An Automated Mobile Game-based Screening Tool for Patients with Alcohol Dependence

JITTRAPOL INTARASIRISAWAT*, CHEE SIANG ANG†, and CHRISTOS EFSTRATIOU, School of Engineering and Digital Arts, University of Kent, United Kingdom

LUKE WILLIAM FEIDHLIM DICKENS, Department of Information Studies, University College London, United Kingdom

NARANCHAYA SRIBURAPAR and DINKAR SHARMA, School of Psychology, University of Kent, United Kingdom

BURACHAI ASAWATHAWEEBOON, Department of Psychology, Faculty of Liberal Arts, Thammasat University, Thailand

Traditional methods for screening and diagnosis of alcohol dependence are typically administered by trained clinicians in medical settings and often rely on interview responses. These self-reports can be unintentionally or deliberately false, and misleading answers can, in turn, lead to inaccurate assessment and diagnosis. In this study, we examine the use of user-game interaction patterns on mobile games to develop an automated diagnostic and screening tool for alcohol-dependent patients. Our approach relies on the capture of interaction patterns during gameplay, while potential patients engage with popular mobile games on smartphones. The captured signals include gameplay performance, touch gestures, and device motion, with the intention of identifying patients with alcohol dependence. We evaluate the classification performance of various supervised learning algorithms on data collected from 40 patients and 40 age-matched healthy adults. The results show that patients with alcohol dependence can be automatically identified accurately using the ensemble of touch, device motion, and gameplay performance features on 3-minute samples (accuracy=0.95, sensitivity=0.95, and specificity=0.95). The present findings provide strong evidence suggesting the potential use of user-game interaction metrics on existing mobile games as discriminant features for developing an implicit measure to identify alcohol dependence conditions. In addition to supporting healthcare professionals in clinical decision-making, the game-based self-screening method could be used as a novel strategy to promote alcohol dependence screening, especially outside of clinical settings.

CCS Concepts: • **Human-centered computing** → *Touch screens*; *Haptic devices*; **Ubiquitous and mobile computing design and evaluation methods**; • **Applied computing** → **Health informatics**.

Additional Key Words and Phrases: screening measures, serious games, alcohol dependence, mobile health

---

*This is the first author

†This is the corresponding author

Authors' addresses: Jittrapol Intarasirisawat, ji81@kent.ac.uk; Chee Siang Ang, c.s.ang@kent.ac.uk; Christos Efstratiou, c.efstratiou@kent.ac.uk, School of Engineering and Digital Arts, University of Kent, Jennison Building, Canterbury, Kent, CT2 7NT, United Kingdom; Luke William Feidhlim Dickens, Department of Information Studies, University College London, Gower Street, London, WC1E 6BT, United Kingdom, l.dickens@ucl.ac.uk; Naranchaya Sriburapar, ns567@kent.ac.uk; Dinkar Sharma, d.sharma@kent.ac.uk, School of Psychology, University of Kent, Keynes College, Canterbury, Kent, CT2 7NP, United Kingdom; Burachai Asawathaweeboon, Department of Psychology, Faculty of Liberal Arts, Thammasat University, Pathumthani, 12121, Thailand, burachai.a@arts.tu.ac.th.

Proc. ACM Interact. Mob. Wearable Ubiquitous Technol., Vol. 4, No. 3, Article 110. Publication date: September 2020.

110

1 INTRODUCTION

Heavy chronic alcohol consumption behaviour is associated with increased risk of diseases, including mental health disorders, physical injuries, as well as other social and legal problems [7]. As defined in the International Classification of Diseases, Tenth Revision (ICD-10) [39], harmful alcohol use is a pattern of alcohol consumption that causes damages to physical or mental functioning, which can lead to adverse social consequences. Alcohol dependence is a more severe form of harmful alcohol use, which often develops after long-term excessive alcohol consumption. Typically, patients with alcohol dependence have impaired control over their use of alcohol and may report multiple unsuccessful attempts to quit despite the obvious harmful consequences. Unfortunately, it is common for individuals with covert symptoms of health conditions to remain unrecognised until serious complications have developed. In England, there are approximately more than half a million dependent drinkers. However, only 18% of those ever receive treatment [12, 13]. Globally, the World Health Organisation (WHO) estimates that 3 million people die prematurely every year as a result of excessive alcohol drinking, which is 5.3% of all global deaths [1]. Diagnostic interviews are standard measures for the diagnosis of alcohol dependence, which require both time and qualified professionals in administration to establish a reliable diagnosis.

Despite being regarded as a quick and easy screening instrument widely used in primary health care, the Alcohol Use Disorders Identification Test (AUDIT) was not explicitly developed to identify alcohol dependence, but rather to identify people with alcohol abuse who are likely to develop dependence symptoms based on hazardous drinking behaviours [22]. Moreover, this questionnaire-based measure significantly relies on retrospective self-reports, which are subject to unintentional or deliberate bias as respondents may be unaware of their actual alcohol intake or even deliberately mislead the counsellors to avoid embarrassment and, thus, may falsely report the amount of their alcohol use [37].

These limitations have led a call for an alternative method that is quick and easy to administer, yet provides a reliable and accurate assessment to improve the screening rate and accuracy for the early detection of alcohol dependence (especially one appropriate for use outside of clinical settings). Our work was motivated by the findings from previous studies which demonstrated that individuals with several medical conditions associated with cognitive impairment exhibited reduced abilities to fine motor movement [34, 45, 48, 49]. In particular, it was found in a recent study that by using a tablet to measure finger dexterity, patients with MCI and Alzheimer's Disease demonstrated significantly slower responses and higher contact duration fluctuation when compared to healthy adults [48]. Furthermore, touch gestures and device movement patterns were used as key features in an automated screening tool to identify children with autism from typically-developing children [6]. Collectively, this body of literature provides encouraging evidence suggesting that the combination of user-game behavioural data passively collected via multisensory input could produce fruitful results for identifying people with cognitively impaired conditions, especially with fine motor disturbance. Most importantly, prior work has demonstrated that alcohol dependence is strongly linked to both declines in cognitive abilities [20, 26], and irregular motor function of patients' hands [11, 33, 52]. Given that the interaction mode in mobile games typically involves touch gestures and device motions, we consider that smartphone games have the potential to capture both of these types of discriminant factors and therefore help develop an automated system for the diagnosis of alcohol dependence.

In this work, we introduce a novel approach to identify patients with alcohol dependence using user-game interaction patterns. We evaluate the classification performance of various supervised learning algorithms

based on the game interaction-based data collected from a total of 40 alcohol-dependent patients and 40 age-matched healthy adults. Furthermore, to examine the feasibility of our automated system as a short self-screening instrument, we also assess and compare the model performance on data with different gameplay durations.

Our results show that patients with alcohol dependence can be automatically identified accurately using an ensemble of touch, device motion, and gameplay performance features on 3-minute samples, with accuracy of 0.95 (sensitivity=0.95, and specificity=0.95). These findings provide strong evidence for the potential use of user-game interaction metrics on existing mobile games as discriminant features for developing an implicit measure to identify alcohol dependence conditions.

To the best of our knowledge, this is the first work to explore the use of gameplay behaviours through gameplay performance, device motion, and touch-based interaction features in an attempt to differentiate between healthy adults and those with alcohol-related problems.

2 RELATED WORK

2.1 Risks Associated with Excessive Alcohol Use and Alcohol Dependence

Apart from social harms, long-term alcohol misuse often poses a substantial risk to several severe physical and mental health conditions, for instance, high blood pressure, stroke, liver disease, neurological complications, depression, including dementia [26, 42]. Previous studies have shown that excessive alcohol consumption can cause cumulative brain damage and cognitive dysfunction through life, even among the young [26, 54]. Uncontrolled shaking of hands, also known as alcohol shake, is common among alcohol dependence. Possible causes of the involuntary hand movement include acute alcohol withdrawal, liver disease, and alcohol-related brain damage. Such hand tremors are most noticeable in a goal-oriented movement toward an object. Alcohol-related brain impairment can also cause problems with coordination, balance, and fine motor movement, including cognitive functioning [11, 33, 52]. It has been found in a growing number of studies that chronic heavy drinkers typically demonstrate deficits in specific cognitive functions, i.e., visuospatial ability, working memory, attention, and executive function [20, 26]. Excessive drinking at younger ages increases the risks of developing long-term cognitive impairment related to alcohol. Unlike other types of cognitive impairment, which typically demonstrate progressive cognitive decline over time, cognitive functions in young chronic alcohol abusers can improve or remain stable with continued alcohol abstinence [26]. Thus, early detection of early-onset alcohol-induced cognitive impairment would greatly facilitate timely intervention [20].

2.2 Limitations of Current Clinical Screening and Diagnostic Instruments

Where sufficient resources exist, alcohol dependence syndrome can be identified through an extensive series of diagnostic measures administered by trained medical staff. The diagnostic procedure often requires diagnostic interviews and other physical or psychological measures to establish a diagnosis. In low-resource settings, such as medical facilities with a shortage of trained clinicians, it is common to use alcohol screening instruments to identify those at risk for alcohol use problems with referral to further diagnostic evaluation for alcohol dependence only if required [24].

The Alcohol Use Disorders Identification Test (AUDIT) is one of the most widely accepted screening tools for alcohol abuse and alcohol dependence in clinical practice. It consists of 10 questions (max score of 40) with regard to alcohol consumption behaviour and social problems caused by chronic drinking and excessive alcohol intake. A score of eight or more is typically used as the cut-off point for identifying individuals with alcohol problems [7]. The AUDIT displayed high sensitivity and specificity values, which were superior to those of other screening instruments [10]. The AUDIT-C, a shortened version of the AUDIT, was developed for use in busy medical settings. It adapts only the first three questions with regard to alcohol consumption of the original AUDIT.

The summed scores of all items can range from 0 to 12. Cut-off scores of 5 or 6 are generally recommended for a positive screen with a high risk of alcohol problems [27].

