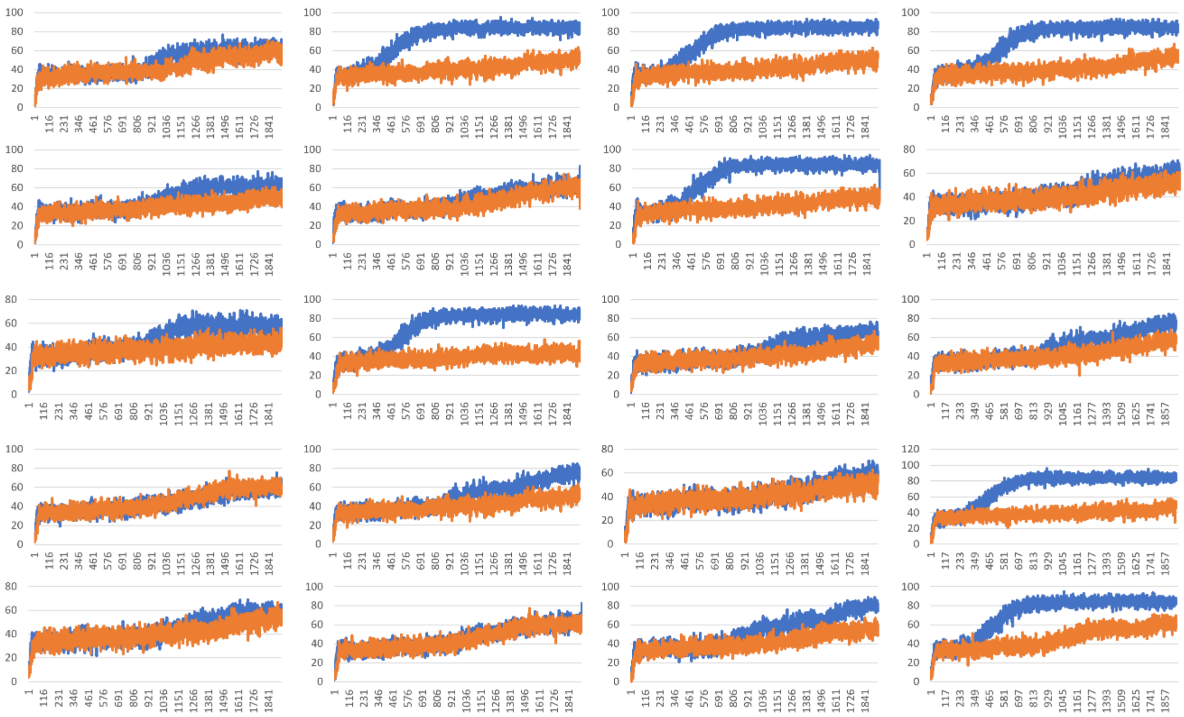


Figure 6-11: Overall win rate for all training scenarios.

One of the main challenges in using DLR agents to test SGs in an automated manner is ensuring that they can adapt to different game content and learn from their experiences as they are exposed to it. This is particularly important in the context of SGs that incorporate PCG techniques, as the effectiveness of these interventions depends on their ability to tailor their content according to specific user needs. In order to apply DLR agents to evaluate the impact of PCG techniques in the content produced by WuF two systems were developed to provide a training environment that would ensure that the agents are exposed in the deterministic nature of the core game while the meta game generates the appropriate content as intended. This approach aims to enable the agents to learn and adapt to different game content and to improve their performance over time. Overall, the goal of this methodology is to assess the feasibility and effectiveness of using DRL agents to play SGs automatically, and to identify the key factors that impact their performance. By understanding these factors, the development of more effective and personalized SGs for health purposes, and the advancement the field of digital health interventions can be facilitated.

Based on the metrics that were monitored during the training of the DRL agents, performance of agents trained on the adaptive version of the SG was found to be superior to the performance of agents trained on the simple version of the SG. Average win rate was found to be significantly higher on average for all training scenarios after a number of training epochs. Three categories were defined according to in-game performance, with agents in category 1 performing significantly better since early in the training process, agents in category 2 performing slightly better later in the training process and agents in category 3 trained in both versions of the SG having similar performance. No cases where agents trained in the adaptive version underperformed in comparison to agents trained in the simple version of the SG were observed. This is a strong indication towards the capacity of the proposed PCG technique that is based on a GA to generate content according to specific player needs in response to their performance in the SG. These results can be a steppingstone towards the development of tools that are able to evaluate the efficiency of adaptive SGs and allow for automated calibration of the PCG parameters. Future steps include the comparison between versions of the SG that include PCG techniques with different parameters in an effort to augment the generated content in an automated and systematic manner. Results from this type of comparison can also be validated through an experimental process with human participants to replicate the results obtained through DRL agent training with human players. The impact of procedurally generating appropriate content could also be generalized in other types of interventions besides SGs and DRL agents could serve as a way to prepare better prepare these techniques for application.

7. Conclusions and future work

In the present Doctoral Thesis, a novel conceptual framework (Fig. 7-1) that allows for real time recognition of engagement during play with adaptive SGs for health is presented. The framework employs data from sensors, interaction with the SG, and user-specific health related data to tailor SG content, through PCG techniques, to personalised user needs. The aim of this approach is to empower the health intervention provided by the adaptive SG by delivering content relevant to user health needs, while enhancing engagement. The literature review conducted during the present Thesis advocates towards the potential of adaptive SGs to serve as state-of-the-art health interventions with the ability to address modern and urgent healthcare needs. Furthermore, recent advancements in sensing technology, as well as the availability of affordable and easy to use wearable devices, able to collect a multitude of health-related data in real time, present an unprecedented opportunity for such approaches. To investigate the feasibility and potential impact of the proposed conceptual framework, two novel SGs aiming to empower self-health management in chronic conditions have been designed and developed. These SGs have been employed, along with a third SG for health, in experimental processes, pilot studies, and simulated gameplay to collect data and investigate the impact of real time recognition of engagement and incorporation of PCG techniques in SGs games for health. In addition, the proposed PCG technique has been evaluated in terms of its ability to generate content which expedites the training of DRL agents for automated testing.

The insights gained from the presented results, along with conclusions drawn from relevant literature, provide convincing arguments towards the value of incorporating a real-time closed engagement feedback loop in adaptive SGs for health, based on sensor and interaction data, as suggested by the proposed conceptual framework. Some of the key conclusions presented in this Doctoral Thesis include the potential superiority of multimodal approaches, which take advantage of heterogeneous data sources, for real-time recognition of engagement in comparison to more traditional unimodal approaches. Additionally, sensor data containing information about specific health-related user needs, can help tailor SG content through PCG techniques. Finally, DRL agents developed for automated testing in SGs for health achieve more efficient learning when exposed to content generated by the proposed PCG technique.

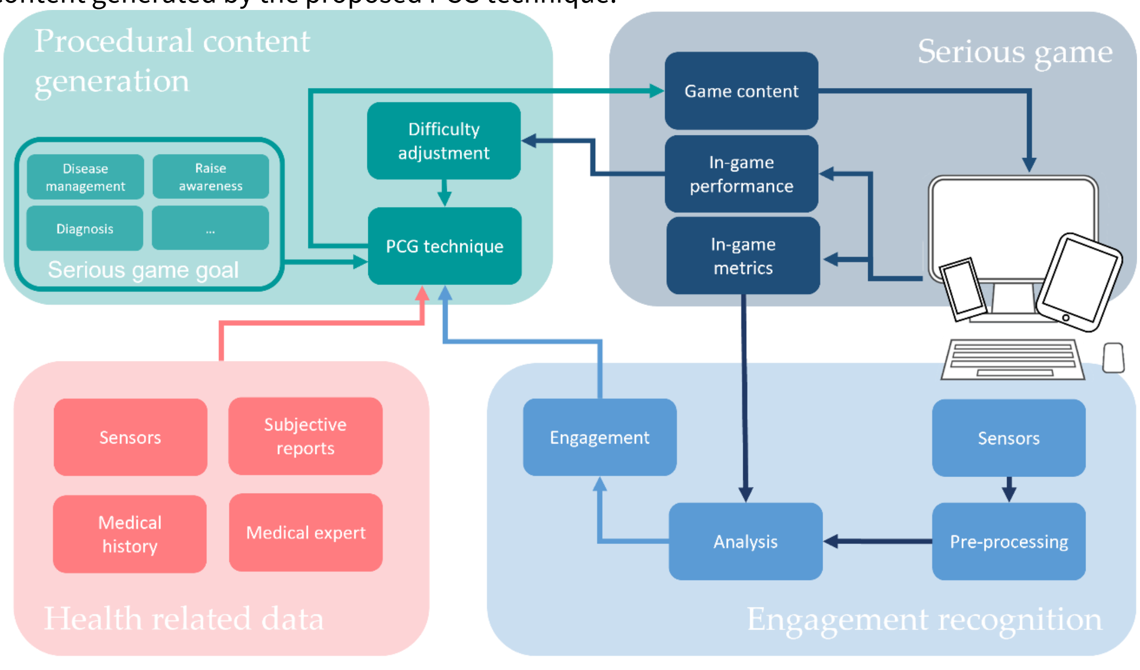


Figure 7-1 : The proposed conceptual framework for adaptivity in serious games for health.