Despite the compelling evidence of the validity and reliability of the tests, at times, clinicians face challenges to evaluate and interpret the results as responses can be ambiguous or evasive. Indeed, the tests rely heavily on self-reports, which can be deliberately falsified to avoid embarrassment and thus are subjective to potential biases [7]. For instance, the actual volume of alcohol consumption is likely to be under-reported among alcohol drinkers [16]. Such response biases could lead to inaccurate screening results. To accurately quantify actual alcohol intake, asking for clarification and clinical observations are often required to probe for the most accurate responses when conducting the assessment [7]. Most importantly, both AUDIT and AUDIT-C were not originally designed with the explicit purpose of identifying alcohol dependence. In fact, these screening measures are used to identify individuals with hazardous and harmful alcohol drinking patterns before developing dependence. Based on the original document, the AUDIT scores of 20 or above are suggested to affirm the need for further diagnostic evaluation for alcohol dependence [7, 22]. In other words, to indicate the possibility of alcohol dependence, the obtained AUDIT scores need to be much greater than the recommended cut-off widely used in primary care.

2.3 The Use of Technologies in Alcohol Screening

In order to overcome these limitations of conventional alcohol screening tools, there has been an extensive number of studies that explored the use of computer technologies to allow more widespread implementation and to improve alcohol screening rates. Given the ubiquitous computing power of modern mobile devices, electronic screening can be used at home or in a waiting area, allowing patients to perform self-screening prior to a clinical appointment [14, 19]. Automatic feedback from such computer-based measures could provide an understanding of their current conditions and prepare them to discuss further when seeing a therapist. Furthermore, one of the reasons for the inadequate implementation of conventional alcohol screening was the fear of offending patients when asked about their drinking behaviour [5]. A relevant study found positive changes in attitudes among nurses when using the computer-assisted approach, as most of them found it easier to obtain patients' information about their alcohol consumption due to its simplicity [8]. Various studies have also argued that health professionals could take advantage of these technology-based approaches to improve the detection of risky drinking behaviours for early intervention in adolescents who are familiar and highly engaged with technologies [19, 32]. In particular, a recent study demonstrated the feasibility of using of behavioural data derived from built-in mobile sensors along with a self-report survey to infer drinking behaviour of young individuals. An overall accuracy of 76.6% was reportedly achieved by their classifier when using a combination of both self-report and sensor features as input [44].

2.4 The Use of Games in Developing Automated Classification Systems in Clinical Research

Over the past few years, there has been emerging interest in the clinical application of games in evaluation and screening for individuals with underlying health conditions, for instance, people with cognitive impairment [2, 31, 51, 53, 55], and children with autism [6]. Vallejo et al. [53] developed a game-based cognitive measure, which simulates daily living tasks. Several in-game tasks were integrated and used to examine various cognitive functioning. They found significant differences in task performance, especially in complex tasks involving extensive cognitive processes between people with dementia and age-matched healthy controls. In particular, patients demonstrated lower gameplay performance in terms of time and percentage of task successfully completed. Regarding usability, the game-based assessment was found to be well-accepted among users, including the elderly population. Their findings appear to be consistent with significant correlations between gameplay performance and standard cognitive measures reported in other studies [25, 31, 51, 55]. While existing studies tend to mainly focus on gameplay performance, researchers have also examined links between gestural and movement patterns

during gameplay and cognitive abilities. For instance, Intarasirisawat et al. [25] conducted a similar study with mobile game-based assessment, investigating relationships between underlying patterns of user-game interaction and cognitive performance. Notably, they reported that a higher degree of phone movement variability during the gameplay was associated with lower cognitive performance on standardised tests. Furthermore, swipe length and swipe speed were found to be significantly correlated with performance on multiple cognitive functions. Another study that employed the motor patterns in an attempt to identify children with autism was conducted by Anzulewicz et al. [6], in which two mobile games were employed on a smart tablet to collect touch and inertial movement patterns of the device during the gameplay. Their best machine learning model could identify children with autism from typically-developing children with 93% accuracy. With respect to the autism motor signature, they reported significant different gestural and inertial movement patterns in autistic children as compared to their age-matched controls. Specifically, children with autism demonstrated greater force, including faster and more distal gestures with greater variation in the given goal-directed motor tasks [6].

These studies highlighted the advantages of game-based measures over traditional approaches. In that, the game-based modalities provided the ability to control stimuli presentation with precise timing and more accurate measurement. Furthermore, the game-based measures were reportedly well-accepted among participants to perform given tasks with minimal instructions. The entertaining nature of the gameplay appears to enhance engagement and sustain participants' attention to complete the given tasks. As such, the game-based approach shows promise as a novel means of supporting healthcare professionals in clinical decision making and screening activities, especially outside clinical settings.

Previous studies have shown that cognitive and motor deficits resulting in continued disruption of hand movement are commonly found among alcoholics. Hence, in the present work, we examined the use of gesture and device motion patterns along with gameplay performance in developing an automated game-based diagnostic tool for rapid and robust identification of individuals with alcohol dependence within a population of healthy adults.

3 APPROACH

Alcohol dependence is often manifested with specific cognitive and physiological changes such as cognitive decline [20, 26] and changes in the motor functions of their hands [11, 33, 52]. We consider that both of these changes can influence the way users interact with mobile games. In particular, prior work has shown that both game performance metrics and characteristics of touch gestures during gameplay are correlated with cognitive decline [25]. Furthermore, potential alcohol dependence related changes in motor function may be manifested in the movement of the mobile device, and the way touch gestures are performed during gameplay.

These considerations have led us to explore the design of a game-based diagnosis tool that aims to identify people with alcohol dependence. Specifically, the intended system is based around the augmentation of existing smartphone games, such as Tetris, Fruit Ninja, and Unblock Puzzle (see section 5 for details), with the ability to collect data about the user performance, touch gestures on the screen, and device movement recorded through the smartphone's accelerometer and gyroscope. Our objective is to develop a classifier that can identify individuals with alcohol dependence, using these signals as input. To this end, we conducted a controlled study involving 40 individuals diagnosed with alcohol dependence and a control group of 40 healthy individuals. The datasets from multiple gameplay sessions were used to train and validate a classifier for the diagnosis of alcohol dependence.

4 DATA COLLECTION

This study protocol was approved by the Research Ethics Review Group at our research institution. All participants gave written informed consent to participate after a complete description of the study.

The study was carried out at two locations, i.e., a speciality substance use disorder hospital and a university, for the patient and the control (healthy participants) groups, respectively. It is worth noting that participants in the patient group were clinically diagnosed with alcohol dependence. Only those who were admitted to a residential treatment program after being detoxified were recruited to enter the study, given the fact that patients, during the detoxification phase, often experience effects of withdrawal and can be easily tired and unable to focus on tasks at hand [3]. The experiments were conducted individually in a quiet room. All participants completed the AUDIT [4] and the Hospital Anxiety and Depression Scale (HADS) [47] questionnaire prior to commencing the data collection session. Those in both group, who scored a total of 8 points or higher in either Anxiety or Depression categories, were excluded from the study. It is worth noting that the use of the Hospital Anxiety and Depression Scale (HADS) was recommended by a clinical consultant who works closely with our research group. In particular, with the aim to evaluate symptoms of both anxiety and depression, the Patient Health Questionnaire (PHQ-9) was not chosen as it was designed to assess only depression severity [18]. Given that Beck Depression Inventory-II (BD-II) consists of 42 items when combining items from both anxiety and depression subscales, BD-II is a longer test when compared to HADS, which has only 14 items in total. Additionally, BD-II is licensed while HADS can be used for free [41]. HADS was, thus, chosen in our study.

A growing number of studies have shown that chronic binge drinkers typically demonstrate deficits in specific cognitive functions, i.e., visuospatial ability, working memory, attention, and executive function [20, 26]. Participants' cognitive function was therefore examined by a research assistant who has received clinical training in cognitive assessment administration. The cognitive assessment battery included the Mini-Mental State Examination (MMSE) [50], the Montreal Cognitive Assessment (MoCA) [38], and the Stroop Colour-Word Test [23]. These cognitive measures assess a variety of cognitive functions, including orientation, attention, memory, language, executive function, and inhibition control. Participants were given an introduction on how to play all three games pre-installed on a mobile phone (Samsung S6) without a screen protector. To reduce the effect of participants' prior game experience in the study, they were asked to familiarise themselves with the games, each game for 5 minutes. Participants were instructed to hold the phone firmly with one hand and play the games with another hand while sitting on a stationary chair without an armrest. The instruction on handgrip and body posture was to eliminate the effects of hand size and finger length on the touch patterns and to maximise phone movement. Only one finger was allowed on the screen at one time to simplify touch interactions in the analysis. No data was recorded during this game training session. Afterwards, participants were asked to play each of the three games for 10 minutes in random order to avoid order effects. After the completion of each game, touch interaction, device motions, as well as gameplay activities collected during the gameplay, were stored on the phone.

5 MOBILE GAMES

In this study, we used three casual games on the mobile platform, i.e., Tetris, Fruit Ninja, and Unblock Puzzle. Previous studies have shown that long-term excessive alcohol use commonly found in alcohol-dependent drinkers is strongly associated with cumulative brain damage impairing cognitive functions (e.g., attention, visuospatial function, executive function and decision-making) [26]. Apart from being highly engaging and easy to learn, the three games were selected based on their particular game mechanics that place specific cognitive demands, particularly in the domains which are often found impaired in patients with alcohol dependence, i.e. visuospatial function in Tetris [28], executive function in Fruit Ninja [29], and problem-solving in Unblock Puzzle. The games were developed based on game assets in the Unity3D asset store to replicate the look and feel of their commercial version. The games were modified to allow touch interactions, device motions, and gameplay activities to be tracked through built-in sensors on the mobile device. Data were continuously recorded with timestamps at the

Fig. 1. A screenshot of a variation of Fruit Ninja in Unity3D asset store. As multiple fruits are thrown in the air in random directions, players must slice all the fruits by swiping their finger across the screen to cut them in half. The game instantly ends if three pieces of fruit are missed, or players slash bombs that are occasionally thrown onto the screen. The game consists of a single level where the game speed remains stable and the number of fruits and bombs are randomly generated throughout the game.

sampling rate of maximal 30Hz. The games automatically ended after 10 minutes and stored the data in JSON files.

6 PARTICIPANTS

There were two groups of participants, an experimental group of alcohol-dependent patients and a control group of healthy adults.