In the following sections conclusions from the work detailed in the current Thesis are presented, along with identified limitations and possible future work. The first section deals with conclusions drawn regarding the adaptive properties incorporated in SGs for health, and in particular, their capacity to affect player experience, enhance engagement, and deliver personalized content. The second section deals with conclusions drawn regarding player modelling, and in particular, the special needs imposed by the nature of the intervention provided by SGs for health while taking into account the effort to achieve real-time recognition of engagement during interaction with them. The final section of this chapter presents future directions that will facilitate and further substantiate the employment of the proposed conceptual framework in building adaptive health interventions that employ SGs.

7.1 Adaptive serious games for health

As discussed in Chapter 1, adaptivity in SGs for health can provide the means for dynamic and personalized interventions that not only deliver tailored content, but also enhance adherence to the intervention. The proposed conceptual framework builds on this notion, by employing data that can help tailor game content towards both of these directions. PCG techniques employed for adaptivity in games hold additional benefits, such as reducing production cost, facilitating user generated content and enabling player driven design approaches, as was evidenced by the design and development processes that were undertaken for the SGs included in the present Thesis. The incorporation of PCG techniques and mechanics facilitating the procedural generation of game content proved to reduce the necessary resources in the SG production process, while providing a systematic way to generate additional game content that can be controlled by a set of rules. In turn, this allowed for easier implementation of changes in the SG suggested by participants during the experimental processes. Besides the observed advantages, the incorporation of PCG techniques in SGs for health proved to present certain challenges as well. Interaction with SGs might share a lot in common with entertainment digital games, however, the difference in their intended use and game content calls for specific approaches in the incorporated mechanics for adaptivity. This is particularly true for PCG techniques employed in SGs for health, where the necessity to ensure that the delivered content is appropriate for the desired intervention is of paramount importance. The incorporation of a control system, able to monitor the delivery of health-related game content to the user is thus necessary, as shown in the incorporation of the GA approach in the SG for type 1 diabetes and obesity. The layered approach of the proposed conceptual framework for adaptivity was found to facilitate the design of such a contingency system. Both SGs that were created during the present thesis have been designed with these considerations in mind, to facilitate adaptivity.

(a) (b)

Figure 7-2: Game mechanics facilitating the incorporation of procedural content generation techniques in the employed serious games: (a) ingredient selection in Express Cooking Train; (b) card selection in Wake Up for the Future.

Regarding the mechanics facilitating adaptivity in ECT, their design approach takes into consideration the SG's goal, namely, to build skills and introduce behaviors with the ultimate goal of promoting effective dietary change in young adults. ECT's design framework is based both on a theoretical background and modern game mechanics to improve player engagement and achieve this desired outcome. To this end, the in-game systems that govern game content benefit from ontological modeling. In this manner, game content such as food ingredients (Fig. 7-2a) can be dynamically altered based on different queries that can be in turn controlled by PCG techniques. Thus, personalized dietary needs tied to the self-management of chronic conditions can be represented through appropriate dynamic game content. Additionally, these queries can be easily controlled and monitored to ensure the appropriate nature of the game content delivered to the user. In terms of the effect these mechanics displayed in the efficiency of the intervention provided by ECT, results from the experimental process indicate that it is on par with the traditional educational intervention that was employed. No statistically significant differences were observed in the post-intervention application of the knowledge questionnaire, however, the fact that no participants received a lower score in the case of ECT could point towards its superiority.

The small number of the participants, as well as differences in their knowledge basis, as depicted in their initial responses to the questionnaire, can be considered a limitation of this experimental process. This can be especially true when taking into consideration the complex nature of ECT and the multiple mechanics that work in conjunction to deliver the desired outcome. Participants only played one session of the SG, limiting thus their interaction with the intervention. ECT's system for ontology-assisted generation of game content was also limited during the experimental process, as all participants were exposed to similar content to avoid risk of bias. However, the free exploration system of the kitchen simulator that can be employed towards PCG was widely accepted among participants and produced an intuitive game experience. Investigation of knowledge retention was also inconclusive, which is to be expected given the limited pool of participants and the relatively small amount of time (one week) that was given for the participants to retake the knowledge questionnaire. Finally, investigation towards the efficiency of ECT was conducted using a version of the SG prior to changes conducted in the UI and tutorial of the game. The increase that was observed in positive fields of GEQ might have contributed towards increased learning and knowledge retention. This preliminary examination of the potential of a complex SG such as ECT, points towards the feasibility of borrowing mechanics and systems from state-of-the-art digital entertainment games with the capacity to adapt their content dynamically as indicated in the proposed conceptual framework.

In terms of game experience and acceptance ECT displayed positive responses in the fields of GEQ. Comparison between the two versions of ECT that were different in terms of user interface demonstrated that the enhanced version produces a better overall experience amongst participants. User feedback was employed to introduce changes in ECT, specifically in terms of player controls, user interface and the tutorial. The kitchen simulation received few negative comments and as a result limited changes were made to its interface to accommodate for them. These changes were found to have an immediate effect on user experience, thus, highlighting the importance of participatory user-driven game design for SGs. Player driven design techniques can be further enhanced with the use of PCG approaches, as procedural generation of game content allows for the development of assets that can be created by users and experts. The improved version of ECT achieved better scores in the GEQ in comparison with the corresponding results of five SGs for the Core module and three SGs for the post-game module. Since it is not easy to compare games with so many differences, this comparison is not expected to provide an indication towards ECT's acceptance, or the improvement between the two versions of the SG, but rather to serve as a tool to investigate the effect of different design choices.

An approach that facilitates adaptivity was also adopted in the design and development of WuF. WuF aims to empower self-health management in OSA and raise awareness about this underdiagnosed condition, mainly targeting older adults. To achieve its goal WuF features simulated debates through flexible card game mechanics as well as an open pixel world that allows for player exploration and interaction with NPCs. The incorporation of the GA-based PCG technique in WuF proved to be facilitated through the SG's card mechanics, which provide an environment that can be easily controlled by available player cards, deck formation, and the card selection process (Fig. 7-2b). Results from the preliminary validation of this technique with human participants indicated differences in game experience between the game versions deployed, despite no visible differences between them. The employed versions featured different variations of the proposed PCG technique in terms of the generated content and dynamic difficulty adjustment. Sense of competence, negative experience and challenge were found to be different among participants according to the generated content and difficulty of the SG, with the observed trends matching the intended design choices. The limited number of participants and the small length of the play sessions can be considered limitations for this study as well. However, the fact that the impact of the procedurally generated content on player experience was observable in this limited playthrough is encouraging regarding the proposed technique's potential to affect the intervention's capabilities.

This initial implementation of the proposed technique for adaptivity based the generated game content solely on data produced through the interaction with the SG. The proposed technique for adaptivity in SGs for health was also investigated in terms of its potential to be applied in different SGs and employ a variety of data. To evaluate the proposed conceptual framework for adaptivity, the GA-based technique was designed with the capacity to process heterogeneous data from a multitude of sources. Additionally, its capacity to represent game content in an abstract manner proved to enhance the technique's generalization capabilities to be incorporated in different types of SGs, as it was intended through its design. To this end, the proposed technique was employed in the Endorse platform and the SG it incorporates, generating game content in the form of available missions and messages displayed to the player. In this instance the PCG technique was not only based on data produced through player interaction, but also takes advantage of health-related data collected through sensors to generate game content. The SG incorporated in the platform aimed to promote self-health management in type 1 diabetes and childhood obesity. Its modular nature, featuring a mission selection system, was found to facilitate the incorporation of the proposed technique as well.

In this instance game content was generated in the form of missions available to the player and messages displayed to them throughout the game. The answers provided in the post-intervention questionnaires indicate an overall acceptance regarding the individualized content provided by the GA, in terms of relevance and usefulness. The participation of health experts in the design of the adaptive content was found to be beneficial, advocating towards the need for participatory SG design. The proposed PCG technique was able to employ sensor data to recognize and promote messages that were relevant to user needs. This was evident in the observed correlations between the generated content and the collected health-related data. Furthermore, high levels of correlations were avoided indicating that content did not reach saturation and was successfully diversified by the algorithm to avoid repetition and maintain engagement in the intervention. These results display the superiority of the proposed approach over simpler rule-based systems in generating tailored content. One of the limitations of this pre-pilot study can be considered the limited number of missions in the employed SG. In next iterations that were evaluated through pilot studies for children suffering from obesity and type 1 diabetes game content was enhanced with more educational and action missions. In conclusion, the importance of making design options that facilitate the enhancement of SGs through adaptivity was largely highlighted in all case studies presented in this Thesis.

Finally, the proposed PCG technique was evaluated in terms of its capacity to facilitate automated testing with the use of DRL agents. According to the monitored metrics during the training, accelerated learning was achieved in the case of the agents exposed to content generated by the PCG technique. The employed version of WuF that incorporated the GA-based technique included similar parameters with the version that achieved better user acceptance on the preliminary experimental process with human players. The capacity of DRL agents to perform differently according to the content provided by the SG is a strong indication towards the capacity of applying automated testing in adaptive SGs for health. Insights gained through this study can be employed to limit the requirements in human testing and facilitate the calibration of PCG techniques in SGs. The investigation of the performance of agents in multiple training scenarios in SG spaces that are controlled by different versions of PCG can shed light in the efficiency of PCG. The application of additional types of agents can also help substantiate this type of testing. The generalization capabilities of automated testing in different games controlled by similar PCG techniques can also be investigated.

7.2 Player modeling in serious games for health

In the present Doctoral Thesis player modeling was conducted towards real-time recognition of engagement. An affective state of engagement was investigated, based on subjective player annotated traces. The presented experimental process aimed at extracting features from sensor data and interaction with the SG, that can lead to real-time recognition of engagement during interaction with ECT. This process led to substantial results that advocate towards the feasibility of the proposed conceptual framework for adaptivity. Results based on the recognition of sitting postures indicate that the assumed sitting postures, the transitions between them, and the overall seated mobility are associated with user-perceived engagement. Associations pointing in that direction were identified across both types of observation frames. Shifts in sitting position were correlated with increased perceived engagement both in continuous and reactive frames. Laying on the back of the chair was correlated with absence of perceived engagement, while the process of shifting to an upwards seated position that did not include the back of the chair points towards an increase in engagement. The significance of these observations appears to be affected by the duration of the observation frames, with increased time yielding more significant results in all observation frames. These results can facilitate the development of systems for real time recognition of engagement during SG play in desktop computers. The increased need for remote working and learning during the COVID-19 pandemic increased the value of such approaches. For their generalization, identifying a general sensor activation threshold for posture monitoring with smart chair can be very important. Results towards this direction are limited in the presented experimental process, as participants were young and featured normal BMIs. Additionally, postures observed and identified during the intervention were not particularly relaxed due to their presence in a research setting. The investigation of additional office setups with multiple chairs incorporating similar sensors can lead to valuable results towards this direction.

Analysis for real time recognition of engagement indicated a significant predictive capacity of both sensor-based sources and in-game metrics towards player perceived engagement. The features of the annotation traces were investigated in terms of their potential to act as the ground truth for engagement during interventions with a SG for health. Differences among them in the presented results were discussed. The duration and nature of the observation frames proved to be significant in recognizing perceived engagement in real time. Game design appears to play a significant part in including appropriate game events that can trigger emotional responses from the player and facilitate real time recognition of engagement. In the presented results from the experimental process, reactive frames seem to produce the majority of features with significant predictive value, with 30 second frames revealing the most significant correlations. To this end, recognition of engagement needs to be tailored in each application to the intervention provided by a SG. Appropriate observation frames and in-game events need to be identified. An automated approach towards the selection of this characteristics could prove valuable in the generalization of such approaches. Features extracted from sensor data and in-game metrics presented a range of significant and highly significant correlations with the player annotated perceived engagement, across all types of observation frames. Comparison between features in a unimodal fashion was not possible. The employed multimodal feature, however, was found to be significantly correlated with features of the annotation trace consistently. Data collected from affordable sensors was found to contain information that can model engagement during interaction with a SG for health. Limitations of this study includes the number of participants and the platform the intervention was tested on. Interaction was only conducted through a keyboard and a mouse; however mobile games are becoming increasingly prevalent. Investigation for real time recognition of engagement during interaction with a SG for health on a mobile platform can yield significant results towards the feasibility of the proposed conceptual framework for adaptivity.

7.3 Future work

The proposed conceptual framework for leveraging sensors towards recognising engagement in real time and procedurally generating content in SGs for health needs to be investigated further. Results collected from the pilot and pre-pilot studies with children suffering from type 1 diabetes and obesity will be applied towards additional evaluation of the proposed PCG technique. The generated SG content in these studies, as in the case of the pre-pilot study that was included in the present Doctoral Thesis, is based on sensor data that monitor disease self-management and lifestyle habits and provide game content tailored to specific clinical needs. Besides children with obesity, these studies include children suffering from type 1 diabetes mellitus that use wearable sensors for blood glucose and insulin intake monitoring, as well as activity trackers. Data collected from these sensors, along with data from the interaction with the SG and responses to user acceptance and feasibility questionnaires will be employed to assess the GA's capacity to deliver personalized content accurately. The efficiency of the intervention will also be evaluated, along with user acceptance and experience regarding the PCG technique incorporated in the SG. The studies include versions of the employed SG with additional content in terms of mini games and messages. More participants were recruited, while updates and enhancements in the employed platform led to increased adherence levels. Analysis on the collected data is expected to provide valuable insight regarding the proposed PCG technique's potential and generalization capabilities, as well as towards the feasibility of the proposed conceptual framework.

Regarding the investigation of real-time recognition of engagement during interaction with SGs for health, the presented analysis has investigated the features' predictive capability in a linear fashion. Supervised learning techniques can be employed to assess the features' capabilities in classifying perceived engagement in a non-linear fashion. Predictive models of engagement can be employed to serve as controllers in constant affective feedback loops and empower adherence to SG-based health interventions. Besides their application in real time, these approaches can also guide SG design by identifying player experience pitfalls in SGs. Advanced techniques in multimodal fusion can be employed to investigate the potential of multimodal features to a greater length via ML. In addition, a multitude of features from a variety of sensor-based data can be employed towards this direction. Deep learning and reinforcement learning methodologies can also assist in this investigation, given their increasing application in the field of multimodal affective recognition. However, larger datasets are necessary in this direction. Issues related with gender representation in participants should also be investigated towards identifying potential impact of gender imbalance on our core findings. The proposed conceptual framework needs to be investigated in terms of recognizing engagement in real time during interaction with other SGs for health as well. Differences in game mechanics and systems are expected to have a great impact on identifying perceived engagement. Regarding the employment of DRL agents, as mentioned in Chapter 6, future directions include the comparison between different versions of the PCG techniques to evaluate their potential in accelerating agent learning. Results from this type of comparison can also be validated through an experimental process with human participants to replicate the results obtained through DRL agent training with human players.

Finally, a novel experimental procedure will be implemented to evaluate the proposed conceptual framework in its entirety. The procedure will employ a novel SG that aims to diagnose Parkinson’s disease and monitor the progression of motor symptoms by collecting tracing samples from smartphone touch screens. The SG is in the process of development and features the proposed GA approach for PCG. Shapes for tracing are generated procedurally in the form of labyrinths and game difficulty is dynamically adjusted. The SG is designed on the premises of the proposed conceptual framework to procedurally generate appropriate game content that maximizes player engagement and addresses their personalized needs according to the identified severity of motor symptoms. The experimental process will deploy sensors for EEG, heart rate and recognition of hand movement, as well as data from interaction with the SG. The collected data will be analysed along with annotated data of perceived engagement in an effort to produce a clinically relevant engagement feedback loop that drives generation of game content in accordance with the proposed framework.

References

[1] L. Sardi, A. Idri, and J. L. Fernández-Alemán, “A systematic review of gamification in e-Health,” *J. Biomed. Inform.*, vol. 71, pp. 31–48, Jul. 2017.

[2] T. M. Fleming *et al.*, “Serious games and gamification for mental health: Current status and promising directions,” *Front. Psychiatry*, vol. 7, p. 215, Jan. 2017.

[3] K. Mitsis *et al.*, “A Multimodal Approach for Real Time Recognition of Engagement towards Adaptive Serious Games for Health,” *Sensors*, vol. 22, no. 7, p. 2472, Mar. 2022.

[4] R. Orji, J. Vassileva, and R. L. Mandryk, “Modeling the efficacy of persuasive strategies for different gamer types in serious games for health,” *User Model. User-adapt. Interact.*, vol. 24, no. 5, pp. 453–498, Oct. 2014.

[5] C. Schrader, J. Brich, J. Frommel, V. Riemer, and K. Rogers, “Rising to the Challenge: An Emotion-Driven Approach Toward Adaptive Serious Games,” *Serious Games Edutainment Appl. Vol. II*, pp. 3–28, Jan. 2017.

[6] P. Sajjadi, A. Eiwas, and O. De Troyer, “Individualization in Serious Games: A Systematic Review of the Literature on the Aspects of the Players to Adapt To,” *Entertain. Comput.*, vol. 41, no. 2, p. 100468, 2021.

[7] S. Çiftci, “Trends of Serious Games Research from 2007 to 2017: A Bibliometric Analysis,” *J. Educ. Train. Stud.*, vol. 6, no. 2, p. 18, Jan. 2018.

[8] P. Wilkinson, “A Brief History of Serious Games,” in *Entertainment Computing and Serious Games*, 2016, pp. 17–41.

[9] F. Laamarti, M. Eid, and A. El Saddik, “An overview of serious games,” *Int. J. Comput. Games Technol.*, no. 11, p. 11, Jan. 2014.

[10] R. T. Hays, “The effectiveness of instructional games: a literature review and discussion,” *Nav. Air Warf. Cent. Train. Syst. Div.*, pp. 1–63, 2005.

[11] P. Moreno-Ger, I. Martinez-Ortiz, M. Freire, B. Manero, and B. Fernandez-Manjon, “Serious games: A journey from research to application,” in *Frontiers in Education Conference*, 2014, pp. 391–394.

[12] S. Gentry *et al.*, “Serious Gaming and Gamification interventions for health professional education,” *Cochrane Database Syst. Rev.*, vol. 23, no. 6, p. e12994, Mar. 2019.

[13] AlliedMarket, “Serious Games Market Size, Share Industry Forecast - 2021–2030,” *Serious Games Market Statistics: 2030*, 2020. [Online]. Available: https://www.alliedmarketresearch.com/serious-games-market. [Accessed: 07-Apr-2022].