6.1 Experimental Group

The study sample was drawn from adult male patients admitted to a residential treatment program for alcohol detox and rehabilitation at the hospital. Patients were only considered for inclusion if they: i) were diagnosed with alcohol dependence according to ICD-10 criteria (coded as F10.2) [39]; ii) abstained from alcohol consumption for at least two weeks prior to the study; and iii) were attending the rehabilitation phase of treatment. Potential study participants were contacted and screened for eligibility by the research assistants who were the counsellors of the treatment program.

6.2 Control Group

Age and gender-matched healthy adults were recruited through public advertising for voluntary participation in the study. Volunteers were given £10 for their research participation upon the completion of the data collection session. Potential participants were excluded if they: i) consumed alcohol in the past 24 hours; ii) had symptoms
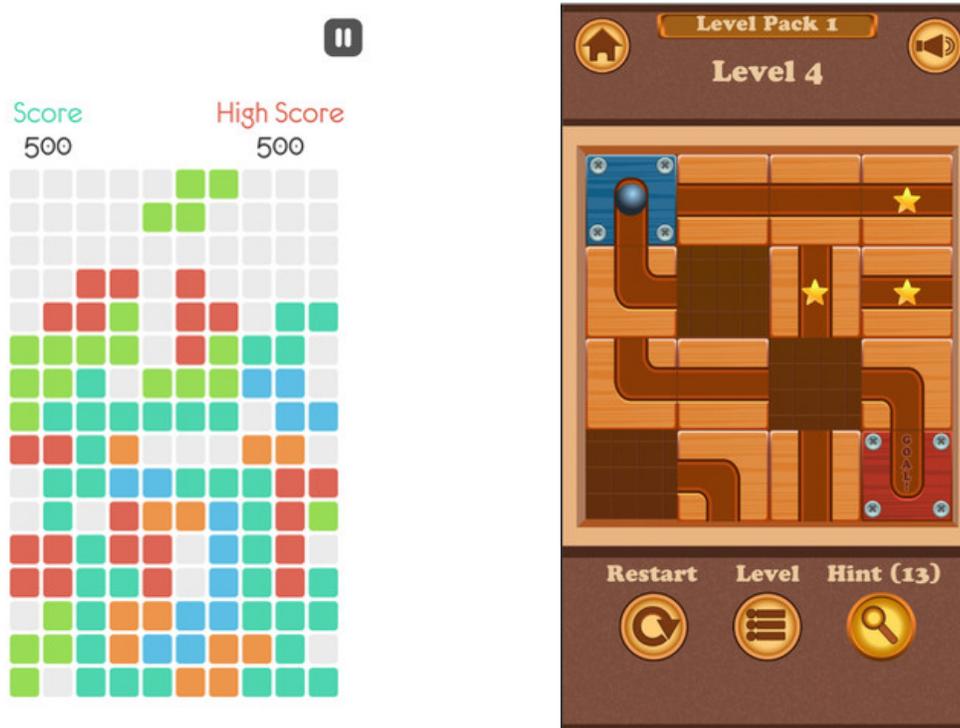
Fig. 2. (Left) A screenshot of a variation of Tetris in the Unity3D asset store. In the game, by moving and rotating the falling shape with finger swipes and taps, players must complete a 10-block horizontal line without gaps. When such a line is created, it disappears, and all blocks above the deleted line will fall. Swiping left/right is to move the shape to the left/right one block unit at a time while swiping down is to force it to fall into the stack instantly. Tapping the screen is to rotate it by 90 degrees. Even though the game consists of only one level where the game speed remains stable along the course of the game, the gameplay becomes more difficult when the falling shapes are stacked up high close to the upper boundary. Therefore, players are required to strategically place falling blocks to complete an entire line horizontally to clear as many lines as possible and get more space to stack the blocks. (Right) A screenshot of a variation of Unblock Puzzle in the Unity3D asset store. Players are required to rearrange the given blocks with different shapes of pipes in order to make the ball roll to the goal block. Once completed, players are moved up to the next level where the level of difficulty gradually increased, e.g. the number of sliding tiles to solve the puzzle. Scores are calculated based on the number of moves to resolve the puzzle in each level.

of hazardous or harmful alcohol use as screened by AUDIT (at the cut-off point of ≥ 8); iii) had a medical history including severe mental illness, drug or alcohol abuse; or iv) were receiving psychoactive medication.

In addition, potential participants in both groups were excluded if they: i) were under 18 years old; ii) had visual conditions affecting daily activities; iii) were diagnosed with Parkinson's disease; iv) had symptoms of generalised anxiety disorder or depression screened by HADS; v) were unable to use smartphones through touch controls; or vi) played video games excessively over 3 hours a day in the past 6 months. Informed consent of 46 patients and 45 healthy participants were acquired, with six patients and five control participants excluded due to exclusion criteria. Thus, the final sample was comprised of 40 alcohol-dependent patients (37 right-handers and 3 left-handers) and 40 healthy participants (37 right-handers and 3 left-handers). All included participants

were male aged between 24 – 65 and had completed primary education (as a minimum). We computed two-tailed independent-samples t-tests for the comparison of alcohol-dependent patients (AD) and healthy controls. The mean age of alcohol-dependent participants (M=42, SD=10.29) was comparable with healthy adults (M=40.75, SD=10.20). There was a significant difference in mean AUDIT scores between patients and controls ($t_{49.649}$=30.135, p < 0.001). Overall, patients with alcohol dependence had significantly lower cognitive performance in all measures as compared to healthy controls (p<0.05). The participants' clinical characteristics are shown in Table 1.

Table 1. Summary of participant characteristics

Variable	AD Patients (n=40)	Healthy Controls (n=40)	P Value (two-tailed)
Gender (female/male)	0/40	0/40	N/A
Handedness (right/left)	37/3	37/3	N/A
Age (years)[means(SD)]	42.08 (10.29)	40.75 (10.20)	0.82
AUDIT [mean(SD)]	35.95 (6.62)	2.28 (2.47)	<0.001
MMSE score [mean(SD)]	26.85 (2.92)	28.10 (1.33)	0.001
MMSE score distribution			
MMSE<=24	7	0	
MMSE 25-30	33	40	
MoCA score [mean(SD)]	21.67 (4.66)	26.33 (2.73)	0.003
MoCA score distribution			
MoCA<=25	32	14	
MoCA 26-30	8	26	
TMTA completion time [mean(SD)]	49.68 (28.71)	25.07 (7.35)	<0.001
Response Inhibition [mean(SD)]	0.29 (1.06)	0.27 (0.22)	0.035

7 DATA PROCESSING AND FEATURE SELECTION

User game interaction patterns were captured for 10 minutes during the gameplay session. Raw data of three different data types, i.e., touch, device motion, and gameplay, were recorded in separated JSON files. In order to avoid inaccurate data analysis, we performed data cleansing to eliminate irrelevant data samples, i.e., the faulty touch data points at the edges of the phone, which were speculated to occur while participants were holding the phone in their palm. The data were then converted into XLS format prior to feature extraction. Features were computed using the entire data samples from each participant. The min-max scaling method was used to transform the features' values to the range of 0 to 1.

7.1 Touch Data

To remove artefacts from the collected samples, touch coordinates were plotted to identify unintentional touches with participants' palm along the edge of the phone. These faulty touches were removed by thresholding the distance and speed between two data points within a given touch interaction. Four measures: count, length, speed and directness index were used to extract features across the four directions of swipes as proposed in [25]. Previous studies have shown that chronic alcohol abuse or withdrawal can cause a tremor or uncontrolled shaking of hands in patients with alcohol dependence [11, 33, 52]. We, therefore, anticipated less consistent touch movement patterns, precisely, a higher degree in movement variability in patients. As a result, the mean value and the standard deviation of features were computed except for the count features. Research findings also suggested that different game mechanics and cognitive demands associated with the given game tasks made some

swipe characteristics patterns more relevant than others[25]. Hence, different subsets of touch-based features were chosen for each game, as listed in Table 2.

Table 2. Extracted touch features for each game

Feature	Mean and SD	Tetris	Fruit Ninja	Unblock Puzzle
Total number of swipes by direction[a]		✓		✓
Total number of swipes in horizontal[b] or vertical[c]		✓		✓
Total number of overall swipes		✓		✓
Total number of taps		✓		
Swipe length by direction[a]	✓	✓		✓
Swipe length in horizontal[b] or vertical[c]	✓	✓		✓
Overall swipe length	✓	✓	✓	✓
Swipe speed by direction[a]	✓	✓		✓
Overall swipe speed	✓	✓	✓	✓
Starting point on the x-axis of swipes by direction[a]	✓	✓		✓
Starting point on the y-axis of swipes by direction[a]	✓	✓		✓
Starting point on the x-axis of swipes in horizontal[b] or vertical[c]	✓	✓		✓
Starting point on the y-axis of swipes in horizontal[b] or vertical[c]	✓	✓		✓
Starting point on the x-axis of overall swipes	✓	✓	✓	✓
Starting point on the y-axis of overall swipes	✓	✓	✓	✓
Centre of mass on the x-axis of overall swipes	✓		✓	
Centre of mass on the y-axis of overall swipes	✓		✓	
Directness index[d] of swipes by direction[a]	✓	✓		✓
Directness index[d] of swipes in horizontal[b] or vertical[c]	✓	✓		✓
Directness index[d] of overall swipes	✓	✓	✓	✓

[a]Right, left and down in Tetris and all four directions in Unblock Puzzle
[b]Denote swipes in both left and right directions
[c]Denote swipes in both up and down directions
[d]A feature to quantify the swipe straightness. Its value is computed by the total distance of an interaction divided by the displacement. As such, the value of 1 indicates a perfectly straight line, while a curved swipe has a greater value.

7.2 Sensor Data

We hypothesised that patients with alcohol dependence would exhibit different device motion patterns from healthy adults during the gameplay session. Therefore, device acceleration and rotational motion were captured using the built-in 3-dimensional motion sensors. Based on the entire samples within the allotted game time, we computed a total of 16 features using the mean and the standard deviation values in each axis, including the magnitude of the 3D vector for acceleration and rotational motion.

7.3 Gameplay Data

Previous studies have shown that gameplay performance was associated with standard cognitive measures [30, 51]. Most alcoholics exhibit cognitive impairment in multiple domains, especially, visuospatial function, executive functions, and memory [33, 40]. Thus, gameplay metrics were included as features. The game interaction and activities are shaped by a set of rules and game mechanics defined explicitly in the game. Therefore, we extracted different sets of gameplay features for each game, as listed in Table 3.