[14] “Serious Games Market | 2022 - 27 | Industry Share, Size, Growth - Mordor Intelligence.” [Online]. Available: https://www.mordorintelligence.com/industry-reports/serious-games-market. [Accessed: 07-Apr-2022].

[15] F. Mäyrä and K. Alha, “Mobile Gaming,” *Video Game Debate 2*, pp. 107–120, Nov. 2020.

[16] M. Ma, L. C. Jain, and P. Anderson, “Future trends of virtual, augmented reality, and games for health,” *Intell. Syst. Ref. Libr.*, vol. 68, pp. 1–6, 2014.

[17] C. C. Abt, *Serious games*. University Press of America, 1987.

[18] D. Djaouti, J. Alvarez, J.-P. Jessel, and O. Rampnoux, “Origins of Serious Games,” *Serious Games Edutainment Appl.*, pp. 25–43, 2011.

[19] D. Djaouti, J. Alvarez, and J.-P. Jessel, “Classifying Serious Games,” in *Handbook of Research on Improving Learning and Motivation through Educational Games:*

Multidisciplinary Approaches, 2011, pp. 118–136.
[20] R. Dörner, S. Göbel, W. Effelsberg, and J. Wiemeyer, *Serious Games: Foundations, Concepts and Practice*. 2016.
[21] M. Hudson, “Non-Serious Serious Games,” *Press Start*, vol. 3, no. 2, pp. 1–21, 2016.
[22] D. R. Michael and S. Chen, *Serious games: games that educate, train, and inform*. Boston, Mass. : Thomson Course Technology, 2006.
[23] S. Tobias, J. D. Fletcher, and A. P. Wind, “Game-based learning,” *Handb. Res. Educ. Commun. Technol. Fourth Ed.*, pp. 485–503, Jan. 2014.
[24] O. V. Anikina and E. V. Yakimenko, “Edutainment as a Modern Technology of Education,” *Procedia - Soc. Behav. Sci.*, vol. 166, pp. 475–479, Jan. 2015.
[25] L. Rice, “Playful Learning,” *J. Educ. Built Environ.*, vol. 4, no. 2, pp. 94–108, Dec. 2009.
[26] M. Slussareff, E. Braad, P. Wilkinson, and B. Strååt, “Games for Learning,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9970 LNCS, pp. 189–211, 2016.
[27] L. von Ahn, “Games with a Purpose,” *Computer (Long. Beach. Calif).*, vol. 39, no. 6, pp. 92–94, Jun. 2006.
[28] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, “From game design elements to gamefulness: Defining ‘gamification,’” in *International Academic MindTrek Conference: Envisioning Future Media Environments*, 2011, pp. 9–15.
[29] B. Morschheuser, L. Hassan, K. Werder, and J. Hamari, “How to design gamification? A method for engineering gamified software,” *Inf. Softw. Technol.*, vol. 95, pp. 219–237, Mar. 2018.
[30] G. Baptista and T. Oliveira, “Gamification and serious games: A literature meta-analysis and integrative model,” *Comput. Human Behav.*, vol. 92, pp. 306–315, Mar. 2019.
[31] D. Liu, R. Santhanam, and J. Webster, “Toward meaningful engagement: A framework for design and research of gamified information systems,” *MIS Q. Manag. Inf. Syst.*, vol. 41, no. 4, pp. 1011–1034, Dec. 2017.
[32] S. McCallum, “Gamification and serious games for personalized health,” in *Studies in Health Technology and Informatics*, 2012, vol. 177, pp. 85–96.
[33] D. Johnson, E. Horton, R. Mulcahy, and M. Foth, “Gamification and serious games within the domain of domestic energy consumption: A systematic review,” *Renew. Sustain. Energy Rev.*, vol. 73, pp. 249–264, Jun. 2017.
[34] M. Fitzgerald and G. Ratcliffe, “Serious games, gamification, and serious mental illness: A scoping review,” *Psychiatr. Serv.*, vol. 71, no. 2, pp. 170–183, Feb. 2020.
[35] R. N. Landers, “Developing a Theory of Gamified Learning: Linking Serious Games and Gamification of Learning,” *Simul. Gaming*, vol. 45, no. 6, pp. 752–768, 2014.
[36] “game_1 noun - Definition, pictures, pronunciation and usage notes | Oxford Advanced Learner’s Dictionary at OxfordLearnersDictionaries.com.” [Online]. Available: https://www.oxfordlearnersdictionaries.com/definition/english/game_1?q=game. [Accessed: 09-May-2022].
[37] G. Hookham and K. Nesbitt, “A Systematic Review of the Definition and Measurement of Engagement in Serious Games,” in *Australasian Computer Science Week Multiconference*, 2019, pp. 1–10.
[38] P. Cairns, “Engagement in digital games,” *Why Engagem. Matters Cross-Disciplinary Perspect. User Engagem. Digit. Media*, pp. 81–104, Jan. 2016.
[39] S.-F. Henrik, “The Player Engagement Process - An Exploration of Continuation Desire in Digital Games,” in *DiGRA International Conference: Think Design Play*, 2011.

[40] E. A. Boyle, T. M. Connolly, T. Hainey, and J. M. Boyle, "Engagement in digital entertainment games: A systematic review," *Comput. Human Behav.*, vol. 28, no. 3, pp. 771–780, May 2012.

[41] A. Perttula, K. Kiili, A. Lindstedt, and P. Tuomi, "Flow experience in game based learning – a systematic literature review," *Int. J. Serious Games*, vol. 4, no. 1, Mar. 2017.

[42] E. Boyle, T. M. Connolly, and T. Hainey, "The role of psychology in understanding the impact of computer games," *Entertain. Comput.*, vol. 2, no. 2, pp. 69–74, Jan. 2011.

[43] V. Wattanasoontorn, I. Boada, R. García, and M. Sbert, "Serious games for health," *Entertain. Comput.*, vol. 4, no. 4, pp. 231–247, Dec. 2013.

[44] C. E. Catalano, A. M. Luccini, and M. Mortara, "Guidelines for an effective design of serious games," *Int. J. Serious Games*, vol. 1, no. 1, Feb. 2014.

[45] K. Mitgutsch and N. Alvarado, "Purposeful by design?: A serious game design assessment framework," in *Foundations of Digital Games 201*, 2012, pp. 121–128.

[46] D. Avila-Pesántez, L. A. Rivera, and M. S. Alban, "Approaches for serious game design: A systematic literature review," *Comput. Educ. J.*, vol. 8, no. 3, 2017.

[47] T. Baranowski, C. Ryan, A. Hoyos-Cespedes, and A. S. Lu, "Nutrition Education and Dietary Behavior Change Games: A Scoping Review," *Games Health J.*, vol. 8, no. 3, pp. 153–176, Jun. 2019.

[48] A. Bandura, *Social foundations of thought and action : a social cognitive theory*. Prentice-Hall, 1986.

[49] K. Starks, "Cognitive behavioral game design: A unified model for designing serious games," *Front. Psychol.*, vol. 5, no. FEB, pp. 1–10, 2014.

[50] C. Cheek *et al.*, "Integrating health behavior theory and design elements in serious games," *JMIR Ment. Heal.*, vol. 2, no. 2, Apr. 2015.

[51] E. Brox, S. T. Konstantinidis, and G. Evertsen, "User-centered design of serious games for older adults following 3 years of experience with exergames for seniors: A study design," *JMIR Serious Games*, vol. 5, no. 1, pp. 1–14, 2017.

[52] Y. Y. N. Ng and C. W. Khong, "A review of affective user-centered design for video games," *Int. Conf. User Sci. Eng. Exp. Eng. Engag.*, pp. 79–84, 2015.

[53] F. Laamarti, M. Eid, and A. El Saddik, "An overview of serious games," *Int. J. Comput. Games Technol.*, vol. 11, p. 11, 2014.

[54] G. Mcallister, P. Mirza-babaei, and J. Avent, "Game Analytics," *Game Anal.*, pp. 621–638, 2013.

[55] K. Kiili, K. Moeller, and M. Ninaus, "Evaluating the effectiveness of a game-based rational number training - In-game metrics as learning indicators," *Comput. Educ.*, vol. 120, pp. 13–28, 2018.

[56] C. S. Loh and Y. Sheng, "Performance metrics for serious games: Will the (real) expert please step forward?," *Int. Conf. Comput. Games AI, Animat. Mobile, Interact. Multimedia, Educ. Serious Games*, pp. 202–206, 2013.

[57] I. Doumanis and S. Smith, "Validation of Games for Behavioral Change : Connecting the Playful and Serious," *Int. J. Serious Games*, vol. 2, no. 3, pp. 63–75, 2015.

[58] I. Mayer *et al.*, "The research and evaluation of serious games: Toward a comprehensive methodology," *Br. J. Educ. Technol.*, vol. 45, no. 3, pp. 502–527, May 2014.

[59] S. Arnab *et al.*, "Mapping learning and game mechanics for serious games analysis," *Br. J. Educ. Technol.*, vol. 46, no. 2, pp. 391–411, Mar. 2015.