Table 3. Extracted gameplay features for each game.

Feature	Description
Tetris	
Max score	The maximum score achieved in the allotted game time
Mean first response x pos	The average position on the x-axis of the first response on the screen of each falling shape
Mean first response y pos	The average position on the y-axis of the first response on the screen of each falling shape
Mean stack height	The average stack height when touch interaction performed on the screen
Mean delta first response y pos stack height	The mean difference between the position on the y-axis of the first response and the stack height
Mean response time	The average response time of all touch interaction on each falling shape
Mean time to first response	The average response time of the first response on the screen of each falling shape
Mean no of taps per shape	The average number of taps performed on each falling shape
Mean no of horizontal swipes per shape	The average number of horizontal swipes performed on each falling shape
Mean no of down swipes per shape	The average number of downward swipes performed on each falling shape
Mean no of swipes per shape	The average number of overall swipes performed on each falling shape
Fruit Ninja	
Max score	The maximum score achieved in the allotted game time
No of times to startover	The number of times the game was started over
Mean overall air time	The average time a ball was in the air including the missing balls
Mean air time before being cut	The average time a ball was in the air before being cut
Mean cut position x	The average cut position in x axis
Mean cut position y	The average cut position in y axis
Unblock Puzzle	
Max level	The maximum level achieved in the given duration
Mean response time	The average response time of all touch interactions in the given duration
Mean response time to complete level	The average response time of touch interactions to complete a level in the given duration
Mean total time to complete level	The average total time to complete a level in the given duration
Mean complete level no of moves	The average number of moves performed to complete a level in the given duration

8 MODEL VALIDATION

In this work, we employed three supervised machine-learning methods to classify patients with alcohol dependence and healthy controls: Logistic Regression (LR), Linear Support Vector Machine (LSVM), and Random Forest (RF). "Alcohol Dependence' and 'Healthy Controls' were used as our class labels. Samples from participants, who were clinically diagnosed with alcohol dependence and were admitted to a residential treatment program, were labelled as 'alcohol dependence'. Samples from healthy participants were labelled as 'healthy control'. Only in the control group of healthy participants, the AUDIT with the cut-off point of ≥ 8 was used to screen symptoms of hazardous or harmful alcohol use. As a result, participants in the control group with an AUDIT score ≥ 8 were excluded from the study. As the observed class labels were balanced (40 alcohol dependence and 40 healthy controls), classification accuracy was used to evaluate the models' performance. K-fold cross-validation technique was applied when building models to generalise our results to unseen data. In particular, stratified K-fold cross-validation (K = 10) was used to generate test sets with a balanced distribution of classes. In each iteration, the models' performance was assessed on a random selection of 4 patients and 4 healthy adults. Feature selection was performed within the cross-validation on the training sets using the Correlation-based Feature Selection algorithm (CFS) [17] to remove collinear features. In each K-fold iteration, grid-search with 10-fold cross-validation was used for hyper-parameter tuning to maximise the model classification accuracy. The models' performance on the completely unseen test sets were evaluated using the corresponding selected features. It is noteworthy that owing to a procedural error during the data collection, the recording was incomplete in one of the participants and therefore was excluded, leaving a final data set of 79 samples to analyse with Unblock Puzzle.

Quantitative comparisons were made on classifier performance over five different feature sets to investigate the effect on accuracy. Particularly, the feature sets included Touch (T), Sensor (S), Gameplay (G), Combined
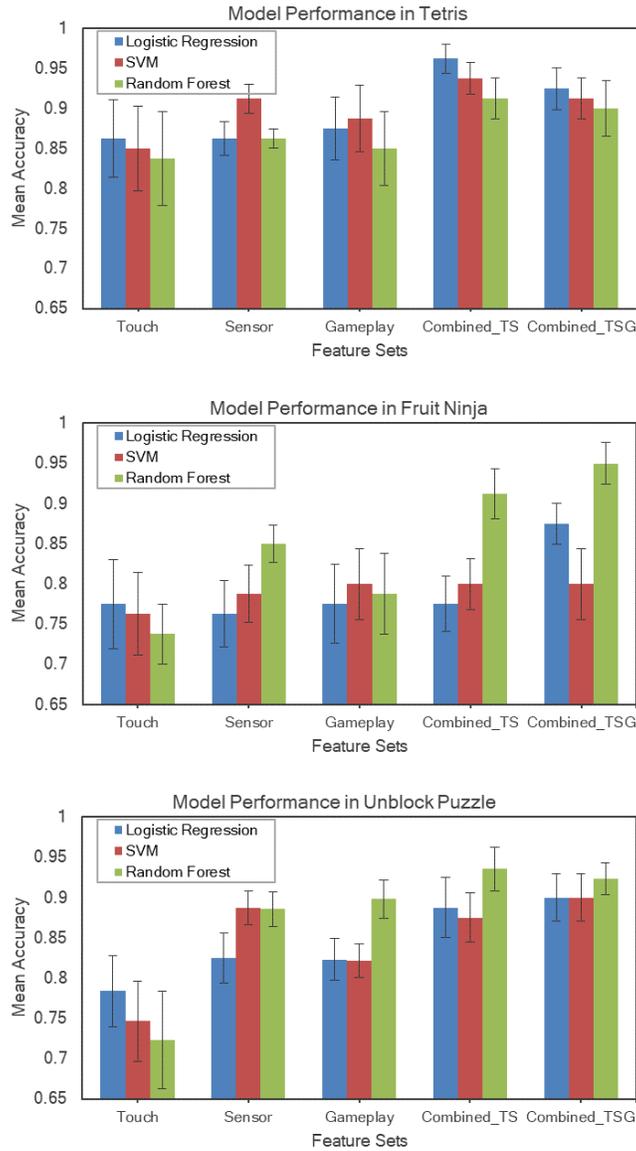
Fig. 3. Illustrates the 10-fold cross-validation performance of each classifier over five different data types using data taken from the three games

Touch and Sensor (TS), and Combined Touch, Sensor and Gameplay (TSG). Mean accuracy and standard error of the mean were computed for each classifier, as shown in Fig. 3.

By comparing the overall results, classifiers built on Tetris features performed better than those of Fruit Ninja and Unblock Puzzle. Using combinations of multiple feature sets (touch, sensor, and gameplay) as input improved

the models' performance in classification accuracy as compared with using a single set of features. Logistic Regression yielded the highest mean accuracy of 0.96 when using the ensemble of touch and sensor features from Tetris. However, with the same feature sets, Random Forest outperformed other classification algorithms in general, where the mean accuracies were well above 0.90 in the other two games. Moreover, the low standard errors of the mean accuracy (0.02 – 0.03) found in these classifiers indicate that the models generalised well against unseen data.

Our main aim in developing these classification models was to investigate their effectiveness as a quick and accurate diagnostic instrument for the identification of alcohol-dependent patients. In current clinical practice, diagnostic interviews are used to diagnose patients with alcohol dependence according to ICD-10 criteria. Despite being regarded as a gold standard, these clinical interviews require trained clinicians to administer and often involve a series of additional measures to establish a reliable diagnosis, e.g., physiological tests and cognitive assessments [24]. In order to support diagnostic team members, this novel game-based diagnostic tool was aimed to provide accurate results within a short time. Therefore, we examined the classifier performance on different gameplay durations, i.e., 3, 4, 5, 6, 7, 8, 9, and 10 minutes. For this reason, the first part of the original 10-minute data was truncated for the given durations accordingly. Based on the findings above, Logistic Regression and Random Forest achieved the best accuracies. Therefore, these classification models were selected and trained on the combinations of multiple feature sets (TS and TSG).

The average classification accuracies of Logistic Regression and Random Forest were evaluated against the combined TS and the combine TSG feature sets, as shown in Table 4 and Table 5, respectively. When using the combined TSG, both classifiers yield better overall performance for the given durations in Fruit Ninja and Unblock Puzzle. In contrast for Tetris, the mean accuracy for TSG was slightly lower than TS for most durations. With regard to classification methods, Random Forest classifiers generally outperformed Logistic Regression models in most cases, especially on input with short gameplay durations of 3 – 6 minutes. In particular, even with the gameplay durations as short as 3 minutes, the random forest was able to give relatively high classification accuracies of 0.94, 0.89, and 0.95 in Tetris, Fruit Ninja, and Unblock Puzzle respectively.

Table 4. Mean accuracies of Logistic Regression Classifiers

Duration	Tetris		Fruit Ninja		Unblock Puzzle	
	TS	TSG	TS	TSG	TS	TSG
10 Minutes	**0.96 (0.02)**	0.93 (0.03)	0.78 (0.03)	**0.88 (0.03)**	0.89 (0.04)	**0.90 (0.03)**
9 Minutes	**0.95 (0.03)**	0.90 (0.03)	0.80 (0.04)	**0.84 (0.03)**	0.88 (0.03)	**0.90 (0.03)**
8 Minutes	**0.93 (0.03)**	0.91 (0.03)	0.83 (0.04)	**0.85 (0.04)**	0.87 (0.04)	**0.90 (0.03)**
7 Minutes	**0.93 (0.03)**	0.91 (0.03)	0.76 (0.04)	**0.84 (0.04)**	0.89 (0.03)	**0.90 (0.03)**
6 Minutes	**0.91 (0.03)**	0.90 (0.03)	0.75 (0.04)	**0.83 (0.04)**	0.89 (0.03)	**0.93 (0.03)**
5 Minutes	0.89 (0.04)	**0.91 (0.03)**	0.80 (0.05)	**0.85 (0.03)**	0.89 (0.03)	0.89 (0.03)
4 Minutes	0.89 (0.03)	**0.90 (0.03)**	0.79 (0.05)	0.79 (0.04)	**0.90 (0.02)**	0.87 (0.02)
3 Minutes	**0.89 (0.03)**	0.88 (0.03)	0.78 (0.04)	**0.81 (0.05)**	0.89 (0.03)	**0.90 (0.03)**

[a](1 standard error either side of the mean)
[b]Table numbers in boldface highlight a better feature set in classification performance between combined TS and combined TSG