[60] C. Girard, J. Ecalle, and A. Magnan, "Serious games as new educational tools: How

effective are they? A meta-analysis of recent studies," *J. Comput. Assist. Learn.*, vol. 29, no. 3, pp. 207–219, Jun. 2013.
[61] E. M. Raybourn and N. Bos, "Design and evaluation challenges of serious games," in *Conference on Human Factors in Computing Systems - Proceedings*, 2005, pp. 2049–2050.
[62] F. Bellotti, B. Kapralos, K. Lee, P. Moreno-Ger, and R. Berta, "Assessment in and of serious games," *Adv. Human-Computer Interact.*, p. 11, Jan. 2013.
[63] L. Nacke, A. Drachen, and S. Gobel, "Methods for Evaluating Gameplay Experience in a Serious Gaming Context," *Electron. J. e-Learning*, vol. 10, no. 2, pp. 172–184, 2012.
[64] J. A. Vargas, L. García-Mundo, M. Genero, and M. Piattini, "A systematic mapping study on Serious Game quality," in *ACM International Conference Proceeding Series*, 2014, pp. 1–10.
[65] K. Larson, "Serious Games and Gamification in the Corporate Training Environment: a Literature Review," *TechTrends*, vol. 64, no. 2, pp. 319–328, Mar. 2020.
[66] A. De Gloria, F. Bellotti, and R. Berta, "Serious Games for education and training," *Int. J. Serious Games*, vol. 1, no. 1, Feb. 2014.
[67] G. P. Papanastasiou, A. S. Drigas, and C. Skianis, "Serious Games in Preschool and Primary Education: Benefits And Impacts on Curriculum Course Syllabus," *Int. J. Emerg. Technol. Learn.*, vol. 12, no. 01, pp. 44–56, Jan. 2017.
[68] R. M. Bottino, M. Ott, and M. Tavella, "Serious Gaming at School: Reflections on Students' Performance, Engagement and Motivation," *Int. J. Game-Based Learn.*, vol. 4, no. 1, pp. 21–36, 2014.
[69] Y. Zhonggen, "A Meta-Analysis of Use of Serious Games in Education over a Decade," *Int. J. Comput. Games Technol.*, vol. 1, pp. 1–8, 2019.
[70] F. Xu, D. Buhalis, and J. Weber, "Serious games and the gamification of tourism," *Tour. Manag.*, vol. 60, pp. 244–256, Jun. 2017.
[71] S. Nuanmeesri, "Development of community tourism enhancement in emerging cities using gamification and adaptive tourism recommendation," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 10, pp. 8549–8563, Nov. 2022.
[72] S. Bampatzia, I. Bourlakos, A. Antoniou, C. Vassilakis, G. Lepouras, and M. Wallace, "Serious games: Valuable tools for cultural heritage," *Lect. Notes Comput. Sci.*, vol. 10056 LNCS, pp. 331–341, 2016.
[73] W. S. Ravyse, A. Seugnet Blignaut, V. Leendertz, and A. Woolner, "Success factors for serious games to enhance learning: a systematic review," *Virtual Real.*, vol. 21, no. 1, pp. 31–58, Mar. 2017.
[74] F. Ricciardi and L. T. De Paolis, "A Comprehensive Review of Serious Games in Health Professions," *Int. J. Comput. Games Technol.*, vol. 9, p. 9, 2014.
[75] N. Sharifzadeh *et al.*, "Health Education Serious Games Targeting Health Care Providers, Patients, and Public Health Users: Scoping Review," *JMIR Serious Games*, vol. 8, no. 1, p. e13459, Mar. 2020.
[76] Y. Wang *et al.*, "Application of Serious Games in Health Care: Scoping Review and Bibliometric Analysis," *Front. Public Heal.*, vol. 10, pp. 1–12, 2022.
[77] D. A. Lieberman and S. J. Brown, "Designing interactive video games for children's health education," *Interact. Technol. new Paradig. Healthc.*, pp. 201–210, 1995.
[78] W. Harris, Lynne; DeShazo, John; Pratt, "Diabetes and obesity: Can video games help?," *Health Informatics J.*, pp. 131–150, 2010.
[79] M. Ulbrich *et al.*, "Advantages of a Training Course for Surgical Planning in Virtual

Reality for Oral and Maxillofacial Surgery : Crossover Study," *JMIR Serious Games*, vol. 11, pp. 1–15, 2023.

[80] X. Zhang and E. Lai, "A Web-based Gaming Approach to Decrease HIV-related Stigma: Game Development and Mixed Methods Evaluation," *JMIR Serious Games*, vol. 10, pp. 1–13, 2022.

[81] M. Orumaa *et al.*, "Impact of Mobile Game FightHPV on Cervical Cancer Screening Attendance: Retrospective Cohort Study," *JMIR Serious Games*, vol. 10, pp. 1–13, 2022.

[82] A. Schättin *et al.*, "Development of a Novel Home-Based Exergame with On-body Feedback: A Usability Study," *JMIR Serious Games*, vol. 10, 2022.

[83] S. Langener, R. Klaassen, J. VanDerNagel, and D. Heylen, "Immersive Virtual Reality Avatars for Embodiment Illusions in People with Mild to Borderline Intellectual Disability: User-Centered Development and Feasibility Study," *JMIR Serious Games*, vol. 10, no. 4, 2022.

[84] S. Warsinsky, M. Schmidt-Kraepelin, S. Rank, S. Thiebes, and A. Sunyaev, "Conceptual Ambiguity Surrounding Gamification and Serious Games in Health Care: Literature Review and Development of Game-Based Intervention Reporting Guidelines (GAMING)," *J Med Internet Res*, vol. 23, no. 9, p. e30390, Sep. 2021.

[85] T. Baranowski, R. Buday, D. I. Thompson, and J. Baranowski, "Playing for Real: Video Games and Stories for Health-Related Behavior Change," *Am. J. Prev. Med.*, vol. 34, no. 1, pp. 74-82.e10, Jan. 2008.

[86] A. Perttula, K. Kiili, A. Lindstedt, and P. Tuomi, "Flow experience in game based learning – a systematic literature review," *Int. J. Serious Games*, vol. 4, no. 1, Mar. 2017.

[87] F. A. M. da Silva, T. S. da Silva, and E. R. Zorzal, "Use of serious games in medicine: a literature revision," *Res. Soc. Dev.*, vol. 10, no. 16, p. e480101624208, 2021.

[88] T. Kaukoranta, J. Smed, and H. Hakonen, "Understanding Pattern Recognition Methods," in *AI Game Programming Wisdom*, 2nd ed., Charles River Media, 2003, pp. 579–589.

[89] B. Bontchev, "Adaptation in affective video games: A literature review," *Cybern. Inf. Technol.*, vol. 16, no. 3, pp. 3–34, 2016.

[90] M. Csikszentmihalyi, *Flow: the psychology of optimal experience*. Harper Perennial Modern Classics, 2008.

[91] V. Shute, F. Ke, and L. Wang, "Assessment and Adaptation in Games," *Instr. Tech. to Facil. Learn. Motiv. Serious Games*, pp. 59–78, 2017.

[92] A. Streicher and J. D. Smeddinck, "Personalized and adaptive serious games," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9970 LNCS, pp. 332–377, 2016.

[93] D. Charles *et al.*, "Player-Centred Game Design: Player Modelling and Adaptive Digital Games," in *DiGRA Conference: Changing Views - Worlds in Play*, 2005.

[94] R. Steinmetz and K. Nahrstedt, "Multimedia Applications," *arXiv:2208.04548 [cs.HC]*, pp. 197–214, 2004.

[95] R. V Aranha *et al.*, "EasyAffecta: A framework to develop serious games for virtual rehabilitation with affective adaptation," *Multimed. Tools Appl. 2022*, pp. 1–26, Jun. 2022.

[96] G. N. Yannakakis, P. Spronck, D. Loiacono, and E. Andre, "Player Modeling," in *Dagstuhl Seminar on Artificial and Computational Intelligence in Games*, 2013.

[97] A. M. Smith, C. Lewis, K. Hullet, and A. Sullivan, "An inclusive view of player modeling," *Proc. 6th Int. Conf. Found. Digit. Games, FDG 2011*, pp. 301–303, 2011.

[98] J. Tao and T. Tan, “Affective computing: A review,” *Lect. Notes Comput. Sci.*, vol. 3784 LNCS, pp. 981–995, 2005.
[99] E. Gurney, “‘What is an Emotion?,’” *Mind*, vol. 9, no. 35, pp. 421–426, Dec. 1884.
[100] C. L. Lisetti, “Affective computing,” *Pattern Anal. Appl.*, vol. 1, no. 1, pp. 71–73, 1998.
[101] E. Hudlicka, “Affective computing for game design,” *Int. North-American Conf. Intell. Games Simul.*, pp. 5–12, 2008.
[102] D. Setiono, D. Saputra, K. Putra, J. V. Moniaga, and A. Chowanda, “Enhancing Player Experience in Game with Affective Computing,” *Procedia Comput. Sci.*, vol. 179, no. 2019, pp. 781–788, 2021.
[103] S. Hamdy and D. King, “Affective games: Adaptation and design,” in *International Conference on Intelligent Games and Simulation*, 2019, pp. 11–18.
[104] M. A. A. Dewan, M. Murshed, and F. Lin, “Engagement detection in online learning: a review,” *Smart Learn. Environ. 2019 61*, vol. 6, no. 1, pp. 1–20, Jan. 2019.
[105] S. Khan and T. Colella, “Inconsistencies in Measuring User Engagement in Virtual Learning - A Critical Review,” *arXiv:2208.04548 [cs.HC]*, 2021.
[106] J. A. Fredricks and W. McColskey, “The Measurement of Student Engagement: A Comparative Analysis of Various Methods and Student Self-report Instruments,” *Handb. Res. Student Engagem.*, pp. 763–782, Jan. 2012.
[107] B. Ahmed, “Analyzing Player Engagement for Biofeedback,” *IEEE Int. Conf. Serious Games Appl. Heal.*, pp. 1–5, 2019.
[108] H. Al Osman, H. Dong, and A. El Saddik, “Ubiquitous Biofeedback Serious Game for Stress Management,” *IEEE Access*, vol. 4, pp. 1274–1286, 2016.
[109] A. Rodriguez *et al.*, “A VR-based serious game for studying emotional regulation in adolescents,” *IEEE Comput. Graph. Appl.*, vol. 35, no. 1, pp. 65–73, Jan. 2015.
[110] H. Schoenau-Fog, “The Player Engagement Process-An Exploration of Continuation Desire in Digital Games,” in *DiGRA International Conference: Think Design Play*, 2011.
[111] R. M. Martey *et al.*, “Measuring Game Engagement: Multiple Methods and Construct Complexity,” *Simul. Gaming*, vol. 45, pp. 528–547, Jan. 2014.
[112] S. D’Mello, P. Chipman, and A. Graesser, “Posture as a predictor of learner’s affective engagement,” *Proc. 29th Annu. Meet. Cogn. Sci. Soc.*, vol. 1, pp. 905–910, 2007.
[113] D. Bibbo, M. Carli, S. Conforto, and F. Battisti, “A sitting posture monitoring instrument to assess different levels of cognitive engagement,” *Sensors*, vol. 19, no. 3, 2019.
[114] B. Cowley, N. Ravaja, and T. Heikura, “Cardiovascular physiology predicts learning effects in a serious game activity,” *Comput. Educ.*, vol. 60, no. 1, pp. 299–309, 2013.
[115] A. M. Porter and P. Goolkasian, “Video games and stress: How stress appraisals and game content affect cardiovascular and emotion outcomes,” *Front. Psychol.*, vol. 10, pp. 1–13, 2019.
[116] P. Bouvier, K. Sehaba, E. Lavoue, and S. George, “Using traces to qualify learner’s engagement in game-based learning,” *IEEE Int. Conf. Adv. Learn. Technol.*, pp. 432–436, 2013.
[117] M. J. Callaghan, N. McShane, and A. G. Eguiluz, “Using game analytics to measure student engagement/retention for engineering education,” *Int. Conf. Remote Eng. Virtual Instrum.*, pp. 297–302, 2014.
[118] M. Hendrikx, S. Meijer, J. Van Der Velden, and A. Iosup, “Procedural content generation for games: A survey,” *ACM Trans. Multimed. Comput. Commun. Appl.*, vol. 9, no. 1, 2013.
[119] G. Smith, E. Gan, A. Othenin-Girard, and J. Whitehead, “PCG-based game design:

Enabling new play experiences through procedural content generation," *ACM Int. Conf. Proceeding Ser.*, 2011.
[120] M. Hafis, H. Tolle, and A. A. Supianto, "A literature review of Empirical Evidence on Procedural Content Generation in Game-Related Implementation," *J. Inf. Technol. Comput. Sci.*, vol. 4, no. 3, pp. 308–328, 2019.
[121] M. Frutos-Pascual and B. G. Zapirain, "Review of the Use of AI Techniques in Serious Games: Decision Making and Machine Learning," *IEEE Trans. Comput. Intell. AI Games*, vol. 9, no. 2, pp. 133–152, Jun. 2017.
[122] J. Liu, S. Snodgrass, A. Khalifa, S. Risi, G. N. Yannakakis, and J. Togelius, "Deep learning for procedural content generation," *Neural Comput. Appl.*, vol. 33, no. 1, pp. 19–37, Jan. 2021.
[123] A. Summerville *et al.*, "Procedural content generation via machine learning (PCGML)," *IEEE Trans. Games*, vol. 10, no. 3, pp. 257–270, Sep. 2018.
[124] S. Singh, A. G. Barto, and N. Chentanez, "Intrinsically motivated reinforcement learning," in *Advances in Neural Information Processing Systems*, 2005.
[125] K. Mitsis, E. Kalafatis, K. Zarkogianni, G. Mourkousis, and K. S. Nikita, "Procedural content generation based on a genetic algorithm in a serious game for obstructive sleep apnea," in *IEEE Conference on Computatonal Intelligence and Games, CIG*, 2020, vol. 2020-August, pp. 694–697.
[126] P. Radanliev *et al.*, "COVID-19 what have we learned? The rise of social machines and connected devices in pandemic management following the concepts of predictive, preventive and personalized medicine.," *EPMA J.*, vol. 11, no. 3, pp. 311–332, Sep. 2020.
[127] I. K. Valavanis, S. G. Mougiakakou, K. A. Grimaldi, and K. S. Nikita, "A multifactorial analysis of obesity as CVD risk factor: use of neural network based methods in a nutrigenetics context," *BMC Bioinformatics*, vol. 11, Sep. 2010.
[128] M. Skevofilakas, K. Zarkogianni, B. G. Karamanos, and K. S. Nikita, "A hybrid Decision Support System for the risk assessment of retinopathy development as a long term complication of Type 1 Diabetes Mellitus," *Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, pp. 6713–6716, 2010.
[129] S. Ahmad, F. Mehmood, F. Khan, and T. K. Whangbo, "Architecting Intelligent Smart Serious Games for Healthcare Applications: A Technical Perspective," *Sensors 2022, Vol. 22, Page 810*, vol. 22, no. 3, p. 810, Jan. 2022.
[130] K. Zarkogianni *et al.*, "A Review of Emerging Technologies for the Management of Diabetes Mellitus," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 12, pp. 2735–2749, Dec. 2015.
[131] A. Waldmann, J. Haladjian, D. Ismailovi, and B. Brügge, "The Square Dance Framework - A framework approach for module-based serious games," in *IADIS International Conference Game and Entertainment Technologies*, 2012, pp. 63–66.
[132] A. J. A. Seyderhelm, K. L. Blackmore, and K. Nesbitt, "Towards Cognitive Adaptive Serious Games: A Conceptual Framework," *Lect. Notes Comput. Sci.*, pp. 331–338, 2019.
[133] A. Yusoff, R. Crowder, L. Gilbert, and G. Wills, "A conceptual framework for serious games," *IEEE Int. Conf. Adv. Learn. Technol.*, pp. 21–23, 2009.
[134] R. Lalwani *et al.*, "I-Mouse: A Framework for Player Assistance in Adaptive Serious Games," *Lect. Notes Comput. Sci.*, pp. 234–238, 2021.
[135] S. Blatsios and R. Ioannis, *An adaptation and personalisation methodology for Serious Games design*. 2019.

[136] I. Afyouni, A. Murad, and A. Einea, “Adaptive Rehabilitation Bots in Serious Games,” *Sensors*, vol. 20, no. 24, p. 7037, Dec. 2020.

[137] S. Ahmad, F. Mehmood, F. Khan, and T. K. Whangbo, “Architecting Intelligent Smart Serious Games for Healthcare Applications: A Technical Perspective,” *Sensors 2022, Vol. 22, Page 810*, vol. 22, no. 3, p. 810, Jan. 2022.

[138] S. Hardy, A. El Saddik, S. Gobel, and R. Steinmetz, “Context aware serious games framework for sport and health,” in *IEEE International Symposium on Medical Measurements and Applications*, 2011, pp. 248–252.

[139] M. Ponsen, “Improving adaptive game AI with evolutionary learning,” *Comput. Sci.*, 2004.

[140] A. Zafar, H. Mujtaba, and M. O. Beg, “Search-based procedural content generation for GVG-LG,” *Appl. Soft Comput.*, vol. 86, p. 105909, Jan. 2020.

[141] A. B. Moghadam and M. K. Rafsanjani, “A genetic approach in procedural content generation for platformer games level creation,” *Conf. Swarm Intell. Evol. Comput.*, pp. 141–146, Jun. 2017.

[142] R. G. de Pontes and H. M. Gomes, “Evolutionary procedural content generation for an endless platform game,” *Brazilian Symp. Games Digit. Entertain. SBGAMES*, vol. 2020-Novem, pp. 80–89, 2020.

[143] R. Lara-Cabrera, C. Cotta, and A. J. Fernández-Leiva, “Using self-adaptive evolutionary algorithms to evolve dynamism-oriented maps for a real time strategy game,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8353 LNCS, pp. 256–263, 2014.

[144] E. Soares De Lima, B. Feijo, and A. L. Furtado, “Procedural Generation of Quests for Games Using Genetic Algorithms and Automated Planning,” in *Brazilian Symposium on Games and Digital Entertainment*, 2019, no. 2016, pp. 144–153.

[145] M. A. Verma and P. W. McOwan, “An adaptive methodology for synthesising mobile phone games using genetic algorithms,” in *IEEE Congress on Evolutionary Computation*, 2005, vol. 1, pp. 864–879.

[146] D. Hooshyar, M. Yousefi, M. Wang, and H. Lim, “A data-driven procedural-content-generation approach for educational games,” *J. Comput. Assist. Learn.*, vol. 34, no. 6, pp. 731–739, Dec. 2018.

[147] M. İnce, “BiLSTM and dynamic fuzzy AHP-GA method for procedural game level generation,” *Neural Comput. Appl.*, vol. 33, no. 15, pp. 9761–9773, 2021.