It is also worth noting that the best performing model on 10-minute samples was the Logistic Regression on combined TS features of Tetris with 0.96 accuracy. However, the model performance continued to decline when

Table 5. Mean accuracies of Random Forest Classifiers

Duration	Tetris		Fruit Ninja		Unblock Puzzle	
	TS	TSG	TS	TSG	TS	TSG
10 Minutes	**0.91 (0.03)**	0.90 (0.03)	0.91 (0.03)	**0.95 (0.03)**	**0.94 (0.03)**	0.92 (0.02)
9 Minutes	0.91 (0.03)	**0.93 (0.03)**	0.84 (0.02)	**0.90 (0.02)**	0.94 (0.03)	**0.96 (0.02)**
8 Minutes	0.91 (0.03)	**0.93 (0.03)**	0.86 (0.02)	**0.93 (0.02)**	0.92 (0.03)	**0.94 (0.02)**
7 Minutes	0.91 (0.03)	0.91 (0.03)	0.88 (0.03)	**0.91 (0.02)**	0.91 (0.03)	**0.95 (0.02)**
6 Minutes	**0.94 (0.02)**	0.91 (0.03)	0.85 (0.03)	**0.86 (0.03)**	0.94 (0.02)	**0.95 (0.02)**
5 Minutes	**0.93 (0.03)**	0.90 (0.04)	0.86 (0.04)	**0.90 (0.03)**	0.92 (0.02)	**0.95 (0.02)**
4 Minutes	0.91 (0.03)	**0.93 (0.03)**	0.85 (0.04)	**0.91 (0.03)**	0.89 (0.04)	**0.93 (0.03)**
3 Minutes	**0.94 (0.03)**	0.93 (0.03)	0.86 (0.04)	**0.89 (0.03)**	0.90 (0.04)	**0.95 (0.02)**

[a](1 standard error either side of the mean)
[b]Table numbers in boldface highlight a better feature set in classification performance between combined TS and combined TSG

reducing the gameplay duration in which the performance dropped to 0.89 when training on 3-minute samples. By contrast, Random Forest classifiers on Unblock Puzzle data using all features performed consistently well above 0.92 across different gameplay durations. The mean accuracies only changed slightly within the range of 0.92 – 0.95 over the given durations. In general, Fruit Ninja seemed to yield the worst classification performance among the three games.

We observe that there is a somewhat mixed picture of classification performance among the three games when reducing the allotted game time. Specifically, irrespective of the algorithm used for the classifier, model performance in Fruit Ninja and Unblock Puzzle seemed to remain relatively stable for different durations (with a slight drop when using shorter intervals of data as input). In Tetris, however, the two classifiers seem to be affected differently with changes in duration. Logistic Regression demonstrated a downward trend in performance similar to the other two games, while the Random Forest classifiers seemed to perform slightly better when using shorter intervals of Tetris data as input. Although the differences are very small to consider them significant, we speculate that such patterns could be explained by the differences in gameplay mechanics and the inherent nature of the two machine learning algorithms. Nevertheless, we do not have sufficient evidence to draw a conclusion.

We move on to examine the sensitivity and specificity of our models, as they are more common measures for assessing the performance of clinical tests [35]. Sensitivity refers to the ability of the clinical test to correctly identify those patients with the condition while specificity tells us how well the tests can correctly identify those without the condition. Therefore, we evaluated and compared the sensitivity and specificity of Logistic Regression and Random Forest classifiers on 3-minute data and 10-minute data, as presented in Table 6 and Table 7.

Overall, all classifiers showed promising classification performance in terms of sensitivity (0.80 – 0.98) and specificity (0.75 – 0.98). In particular, using 10-minute samples, the Logistic Regression classifier on combined TS features of Tetris yielded the highest sensitivity of 0.98 with a specificity of 0.95. On the contrary, the highest specificity of 0.98 was obtained from the classifiers using the ensemble of all features of Fruit Ninja and Unblock Puzzle but with lower sensitivities of 0.93 and 0.83, respectively. These results suggest that the Logistic Regression classifier on 10-minute combined TS features of Tetris has the best capability in screening patients. On the other hand, by using 10-minute samples and all feature sets, the Logistic Regression model on Unlock Puzzle and the Random Forest model on Fruit Ninja seem to be better options for diagnosis purposes in which a high specificity value is preferred.

Table 6. Sensitivity and specificity of Logistic Regression Classifiers

Game	3-Minute Samples				10-Minute Samples			
	Combined TS		Combined TSG		Combined TS		Combined TSG	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
Tetris	0.85 (0.04)	0.93 (0.04)	0.83 (0.06)	0.93 (0.04)	0.98 (0.02)	0.95 (0.03)	0.90 (0.05)	0.95 (0.03)
Fruit Ninja	0.80 (0.05)	0.75 (0.09)	0.80 (0.05)	0.83 (0.08)	0.80 (0.06)	0.75 (0.07)	0.90 (0.04)	0.85 (0.05)
Unblock Puzzle	0.83 (0.06)	0.95 (0.03)	0.88 (0.06)	0.93 (0.04)	0.83 (0.05)	0.95 (0.05)	0.83 (0.06)	0.98 (0.02)

[a](1 standard error either side of the mean)

Table 7. Sensitivity and specificity of Random Forest Classifiers

Game	3-Minute Samples				10-Minute Samples			
	Combined TS		Combined TSG		Combined TS		Combined TSG	
	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity
Tetris	0.98 (0.02)	0.90 (0.05)	0.93 (0.04)	0.93 (0.04)	0.95 (0.03)	0.88 (0.05)	0.90 (0.05)	0.90 (0.04)
Fruit Ninja	0.85 (0.05)	0.88 (0.07)	0.85 (0.05)	0.93 (0.05)	0.88 (0.06)	0.95 (0.03)	0.93 (0.05)	0.98 (0.02)
Unblock Puzzle	0.93 (0.04)	0.88 (0.07)	0.95 (0.03)	0.95 (0.03)	0.93 (0.05)	0.95 (0.03)	0.93 (0.04)	0.93 (0.04)

[a](1 standard error either side of the mean)

Considering our aim to reduce the administration time, the classifiers on 3-minute samples still showed promising results, particularly those based on the Random Forest algorithm. Using combined TSG features of Unblock Puzzle, the model achieved a sensitivity of 0.95 and specificity of 0.95. Interestingly, the aforementioned algorithm on the combined TS features of Tetris had the highest sensitivity of 0.98 but a lower specificity value of 0.90. These results demonstrate that even for the relatively short duration of 3 minutes, our game-based classifiers yielded consistently high sensitivities and specificities with low standard errors of the mean. The 10-fold cross-validation performance of the models indicates that our game-based approach provides the ability to effectively diagnose alcohol-dependent patients with overall good generalisation capabilities and short administration time.

9 FEATURE IMPORTANCE

Further investigations into appropriate game features and their importance can provide useful insights to improve classification performance, model complexity, and speed. To this end, we investigate which user-game interaction features are most important for building the automated game-based measure for screening patients with alcohol dependence. The analysis was conducted on the Random Forest classifier on the ensemble of all feature sets of Unblock Puzzle as the model shows the best classification performance among the compared classifiers on 3-minute samples.

Tree-based models such as the Random Forests use node impurity or information gain to split nodes and construct an ensemble of decision trees. For each tree, feature importance is determined by the amount of information gained from the feature weighted by the number of observations in the node. The model provides feature importance scores by averaging values of feature importance across all decision trees. The higher the importance score, the more informative the feature is. In this work, we ran 10-fold cross-validation on Random Forests and ranked the feature importance scores using the *feature_importances_* attribute in the scikit-learn
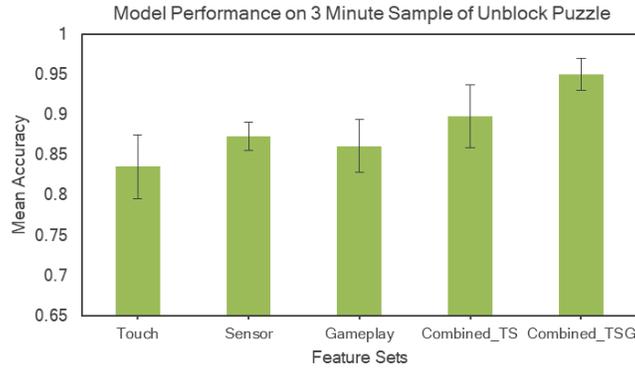
Fig. 4. Illustrates the 10-fold cross-validation performance of Random Forest Classifier over five different data types using 3-minute data taken from Unblock Puzzle.

machine learning library. Feature selection was performed in the data processing step to maximise the classification accuracy by removing redundant features and thus improve the interpretability of the classification model. In that, the Correlation-based Feature Selection (CFS) technique was performed within the cross-validation to select a subset of relevant features out of all original features.

We extracted feature importance values at each cross-validation iteration and determined the overall ten top-ranked features with the highest median variable importance. We statistically tested whether there was a significant difference between patients and control groups. Due to the small sample size, the Shapiro-Wilk test was used to determine the normality of the selected features. In a handful of the features (8 out of 10), the results indicated that the data significantly deviated from normal distributions ($p<0.05$). Therefore, the statistical results were derived from the non-parametric Mann Whitney U Test, as shown in Table 8.

Table 8. Top 10 features with the highest median feature importance

Feature	Type	Feature Importance	Median Feature Value		U Test Statistic	P Value
			Controls	Patients	($n_{control}$=40,$n_{patients}$=40)	(two-tailed)
Mean Acceleration Magnitude	Sensor	0.228	0.996	1.028	202	<.001
Mean Response Time to Complete Level	Gameplay	0.174	1963.84	5787.93	60	<.001
Mean Response Time	Gameplay	0.126	1598.60	4377.47	142	<.001
Max Level	Gameplay	0.097	9	4	140	<.001
Mean Speed Swipe Up	Touch	0.048	2945.74	1627.54	196	<.001
Mean Speed Swipe Right	Touch	0.038	2623.65	1642.09	232	<.001
Mean Total Time to Complete Level	Gameplay	0.027	12468.67	25135.9	276	<.001
Mean Length Swipe Left	Touch	0.020	553.63	428.32	331	<.001
Mean Speed Swipe Left	Touch	0.019	2541.73	1501.30	273	<.001
Standard Deviation Acceleration Magnitude	Sensor	0.013	0.040	0.041	762	0.86

[a]Bonferroni critical value $p<0.005$ (0.05/10)

Results of the statistical analysis revealed significant differences ($p<0.001$) in all these features between the two groups except for the standard deviation acceleration magnitude. The mean acceleration magnitude by far is the strongest predictive variable with the highest median feature importance value. As such, patients tend to exhibit a statistically larger degree of device movement (1.028) comparing to healthy controls (0.996).