[148] B. Hssina and M. Erritali, “A Personalized Pedagogical Objectives Based on a Genetic Algorithm in an Adaptive Learning System,” *Procedia Comput. Sci.*, vol. 151, pp. 1152–1157, Jan. 2019.

[149] K. Mitsis, K. Zarkogianni, N. Bountouni, M. Athanasiou, and K. S. Nikita, “An Ontology-Based Serious Game Design for the Development of Nutrition and Food Literacy Skills,” *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, pp. 1405–1408, Jul. 2019.

[150] E. Kalafatis *et al.*, “Artificial Intelligence based procedural content generation in serious games for health: The case of childhood obesity,” in *International Conference on Wireless Mobile Communication and Healthcare*, 2022.

[151] World Health Organization, “The Nutrition Challenge. Food system solutions,” *World Heal. Organ.*, vol. 22, no. 1, pp. 4–7, 2018.

[152] D. Mozaffarian, “Dietary and Policy Priorities for Cardiovascular Disease, Diabetes, and Obesity – A Comprehensive Review,” *Circulation*, vol. 133, no. 2, p. 187, Jan. 2016.

[153] F. Deer, T. Falkenberg, B. Mcmillan, and L. Sims, *Sustainable well-being: concepts,*

issues, and educational practices. ESWB-press, 2014.
[154] C. Krause, K. Sommerhalder, S. Beer-Borst, and T. Abel, “Just a subtle difference? Findings from a systematic review on definitions of nutrition literacy and food literacy,” *Health Promot. Int.*, vol. 33, no. 3, pp. 378–389, 2018.
[155] K. J. Silk, J. Sherry, B. Winn, N. Keesecker, M. A. Horodynski, and A. Sayir, “Increasing nutrition literacy: testing the effectiveness of print, web site, and game modalities,” *J. Nutr. Educ. Behav.*, vol. 40, no. 1, pp. 3–10, Jan. 2008.
[156] E. Y. N. Yuen, M. Thomson, and H. Gardiner, “Measuring Nutrition and Food Literacy in Adults: A Systematic Review and Appraisal of Existing Measurement Tools,” *Heal. Lit. Res. Pract.*, vol. 2, no. 3, Jul. 2018.
[157] E. Truman, D. Lane, and C. Elliott, “Defining food literacy: A scoping review,” *Appetite*, vol. 116, pp. 365–371, Sep. 2017.
[158] D. Thompson *et al.*, “Serious video games for health: How behavioral science guided the development of a serious video game,” *Simul. Gaming*, vol. 41, no. 4, pp. 587–606, 2010.
[159] “Bbc - Ontologies - Food Ontology - RINWUNS.” [Online]. Available: https://www.rinwuns.com/bbc-ontologies-food-ontology/. [Accessed: 30-Jun-2022].
[160] S. Heintz and E. L. C. Law, “The game genre map: A revised game classification,” *Annu. Symp. Comput. Interact. Play*, pp. 175–184, Oct. 2015.
[161] E. L. Deci and R. M. Ryan, “Self-determination theory: A macrotheory of human motivation, development, and health,” *Can. Psychol.*, vol. 49, no. 3, pp. 182–185, Aug. 2008.
[162] E. Cortés-Reyes, K. Parrado-Bermúdez, and F. Escobar-Córdoba, “New perspectives in the treatment of obstructive sleep apnea–hypopnea syndrome,” *Colomb. J. Anesthesiol.*, vol. 45, no. 1, pp. 62–71, Jan. 2017.
[163] A. V. Benjafield *et al.*, “Estimation of the global prevalence and burden of obstructive sleep apnoea: a literature-based analysis,” *Lancet. Respir. Med.*, vol. 7, no. 8, pp. 687–698, Aug. 2019.
[164] K. N. K. Mitsis, E. Kalafatis, I. Giannaki, I. Malama, M. Christodoulakis, K. Zarkogianni, “A serious game approach to raise awareness and promote self-managementof obstructive sleep apnea,” in *Evevit*, 2019, p. 80.
[165] M. A. Camara, Y. Castillo, D. Blanco-Almazan, L. Estrada, and R. Jane, “MHealth tools for monitoring Obstructive Sleep Apnea patients at home: Proof-of-concept,” in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, 2017, pp. 1555–1558.
[166] K. Archontogeorgis, E. Nena, N. Papanas, and P. Steiropoulos, “Biomarkers to improve diagnosis and monitoring of obstructive sleep apnea syndrome: Current status and future perspectives,” *Pulm. Med.*, vol. 2014, 2014.
[167] J. Rebelo, P. D. Gaspar, V. N. G. J. Soares, and J. M. L. P. Caldeira, “A Novel mHealth Approach for the Monitoring and Assisted Therapeutics of Obstructive Sleep Apnea,” *Appl. Sci. 2022, Vol. 12, Page 10257*, vol. 12, no. 20, p. 10257, Oct. 2022.
[168] M. Kordaki and A. Gousiou, “Digital card games in education: A ten year systematic review,” *Comput. Educ.*, vol. 109, pp. 122–161, Jun. 2017.
[169] H. Rastegarpour and P. Marashi, “The effect of card games and computer games on learning of chemistry concepts,” *Procedia - Soc. Behav. Sci.*, vol. 31, pp. 597–601, Jan. 2012.
[170] E. Marchetti and A. Valente, “Learning via Game Design: From Digital to Card Games

and Back Again," *Electron. J. e-Learning*, vol. 13, no. 3, p. 167-180, Mar. 2015.
[171] C. M. Odenweller, C. T. Hsu, and S. E. Dicarlo, "Educational card games for understanding gastrointestinal physiology.," *Am. J. Physiol.*, vol. 275, no. 6, pp. 78–84, 1998.
[172] I. Vasilakis *et al.*, "The ENDORSE Feasibility Pilot Trial: Assessing the Implementation of Serious Games Strategy and Artificial Itelligence-Based Telemedicine in Glycemic Control Improvement.No Title," in *Diabetes Technology & Therapeutics*, 2022.
[173] D. Nathan *et al.*, "The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus.," *N. Engl. J. Med.*, vol. 329, no. 14, pp. 977–986, Sep. 1993.
[174] S. M. Fruh, "Obesity: Risk factors, complications, and strategies for sustainable long-term weight management," *J. Am. Assoc. Nurse Pract.*, vol. 29, no. 1, p. S3, Oct. 2017.
[175] D. Hooshyar, M. Yousefi, and H. Lim, "Data-driven approaches to game player modeling: A systematic literature review," *ACM Comput. Surv.*, vol. 50, no. 6, 2018.
[176] G. N. Yannakakis and J. Togelius, "Artificial intelligence and games," *Artif. Intell. Games*, pp. 1–337, 2018.
[177] "Procedural Content Generation in Serious Games | EAI Blog." [Online]. Available: https://blog.eai-conferences.org/2015/05/27/procedural-content-generation-in-serious-games/. [Accessed: 30-Oct-2021].
[178] K. Mitsis, K. Zarkogianni, K. Dalakleidi, G. Mourkousis, and K. S. Nikita, "Evaluation of a serious game promoting nutrition and food literacy: Experiment design and preliminary results," *IEEE Int. Conf. Bioinforma. Bioeng.*, pp. 497–502, Oct. 2019.
[179] N. Kliemann, J. Wardle, F. Johnson, and H. Croker, "Reliability and validity of a revised version of the General Nutrition Knowledge Questionnaire," *Eur. J. Clin. Nutr.*, vol. 70, no. 10, pp. 1174–1180, Oct. 2016.
[180] T. Lazou, M. Georgiadis, K. Pentieva, A. McKevitt, and E. Iossifidou, "Food safety knowledge and food-handling practices of Greek university students: A questionnaire-based survey," *Food Control*, vol. 28, no. 2, pp. 400–411, Dec. 2012.
[181] E. Y. N. Yuen, M. Thomson, and H. Gardiner, "Measuring Nutrition and Food Literacy in Adults: A Systematic Review and Appraisal of Existing Measurement Tools," *Heal. Lit. Res. Pract.*, vol. 2, no. 3, Jul. 2018.
[182] W. A. Ijsselsteijn, D. Kort, and Y. A. W. & Poels, "Game experience questionnaire," *Hum. Technol. Interact.*, 2013.
[183] P. Lopes, G. N. Yannakakis, and A. Liapis, "RankTrace: Relative and unbounded affect annotation," *2017 7th Int. Conf. Affect. Comput. Intell. Interact. ACII 2017*, vol. 2018-Janua, pp. 158–163, 2018.
[184] "Mega 2560 Rev3 | Arduino Documentation | Arduino Documentation." [Online]. Available: https://docs.arduino.cc/hardware/mega-2560. [Accessed: 08-Jan-2022].
[185] E. Fragkiadakis, K. V. Dalakleidi, and K. S. Nikita, "Design and Development of a Sitting Posture Recognition System," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2019, pp. 3364–3367.
[186] "Sensitronics, 2019." [Online]. Available: https://www.sensitronics.com/fsr101.htm. [Accessed: 24-Dec-2021].
[187] "PulseSensor." [Online]. Available: https://pulsesensor.com/. [Accessed: 24-Dec-2021].
[188] F. Shaffer and J. P. Ginsberg, "An Overview of Heart Rate Variability Metrics and Norms," *Front. Public Heal.*, vol. 5, no. September, pp. 1–17, 2017.