The finding is consistent with previous research, where hand tremors are commonly found in patients with alcohol dependence [11, 33, 52]. It is therefore not unexpected to find that the device motion feature could be an important marker to discriminate patients from healthy controls. With regard to the gameplay performance, patients demonstrated lower performance comparing to healthy adults in terms of the mean response time to complete level (patients=5787.93, controls=1963.84), the overall response time (patients=4377.47, controls=1598.60), the max level completed (patients=4, controls=9), including the mean total response time (patients=25135.9, controls=12468.67). These seem to be coherent with the findings in several studies [11, 33, 52]. In that, long-term excessive alcohol consumption can cause problems with cognitive functioning, in this context, adversely affecting gameplay performance in such high cognitive demand tasks. Furthermore, significant differences in touch patterns between both groups were observed. Specifically, patients' swipe patterns are significantly shorter and slower as compared to healthy controls.

In addition, we compared the classification performance over different feature sets to investigate whether the use of a combination of all feature sets can significantly improve classification accuracy comparing to the use of device motion-based features alone. As shown in Fig 4, when using only a single set of features as input, the classification performance is 0.84, 0.87, 0.86 for touch, sensor, and gameplay, respectively. By using the combination of all sources of features, the classification accuracy significantly improved to 0.95. These results appear consistent with the findings we found earlier in 10-min samples, where the model performance was maximised by using an ensemble of multiple feature sets as input.

10 PARTICIPANTS' FEEDBACK

At the end of the experiment session, a brief interview was carried out to obtain participants' views regarding the acceptability and practical application of the game-based approach. Most participants (65 out of 80) reported enjoying themselves and being engaged with the games. Moreover, the game-based method was reported to reduce their anxiety as compared to the paper-based version. Nevertheless, the majority of them (69 out of 80) found their game experience in Tetris quite frustrating due to difficulty in controlling the Tetris block movement with touch. Unlike the traditional controls, on a touchscreen, players have to drag their finger left or right to move the falling shape. As such, many participants expected that the touched piece would move proportionally based on the distance the finger has moved on the screen. The fact that the block only moves one block unit at a time might be confusing, which may lead to a less engaging game experience. More training may have been beneficial to allow participants to familiarise themselves with the game controls. On the other hand, several participants highly enjoyed the Unblock Puzzle as the in-game tasks challenge their problem-solving skills and creativity. Despite requiring improvement in certain areas, most participants nonetheless felt comfortable and relaxed during the gameplay. These overall positive feedbacks suggest that the game-based screening platform is well-accepted with regard to usability.

11 DISCUSSION

This work presents a novel approach based on machine learning algorithms to provide an alternative method for home-based screening for alcohol dependence. In this study, we examined the use of user-game interaction features from the three popular mobile games to classify patients and healthy adults. The ensemble of features extracted from 10-minute samples through touch interaction, device motion, and gameplay performance was used as input for the proposed classification models. A set of evaluation metrics, specifically the accuracy, sensitivity, and specificity, were used to quantify the classifier performance on entirely independent data sets using a stratified 10-fold cross-validation scheme. Our Logistic Regression classifier based on Tetris samples was capable of identifying the alcohol dependence condition with the accuracy of 0.96. The relatively high sensitivity of 0.98 indicates that our game-based model could be used as an effective screening tool for such a condition. A practical

implementation of such a screening instrument should not only provide accurate and reliable outcomes but should also be inexpensive and time-efficient. Our classifiers show the promise to provide such an instrument as they perform consistently well even with sample lengths of 3 minutes in terms of accuracy, sensitivity, and specificity.

Generally, in medical screening and diagnosis, there is always a trade-off between specificity and sensitivity, and the decision must be made based on their relative clinical importance. In the screening process, it is important to identify as many patients with the condition as possible. Missing cases could lead to delayed treatment or no treatment at all. As a result, for early detection of health conditions, a screening tool with a high sensitivity value is preferred. In contrast, high specificity allows healthcare professionals to regard individuals as having a condition confidently. Thus, specificity is a useful attribute in a diagnostic procedure to avoid provoking anxiety, including unnecessary follow-up tests and treatment [35]. Despite a slightly lower specificity of 0.90, results show that TS features of Tetris samples with the Random forest classifier has the best screening properties (sensitivity of 0.98) when compared to other classifiers. Nevertheless, considering the trade-off between false positive and false negatives, we recommend the Random Forest model using the combined TSG features of Unblock Puzzle as it has the best combination of specificity (0.95) and sensitivity (0.95) for a reliable screening instrument.

11.1 Comparison to Conventional Screening Tools

To fully evaluate the performance and practicality of our proposed game-based method, it is essential to compare the results with current clinical screening tests. In 2009, a systematic review of alcohol use disorder screening instruments evaluated an exhaustive number of studies on the validation of AUDIT and AUDIT-C as screening instruments for alcohol-related problems [10]. Derived from the review, Table 9 reveals the screening and diagnostic performance of AUDIT and AUDIT-C for identifying alcohol dependence found in previous studies [15, 43, 46]. The findings show that both tests were sensitive to alcohol dependence with sensitivity above 0.82 at relative cut-off points. Of note, at a cut-off point of 5, the AUDIT-C yielded a relatively high sensitivity of 0.94, which appears to be superior to that of the full-scale version. Nonetheless, the specificity was only 0.58 [46]. In other words, 42% of non-alcohol dependent individuals were misclassified. Slightly lower sensitivity values were found in other studies, i.e., 0.88 [15, 43] and 0.82 [9]. However, the results yielded significant improvement in specificity (>= 0.77). Among these studies, Rumpf et al. [43] reported a sensitivity of 0.88 and a specificity of 0.81 using an optimal cut-off point of four, which is much lower than the standard cut-off point of >=8 for the AUDIT [4]. Nevertheless, these studies seem to present varying empirical results with very low specificities. One possible explanation could be that these conventional screening tools typically involve questionnaire techniques which heavily rely on respondents' feedback. It was found that several patients with alcohol-related problems are unable to quantify the actual amount of their alcohol consumption [16]. Their responses may also be confounded with dishonest answers to avoid embarrassment. Thus, these self-reported response biases may lead to inaccurate and unreliable results [36].

Table 9. Sensitivity and specificity of paper-based screening tools for alcohol dependence

Study	Test	Cutoff	Sensitivity	Specificity
Rumpf et al. [43]	AUDIT	4	0.88	0.81
Giang et al. [15]	AUDIT	7 or 8	0.88	0.77
Dawson et al. [9]	AUDIT-C	6	0.82	0.80
Seale et al. [46]	AUDIT-C	5	0.94	0.58

Comparing to the results obtained from our game-based classification models, based on 3-minute samples from the Unblock Puzzle game, the Random Forest classifier was found to have identified the alcohol dependence condition at a greater sensitivity (0.95) and specificity (0.95) than the conventional paper-based instruments. Besides the excellent sensitivity for such a condition, the screening process using our proposed game-based approach is as brief as 3 minutes, which is comparable or even faster than the paper-based modalities.

11.2 Comparison to EEG-based Screening Tools

Other than the traditional paper-based tools, a number of literature reviews reported the clinical relevance between various electroencephalogram (EEG) features and the alcohol user disorder, for instance, inter-hemispheric coherence and phase delay [21]. Another review, published in 2017, identified several studies that evaluated the clinical application of EEG-based methods to classify patients with alcohol problems and healthy adults [36]. Various machine learning techniques were employed on EEG signals recorded from multiple scalp electrodes in response to stimuli or cognitive events. The EEG-based classifiers showed a great capability in detecting AUD patients. Overall, the machine learning techniques utilised in these studies achieved relatively high accuracies (>0.90) [36]. Similarly, the classification models of Mumtaz et al. [37] employed the EEG features computed from each of 19 electrode signals. These EEG data were captured in participants' resting state with 5-min of eye-closed and 5 min of eyes-open. By integrating all EEG features, their classifiers yielded similar results with an accuracy of 0.89. Even though the EEG modalities show very promising results, these technology-based screening methods have a number of barriers to applying in field settings. For instance, artefacts caused by eye movement and muscle activities need to be removed prior to data analysis [37]. Most importantly, the EEG data acquisition often takes a long time for equipment set-up and calibration to ensure precise placement of multiple electrodes on the scalp. For these reasons, it can be administered only by trained clinical staff within clinical premises. These limitations make the EEG approach less effective and not popular in clinical practice as compared to the conventional paper-based tools [36].

On the contrary, our proposed method exploits the sensing capabilities of mobile devices to passively collect the user-game interaction behaviours from the 3-minute gameplay. With the entertaining features of the games, we anticipated that individuals would be less likely to have test anxiety as reported in [30] and, therefore, this game-based measure could potentially promote self-screening for early detection of potential alcohol dependence. As well as providing accurate outcomes, the game-based approach is relatively fast and inexpensive (as it can be installed on almost any smartphone). In addition, the game-based screening instrument is robust, portable and requires only minimal effort to implement. As such, our automated system could be adapted by health professionals as a home-based screening tool to help identify individuals who are at risk of having a subtle sign of the alcohol dependence condition.

11.3 Limitations and Future Work

Despite the promising results, the findings of this study are subject to two major limitations. First, since almost all patients admitted to the speciality facility where the experiment was carried out are male. Hence we restricted our sample collection to males exclusively. Second, in order to remove the possible confounding factors, individuals with excessive gaming experience or symptoms of anxiety or depression were excluded from our study. As our proposed technique focuses on the use of touch interaction and device motion, a tremor symptom found in other cognitive disorders could have an adverse impact on our technique and potentially confound our results. Therefore, we also excluded individuals with health conditions often exhibiting hand tremors (e.g. Parkinson's disease) from our study. In addition, impaired visuo-motor coordination and hand-eye coordination could adversely affect participants' gameplay performance and user-game behavioural patterns. Therefore, visually impaired individuals were excluded. For these reasons, our results should be interpreted with caution. Due to this limited

scope, our results cannot be generalised to the entire relevant subgroup of our target population. These limitations could be overcome by further research to investigate the potential effects of these factors on model classification performance and validate the findings we drew from this study.