[189] M. L. Munoz *et al.*, “Validity of (Ultra-)Short Recordings for Heart Rate Variability Measurements,” *PLoS One*, vol. 10, no. 9, Sep. 2015.
[190] L. Salahuddin, J. Cho, M. G. Jeong, and D. Kim, “Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings,” in *Annual International Conference of the IEEE Engineering in Medicine and Biology*, 2007, pp. 4656–4659.
[191] G. N. Yannakakis and J. Hallam, “Towards optimizing entertainment in computer games,” *Appl. Artif. Intell.*, vol. 21, no. 10, pp. 933–971, Nov. 2007.
[192] S. Poria, E. Cambria, R. Bajpai, and A. Hussain, “A review of affective computing: From unimodal analysis to multimodal fusion,” *Inf. Fusion*, vol. 37, pp. 98–125, Sep. 2017.
[193] J. Hernandez-Gonzalez, I. Inza, and J. A. Lozano, “A Note on the Behavior of Majority Voting in Multi-Class Domains with Biased Annotators,” *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 1, pp. 195–200, Jan. 2019.
[194] A. DeSmet *et al.*, “A meta-analysis of serious digital games for healthy lifestyle promotion,” *Prev. Med. (Baltim).*, vol. 69, pp. 95–107, Dec. 2014.
[195] A. C. P. D. Cargnin, M. Cordeiro d’Ornellas, “A Serious Game for Upper Limb Stroke Rehabilitation Using Biofeedback and Mirror-Neurons Based Training,” *Stud. Health Technol. Inform.*, vol. 216, pp. 348–52, 2015.
[196] S. Nosier and R. Salah, “Forecasting Covid-19 Infections and Deaths Horizon in Egypt,” *medRxiv*, Sep. 2020.
[197] A. Fuchslocher, K. Gerling, M. Masuch, and N. Krämer, “Evaluating social games for kids and teenagers diagnosed with cancer,” *IEEE Int. Conf. Serious Games Appl. Heal.*, pp. 1–4, 2011.
[198] L. G. R. De Lima, A. De Lima Salgado, and A. P. Freire, “Evaluation of the user experience and Intrinsic motivation with educational and mainstream digital games,” *Lat. Am. Conf. Hum. Comput. Interact.*, vol. 11, p. 1, Nov. 2015.
[199] E. N. Castellar, K. Oksanen, and J. Van Looy, “Assessing game experience: Heart rate variability, in-game behavior and self-report measures,” *Int. Work. Qual. Multimed. Exp.*, pp. 292–296, Dec. 2014.
[200] C. Alonso-Fernández, A. R. Cano, A. Calvo-Morata, M. Freire, I. Martínez-Ortiz, and B. Fernández-Manjón, “Lessons learned applying learning analytics to assess serious games,” *Comput. Human Behav.*, vol. 99, pp. 301–309, Oct. 2019.
[201] S. D’mello and A. Graesser, “Mining bodily patterns of affective experience during learning,” in *International Conference on Educational Data Mining*, 2010, pp. 31–40.
[202] A. B. Moghadam and M. K. Rafsanjani, “A genetic approach in procedural content generation for platformer games level creation,” *Conf. Swarm Intell. Evol. Comput.*, pp. 141–146, Jun. 2017.
[203] Y. H. Pereira, R. Ueda, L. B. Galhardi, and J. D. Brancher, “Using Procedural Content Generation for Storytelling in a Serious Game Called Orange Care,” in *Brazilian Symposium on Games and Digital Entertainment*, 2019, pp. 192–197.
[204] S. Carlier, S. Van der Paelt, F. Ongenae, F. De Backere, and F. De Turck, “Empowering Children with ASD and Their Parents: Design of a Serious Game for Anxiety and Stress Reduction,” *Sensors*, vol. 20, no. 4, Feb. 2020.
[205] M. L. Dontje, M. De Groot, R. R. Lengton, C. P. Van Der Schans, and W. P. Krijnen, “Measuring steps with the Fitbit activity tracker: an inter-device reliability study,” *J. Med. Eng. Technol.*, vol. 39, no. 5, pp. 286–290, Jul. 2015.
[206] S. P. Shultz, R. C. Browning, Y. Schutz, C. Maffeis, and A. P. Hills, “Childhood obesity and walking: guidelines and challenges,” *Int. J. Pediatr. Obes.*, vol. 6, no. 5–6, pp. 332–341,

Oct. 2011.
[207] M. Hirshkowitz *et al.*, “National Sleep Foundation’s sleep time duration recommendations: methodology and results summary,” *Sleep Heal.*, vol. 1, no. 1, pp. 40–43, Mar. 2015.
[208] S. Fizek, “Automated State of Play,” *Digit. Cult. Soc.*, vol. 4, no. 1, pp. 201–214, Mar. 2018.
[209] S. Ariyurek, A. Betin-Can, and E. Surer, “Automated Video Game Testing Using Synthetic and Humanlike Agents,” in *IEEE Transactions on Games*, 2021, vol. 13, no. 1, pp. 50–67.
[210] C. Politowski, F. Petrillo, and Y.-G. Guéhéneuc, “A Survey of Video Game Testing,” *arXiv:2103.06431 [cs.SE]*, 2021.
[211] A. Zook, E. Fruchter, and M. O. Riedl, “Automatic Playtesting for Game Parameter Tuning via Active Learning,” *arXiv:1908.01417 [cs.AI]*, 2019.
[212] J. Pfau, A. Liapis, G. Volkmar, G. N. Yannakakis, and R. Malaka, “Dungeons Replicants: Automated Game Balancing via Deep Player Behavior Modeling,” *IEEE Conf. Comput. Intell. Games*, pp. 431–438, 2020.
[213] J. T. Kristensen, A. Valdivia, and P. Burelli, “Estimating Player Completion Rate in Mobile Puzzle Games Using Reinforcement Learning,” *IEEE Conf. Comput. Intell. Games*, pp. 636–639, 2020.
[214] S. G. Ansari, “Toward Automated Assessment of User Experience in Extended Reality,” *Int. Conf. Softw. Testing, Verif. Valid.*, pp. 430–432, 2020.
[215] J. Bergdahl, C. Gordillo, K. Tollmar, and L. Gisslen, “Augmenting Automated Game Testing with Deep Reinforcement Learning,” in *IEEE Conference on Computatonal Intelligence and Games*, 2020, pp. 600–603.
[216] B. Krylov, M. Abramov, and A. Khlobystova, “Automated Player Activity Analysis for a Serious Game About Social Engineering,” *Stud. Syst. Decis. Control*, vol. 337, pp. 587–599, 2021.
[217] R. Qasrawi, M. Amro, and R. Jayousi, “Automatic analytics model for learning skills analysis using game player data and robotic process automation in a serious game for education,” *Int. Conf. Promis. Electron. Technol.*, pp. 94–98, Dec. 2020.
[218] E. Rowe, W. J. Hawkins, R. S. Baker, J. Asbell-Clarke, and E. Kasman, “Building Automated Detectors of Gameplay Strategies to Measure Implicit Science Learning,” in *Educational Data Mining*, 2014.
[219] K. Sipiyaruk, S. Hatzipanagos, P. A. Reynolds, and J. E. Gallagher, “Serious Games and the COVID-19 Pandemic in Dental Education: An Integrative Review of the Literature,” *Computers*, vol. 10, no. 4, p. 42, Apr. 2021.
[220] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, “Deep reinforcement learning: A brief survey,” *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 26–38, 2017.
[221] E. F. Morales and J. H. Zaragoza, “An introduction to reinforcement learning,” *Decis. Theory Model. Appl. Artif. Intell. Concepts Solut.*, pp. 63–80, 2011.
[222] Y. Fenjiro and H. Benbrahim, “Deep Reinforcement Learning Overview of the State of the Art,” *J. Autom. Mob. Robot. Intell. Syst.*, vol. 12, no. 3, 2018.
[223] K. Shao, Z. Tang, Y. Zhu, N. Li, and D. Zhao, “A Survey of Deep Reinforcement Learning in Video Games,” *arXiv:1912.10944 [cs.MA]*, pp. 1–13, 2019.
[224] V. Mnih *et al.*, “Playing Atari with Deep Reinforcement Learning,” *arXiv:1312.5602 [cs.LG]*, Dec. 2013.

[225] D. Silver *et al.*, “Mastering the game of Go with deep neural networks and tree search,” *Nature*, vol. 529, no. 7587, pp. 484–489, Jan. 2016.
[226] G. Lample and D. S. Chaplot, “Playing FPS Games with Deep Reinforcement Learning,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2017, vol. 31, no. 1, pp. 2140–2146.
[227] A. Dobrovsky *et al.*, “Improving Adaptive Gameplay in Serious Games Through Interactive Deep Reinforcement Learning,” in *Cognitive Infocommunications, Theory and Applications*, Springer, Cham, 2019, pp. 411–432.
[228] M. Thielscher, “A General Game Description Language for incomplete information games,” in *Proceedings of the National Conference on Artificial Intelligence*, 2010, vol. 2, pp. 994–999.
[229] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal Policy Optimization Algorithms,” *arXiv:1707.06347 [cs.LG]*, pp. 1–12, 2017.
[230] “Gymnasium Documentation.” [Online]. Available: https://gymnasium.farama.org/. [Accessed: 29-Jan-2023].