It is worth noting that the variations in levels of game difficulty over the course of the games could potentially have impacts on classification performance in these games. For example, in Tetris, due to there being a shorter amount of time to place blocks, there are periods of greater and lesser difficulty, partly depending on early choices and partly on chance, within the game. It is speculated that players with better visuospatial function can anticipate and mitigate against these periods of greater difficulty and hence its choice as a tool to distinguish between those with impairment of that function and those without. In Unblock Puzzle, given that the game places demands on cognitive function related to problem-solving, we speculate that as the level of difficulty gradually increases when progressing through levels, patients with impaired problem-solving abilities would take longer to solve the puzzle, leading to completing fewer levels when compared to healthy individuals. This is indeed in line with the results in Table 8, where the maximum level achieved by alcohol-dependent patients was found to be significantly lower than that of healthy participants. However, in the present work, we do not have sufficient data to perform an in-depth analysis. This issue might be addressed in future research.

Future research could examine the feasibility of using such gameplay behaviour patterns to determine the risk of relapse in detoxified alcohol-dependent patients. The practical implementation in a large-scale setting may require additional features such as an automatic notification system to notify the results to clinicians so that timely intervention and appropriate treatment strategies can be given to prevent further development of more severe conditions.

12 CONCLUSION

We believe that this is the first work that evaluates the application of user-game interaction patterns in game-based classifiers for screening and diagnosis of an alcohol dependence condition. We present our novel game-based approach as a fast and inexpensive home-based self-screening tool. In comparison with the paper-based and EEG-based methods, our models demonstrated equal or superior classification performance in terms of accuracy, sensitivity, and specificity. We envision that our home-based self-screening instrument could be used to establish a practical and cost-effective screening strategy to increase screening rates for early detection of alcohol dependence in primary care, hospitals or rehab centres.

ACKNOWLEDGMENTS

REFERENCES

[1] 2018. Alcohol. url: https://www.who.int/news-room/fact-sheets/detail/alcohol;. https://www.who.int/news-room/fact-sheets/detail/alcohol; Accessed: 4-May-2020.
[2] 2019. *Predicting Cognitive Impairment Level after a Serious Game-based Therapy in Chronic Stroke Survivors.*
[3] 2020. Alcohol Rehab. url: https://www.drugrehab.com/addiction/alcohol/rehab/;. https://www.drugrehab.com/addiction/alcohol/rehab/; Accessed: 4-May-2020.
[4] John P. Allen, Raye Z. Litten, Joanne B. Fertig, and Thomas Babor. 1997. A Review of Research on the Alcohol Use Disorders Identification Test (AUDIT). *Alcoholism: Clinical and Experimental Research* 21, 4 (June 1997), 613–619. https://doi.org/10.1111/j.1530-0277.1997.tb03811.x
[5] S. Anderson, D.R. Eadie, A.M. MacKintosh, and S. Haw. 2001. Management of alcohol misuse in Scotland: the role of A&E nurses. *Accident and Emergency Nursing* 9, 2 (April 2001), 92–100. https://doi.org/10.1054/aaen.2000.0210
[6] Anna Anzulewicz, Krzysztof Sobota, and Jonathan T. Delafield-Butt. 2016. Toward the Autism Motor Signature: Gesture patterns during smart tablet gameplay identify children with autism. *Scientific reports* 6 (2016), 31107.

[7] Thomas F Babor, Juan Ramon de la Fuente, John Saunders, and Marcus Grant. 2001. The Alcohol Use Disorders Identification Test: Guidelines for use in. *Primary care* (2001).

[8] Preben Bendtsen, Marika Holmqvist, and Kjell Johansson. 2007. Implementation of computerized alcohol screening and advice in an emergency department – a nursing staff perspective. *Accident and Emergency Nursing* 15, 1 (Jan. 2007), 3–9. https://doi.org/10.1016/j.aaen.2006.09.004

[9] Deborah A. Dawson, Bridget F. Grant, Frederick S. Stinson, and Yuan Zhou. 2005. Effectiveness of the Derived Alcohol Use Disorders Identification Test (AUDIT-C) in Screening for Alcohol Use Disorders and Risk Drinking in the US General Population. *Alcoholism: Clinical & Experimental Research* 29, 5 (May 2005), 844–854. https://doi.org/10.1097/01.alc.0000164374.32229.a2

[10] Carolina de Meneses-Gaya, Antonio Waldo Zuardi, Sonia Regina Loureiro, and José Alexandre S. Crippa. 2009. Alcohol Use Disorders Identification Test (AUDIT): An updated systematic review of psychometric properties. *Psychology & Neuroscience* 2, 1 (Jan. 2009), 83–97. https://doi.org/10.3922/j.psns.2009.1.12

[11] Andres Deik, Rachel Saunders-Pullman, and Marta San Luciano. 2012. Substance abuse and movement disorders: complex interactions and comorbidities. *Current drug abuse reviews* 5, 3 (2012), 243–253.

[12] Public Health England. [n.d.]. Alcohol dependence prevalence in England. url: https://www.gov.uk/government/publications/alcohol-dependence-prevalence-in-england;. https://www.gov.uk/government/publications/alcohol-dependence-prevalence-in-england; Accessed: 5-Feb-2020.

[13] Public Health England. [n.d.]. Public Health Profiles. url: https://fingertips.phe.org.uk/public-health-dashboard-ft#page/11/gid/1938133154/pat/6/par/E12000006/ati/102/are/E10000015/iid/93011/age/168/sex/4;. https://fingertips.phe.org.uk/public-health-dashboard-ft#page/11/gid/1938133154/pat/6/par/E12000006/ati/102/are/E10000015/iid/93011/age/168/sex/4; Accessed: 5-Feb-2020.

[14] Marcus Forsythe and Geraldine A. Lee. 2012. The evidence for implementing alcohol screening and intervention in the emergency department – Time to act. *International Emergency Nursing* 20, 3 (July 2012), 167–172. https://doi.org/10.1016/j.ienj.2011.09.006

[15] KIM BAO GIANG, FREDIK SPAK, TRUONG VIET DZUNG, and PETER ALLEBECK. 2005. THE USE OF AUDIT TO ASSESS LEVEL OF ALCOHOL PROBLEMS IN RURAL VIETNAM. *Alcohol and Alcoholism* 40, 6 (Aug. 2005), 578–583. https://doi.org/10.1093/alcalc/agh198

[16] Conor Gilligan, Kristen G. Anderson, Benjamin O. Ladd, Yun Ming Yong, and Michael David. 2019. Inaccuracies in survey reporting of alcohol consumption. *BMC Public Health* 19, 1 (Dec. 2019). https://doi.org/10.1186/s12889-019-7987-3

[17] Mark Andrew Hall. 1999. Correlation-based feature selection for machine learning. (1999).

[18] Maja Hansson, Jayanti Chotai, Annika Nordström, and Owe Bodlund. 2009. Comparison of two self-rating scales to detect depression: HADS and PHQ-9. *British Journal of General Practice* 59, 566 (Sept. 2009), e283–e288. https://doi.org/10.3399/bjgp09x454070

[19] Sion Kim Harris and John R Knight. 2014. Putting the screen in screening: technology-based alcohol screening and brief interventions in medical settings. *Alcohol Research: Current Reviews* 36, 1 (2014), 63.

[20] Daniel F. Hermens, Jim Lagopoulos, Juliette Tobias-Webb, Tamara De Regt, Glenys Dore, Lisa Juckes, Noeline Latt, and Ian B. Hickie. 2013. Pathways to alcohol-induced brain impairment in young people: A review. *Cortex* 49, 1 (jan 2013), 3–17. https://doi.org/10.1016/j.cortex.2012.05.021

[21] Adianes Herrera-Díaz, Raúl Mendoza-Quiñones, Lester Melie-Garcia, Eduardo Martínez-Montes, Gretel Sanabria-Diaz, Yuniel Romero-Quintana, Iraklys Salazar-Guerra, Mario Carballoso-Acosta, and Antonio Caballero-Moreno. 2015. Functional Connectivity and Quantitative EEG in Women with Alcohol Use Disorders: A Resting-State Study. *Brain Topography* 29, 3 (Dec. 2015), 368–381. https://doi.org/10.1007/s10548-015-0467-x

[22] John C. Higgins-Biddle and Thomas F. Babor. 2018. A review of the Alcohol Use Disorders Identification Test (AUDIT), AUDIT-C, and USAUDIT for screening in the United States: Past issues and future directions. *The American Journal of Drug and Alcohol Abuse* 44, 6 (May 2018), 578–586. https://doi.org/10.1080/00952990.2018.1456545

[23] Susan Homack and Cynthia A. Riccio. 2004. A meta-analysis of the sensitivity and specificity of the Stroop Color and Word Test with children. *Archives of clinical Neuropsychology* 19, 6 (2004), 725–743.

[24] Katia Iglesias, Frank Sporkert, Jean-Bernard Daeppen, Gerhard Gmel, and Stephanie Baggio. 2018. Comparison of self-reported measures of alcohol-related dependence among young Swiss men: a study protocol for a cross-sectional controlled sample. *BMJ Open* 8, 7 (2018). https://doi.org/10.1136/bmjopen-2018-023632 arXiv:https://bmjopen.bmj.com/content/8/7/e023632.full.pdf

[25] Jittrapol Intarasirisawat, Chee Siang Ang, Christos Efstratiou, Luke William Feidhlim Dickens, and Rupert Page. 2019. Exploring the Touch and Motion Features in Game-Based Cognitive Assessments. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (Sept. 2019), 1–25. https://doi.org/10.1145/3351245

[26] Sameer Jauhar, E. Jane Marshall, and Iain D. Smith. 2014. Alcohol and cognitive impairment. *Advances in Psychiatric Treatment* 20, 5 (sep 2014), 304–313. https://doi.org/10.1192/apt.bp.113.011981

[27] Zarnie Khadjesari, Ian R. White, Jim McCambridge, Louise Marston, Paul Wallace, Christine Godfrey, and Elizabeth Murray. 2017. Validation of the AUDIT-C in adults seeking help with their drinking online. *Addiction Science & Clinical Practice* 12, 1 (Jan. 2017). https://doi.org/10.1186/s13722-016-0066-5

[28] Alex Lau-Zhu, Emily A. Holmes, Sally Butterfield, and Joni Holmes. 2017. Selective Association Between Tetris Game Play and Visuospatial Working Memory: A Preliminary Investigation. *Applied Cognitive Psychology* 31, 4 (jul 2017), 438–445. https://doi.org/10.1002/acp.3339

[29] Qian Liu, Xinyi Zhu, Albert Ziegler, and Jiannong Shi. 2015. The effects of inhibitory control training for preschoolers on reasoning ability and neural activity. *Scientific Reports* 5, 1 (sep 2015). https://doi.org/10.1038/srep14200

[30] Jim Lumsden, Elizabeth A Edwards, Natalia S Lawrence, David Coyle, and Marcus R Munafò. 2016. Gamification of Cognitive Assessment and Cognitive Training: A Systematic Review of Applications and Efficacy. *JMIR Serious Games* 4, 2 (jul 2016), e11. https://doi.org/10.2196/games.5888

[31] V. Manera, P. D. Petit, A. Derreumaux, I. Orvieto, M. Romagnoli, G. Lyttle, R. David, and P. H. Robert. 2015. 'Kitchen and cooking,' a serious game for mild cognitive impairment and Alzheimer's disease: a pilot study. *Frontiers in aging neuroscience* 7 (Mar 17 2015), 24. https://doi.org/10.3389/fnagi.2015.00024[doi] LR: 20170220; JID: 101525824; OID: NLM: PMC4362400; OTO: NOTNLM; 2015/01/14 [received]; 2015/02/22 [accepted]; epublish.

[32] Lisa A Marsch, Warren K Bickel, and Michael J Grabinski. 2007. Application of interactive, computer technology to adolescent substance abuse prevention and treatment. *Adolescent Medicine: State of the Art Reviews* 18, 2 (2007), 342.

[33] Peter R Martin, Charles K Singleton, and Susanne Hiller-Sturmhöfel. 2003. The role of thiamine deficiency in alcoholic brain disease. *Alcohol research & health : the journal of the National Institute on Alcohol Abuse and Alcoholism* 27, 2 (2003), 134–42.

[34] P Mavrogiorgou, R Mergl, P Tigges, J El Husseini, A Schröter, G Juckel, M Zaudig, and U Hegerl. 2001. Kinematic analysis of handwriting movements in patients with obsessive-compulsive disorder. *Journal of neurology, neurosurgery, and psychiatry* 70, 5 (May 2001), 605–12.

[35] L. Daniel Maxim, Ron Niebo, and Mark J. Utell. 2014. Screening tests: a review with examples. *Inhalation Toxicology* 26, 13 (Sept. 2014), 811–828. https://doi.org/10.3109/08958378.2014.955932

[36] Wajid Mumtaz, Pham Lam Vuong, Aamir Saeed Malik, and Rusdi Bin Abd Rashid. 2017. A review on EEG-based methods for screening and diagnosing alcohol use disorder. *Cognitive Neurodynamics* 12, 2 (Dec. 2017), 141–156. https://doi.org/10.1007/s11571-017-9465-x

[37] Wajid Mumtaz, Pham Lam Vuong, Likun Xia, Aamir Saeed Malik, and Rusdi Bin Abd Rashid. 2016. An EEG-based machine learning method to screen alcohol use disorder. *Cognitive Neurodynamics* 11, 2 (Oct. 2016), 161–171. https://doi.org/10.1007/s11571-016-9416-y

[38] Ziad S. Nasreddine, Natalie A. Phillips, Valérie Bédirian, Simon Charbonneau, Victor Whitehead, Isabelle Collin, Jeffrey L. Cummings, and Howard Chertkow. 2005. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society* 53, 4 (2005), 695–699.

[39] World Health Organization. 1992. *ICD 10: International Statistical Classification of Diseases and Related Health Problems Volume 1.* Amer Psychiatric Pub Inc. https://www.amazon.com/ICD-10-International-Statistical-Classification/dp/9241544198?SubscriptionId=AKIAIOBINVZYXZQZ2U3A&tag=chimbori05-20&linkCode=xm2&camp=2025&creative=165953&creativeASIN=9241544198

[40] M. Oscar-Berman, B. Shagrin, D. L. Evert, and C. Epstein. 1997. Impairments of brain and behavior: the neurological effects of alcohol. *Alcohol Health Res World* 21, 1 (1997), 65–75.

[41] Tina Phan, Owen Carter, Claire Adams, Grant Waterer, Li Ping Chung, Maxine Hawkins, Cobie Rudd, Mel Ziman, and Natalie Strobel. 2016. Discriminant validity of the Hospital Anxiety and Depression Scale, Beck Depression Inventory (II) and Beck Anxiety Inventory to confirmed clinical diagnosis of depression and anxiety in patients with chronic obstructive pulmonary disease. *Chronic Respiratory Disease* 13, 3 (June 2016), 220–228. https://doi.org/10.1177/1479972316634604

[42] J. Rehm. 2011. The risks associated with alcohol use and alcoholism. *Alcohol Res Health* 34, 2 (2011), 135–143.

[43] H.-J. Rumpf. 2002. SCREENING FOR ALCOHOL USE DISORDERS AND AT-RISK DRINKING IN THE GENERAL POPULATION: PSYCHOMETRIC PERFORMANCE OF THREE QUESTIONNAIRES. *Alcohol and Alcoholism* 37, 3 (May 2002), 261–268. https://doi.org/10.1093/alcalc/37.3.261

[44] Darshan Santani, Trinh-Minh-Tri Do, Florian Labhart, Sara Landolt, Emmanuel Kuntsche, and Daniel Gatica-Perez. 2018. DrinkSense: Characterizing Youth Drinking Behavior Using Smartphones. *IEEE Transactions on Mobile Computing* 17, 10 (Oct. 2018), 2279–2292. https://doi.org/10.1109/tmc.2018.2797901

[45] A. Schröter, R. Mergl, K. Bürger, H. Hampel, H.-J. Möller, and U. Hegerl. 2003. Kinematic Analysis of Handwriting Movements in Patients with Alzheimer's Disease, Mild Cognitive Impairment, Depression and Healthy Subjects. *Dementia and Geriatric Cognitive Disorders* 15, 3 (2003), 132–142. https://doi.org/10.1159/000068484

[46] J. Paul Seale, John M. Boltri, Sylvia Shellenberger, Mary M. Velasquez, Monica Cornelius, Monique Guyinn, Ike Okosun, and Heather Sumner. 2006. Primary Care Validation of a Single Screening Question for Drinkers. *Journal of Studies on Alcohol* 67, 5 (Sept. 2006), 778–784. https://doi.org/10.15288/jsa.2006.67.778

[47] R Philip Snaith. 2003. *Health and Quality of Life Outcomes* 1, 1 (2003), 29. https://doi.org/10.1186/1477-7525-1-29

[48] Shota Suzumura, Aiko Osawa, Natsumi Maeda, Yuko Sano, Akihiko Kandori, Tomohiko Mizuguchi, Ying Yin, and Izumi Kondo. 2018. Differences among patients with Alzheimer's disease, older adults with mild cognitive impairment and healthy older adults in finger dexterity. *Geriatrics & Gerontology International* 18, 6 (March 2018), 907–914. https://doi.org/10.1111/ggi.13277

[49] Peter Tigges, Roland Mergl, Thomas Frodl, Eva M. Meisenzahl, Jürgen Gallinat, Andreas Schröter, Michael Riedel, Norbert Müller, Hans-Jürgen Möller, and Ulrich Hegerl. 2000. Digitized analysis of abnormal hand–motor performance in schizophrenic patients. *Schizophrenia Research* 45, 1-2 (sep 2000), 133–143. https://doi.org/10.1016/s0920-9964(99)00185-1

[50] Tom N. Tombaugh and Nancy J. McIntyre. 1992. The mini-mental state examination: a comprehensive review. *Journal of the American Geriatrics Society* 40, 9 (1992), 922–935.

[51] T. Tong, M. Chignell, M. C. Tierney, and J. Lee. 2016. A Serious Game for Clinical Assessment of Cognitive Status: Validation Study. *JMIR serious games* 4, 1 (May 27 2016), e7. https://doi.org/10.2196/games.5006[doi] LR: 20170220; JID: 101645255; OID: NLM: PMC4902858; OTO: NOTNLM; 2015/08/05 [received]; 2016/02/29 [accepted]; 2015/11/30 [revised]; epublish.

[52] Louis A Trevisan, Nashaat Boutros, Ismene L Petrakis, and John H Krystal. 1998. Complications of alcohol withdrawal: pathophysiological insights. *Alcohol Research* 22, 1 (1998), 61.

[53] Vanessa Vallejo, Patric Wyss, Luca Rampa, Andrei V. Mitache, René M. Müri, Urs P. Mosimann, and Tobias Nef. 2017. Evaluation of a novel Serious Game based assessment tool for patients with Alzheimer's disease. *PLOS ONE* 12, 5 (may 2017), e0175999. https://doi.org/10.1371/journal.pone.0175999

[54] Donald W. Zeigler, Claire C. Wang, Richard A. Yoast, Barry D. Dickinson, Mary Anne McCaffree, Carolyn B. Robinowitz, and Melvyn L. Sterling. 2005. The neurocognitive effects of alcohol on adolescents and college students. *Preventive Medicine* 40, 1 (jan 2005), 23–32. https://doi.org/10.1016/j.ypmed.2004.04.044

[55] Stelios Zygouris, Dimitrios Giakoumis, Konstantinos Votis, Stefanos Doumpoulakis, Konstantinos Ntovas, Sofia Segkouli, Charalampos Karagiannidis, Dimitrios Tzovaras, and Magda Tsolaki. 2015. Can a virtual reality cognitive training application fulfill a dual role? Using the virtual supermarket cognitive training application as a screening tool for mild cognitive impairment. *Journal of Alzheimer's Disease* 44, 4 (2015), 1333–1347.