JMIR Mental Health

Journal Impact Factor (JIF) (2023): 4.8
Volume 2 (2015), Issue 2 ISSN 2368-7959 Editor in Chief: John Torous, MD, MBI

Contents

Original Papers

Effectiveness of Computer-Assisted Therapy for Substance Dependence Using Breaking Free Online: Subgroup Analyses of a Heterogeneous Sample of Service Users (e13)
Sarah Elison, Glyn Davies, Jonathan Ward. . . . 3

An Online, Moderated Peer-to-Peer Support Bulletin Board for Depression: User-Perceived Advantages and Disadvantages (e14)
Kathleen Griffiths, Julia Reynolds, Sara Vassallo. . . . 15

Integrating Health Behavior Theory and Design Elements in Serious Games (e11)
Colleen Cheek, Theresa Fleming, Mathijs Lucassen, Heather Bridgman, Karolina Stasiak, Matthew Shepherd, Peter Orpin. . . . 37

Identifying Chinese Microblog Users With High Suicide Probability Using Internet-Based Profile and Linguistic Features: Classification Model (e17)
Li Guan, Bibo Hao, Qijin Cheng, Paul Yip, Tingshao Zhu. . . . 58

Word Recall: Cognitive Performance Within Internet Surveys (e20)
Shannon Runge, Benjamin Craig, Heather Jim. . . . 68

Factor Structure of the Internet Addiction Test in Online Gamers and Poker Players (e12)
Yasser Khazaal, Sophia Achab, Joel Billieux, Gabriel Thorens, Daniele Zullino, Magali Dufour, Stéphane Rothen. . . . 98

Twitter: A Novel Tool for Studying the Health and Social Needs of Transgender Communities (e16)
Evan Krueger, Sean Young. . . . 111

Viewpoint

Adjustment Disorders Are Uniquely Suited for eHealth Interventions: Concept and Case Study (e15)
Andreas Maercker, Rahel Bachem, Louisa Lorenz, Christian Moser, Thomas Berger. . . . 26

Review

Online Peer-to-Peer Support for Young People With Mental Health Problems: A Systematic Review (e19)
Kathina Ali, Louise Farrer, Amelia Gulliver, Kathleen Griffiths. . . . 81

XSL•FO
RenderX

Short Paper

Preferences of Young Adults With First-Episode Psychosis for Receiving Specialized Mental Health Services Using Technology: A Survey Study (e18)
Shalini Lal, Jennifer Dell'Elce, Natasha Tucci, Rebecca Fuhrer, Robyn Tamblyn, Ashok Malla. 90

XSL•FO
RenderX

Original Paper

Effectiveness of Computer-Assisted Therapy for Substance Dependence Using Breaking Free Online: Subgroup Analyses of a Heterogeneous Sample of Service Users

Sarah Elison[1], BSc (Hons), MSc, MPhil, PhD; Glyn Davies[1], BSc (Hons), MSc (Econ); Jonathan Ward[1], BA (Econ), MSc, DClinPsy

Breaking Free Online, Manchester, United Kingdom

Corresponding Author:
Sarah Elison, BSc (Hons), MSc, MPhil, PhD
Breaking Free Online
274 Deansgate
Manchester, M3 4JB
United Kingdom
Phone: 44 0161 834 ext 4647
Fax: 44 0161 834 4647
Email: selison@breakingfreegroup.com

Abstract

Background: Substance misuse services within the United Kingdom have traditionally been oriented to opiate and crack users, and attended predominantly by male service users. Groups who do not fit this demographic, such as women or those whose primary drug of choice is neither heroin nor crack, have tended to be underrepresented in services. In addition, there can be stigma associated with traditional opiate and crack-centric services. Therefore, the computerized treatment and recovery program, Breaking Free Online (BFO), was developed to enable service users to access confidential support for dependence on a wide range of substances. BFO is delivered as computer-assisted therapy (CAT), or, where appropriate, used as self-help.

Objective: The aim of this study was to report psychometric outcomes data from 393 service users accessing online support for substance misuse via BFO.

Methods: Following initial referral to substance misuse services, all participants were supported in setting up a BFO login by a practitioner or peer mentor, and, where required, assisted as they completed an online baseline assessment battery contained within the BFO program. Following a period of engagement with BFO, all participants completed the same battery of assessments, and changes in the scores on these assessments were examined.

Results: Significant improvements were found across the 393 service users in several areas of psychosocial functioning, including quality of life, severity of alcohol and drug dependence, depression, and anxiety (P=<.001 across all aspects of functioning). Additionally, significant improvements were found within specific subgroups of participants, including females (P=.001-<.001), males (P=.004-<.001), service users reporting alcohol dependence (P=.002-<.001), opiate and crack dependence (P=.014-<.001), and those seeking support for other substances that may be less well represented in the substance misuse sector (P=.001-<.001).

Conclusions: Data from this study indicates that BFO is an effective clinical treatment for a wide range of individuals requiring support for substance misuse. Further work is currently underway to examine more closely the clinical effectiveness of the program.

(JMIR Mental Health 2015;2(2):e13) doi:10.2196/mental.4355

KEYWORDS

substance misuse; computer-assisted therapy; treatment; psychosocial interventions

Introduction

In recent years, the substance misuse sector has been subject to changes that may have long-term implications for societal attitudes to substance misuse, and the ways in which individuals facing such difficulties may best be supported [1]. For example, there seems to be increasing recognition that the range of substances of misuse and dependence may be wider than alcohol, opiates, and crack. Regulatory bodies such as the National Institute for Health and Care Excellence (NICE) and

XSL•FO
RenderX

the National Treatment Agency (NTA) have published guidelines around the treatment of cannabis and stimulants [2], and prescription and over the counter medication [3] dependencies. However, substance misuse services have not traditionally been designed for non-opiate or crack-using individuals [4]. Such services have been designed primarily to provide support to opiate and crack-using individuals, who represent the more traditionally identifiable substance dependent groups within society. Although services may have some provisions for non-opiate or crack-using individuals, they continue to be perceived as being oriented towards opiate and crack dependencies [5].

There is growing interest in the psychosocial issues that drive substance dependence, with dependence now being seen by some as a symptom of underlying mental health issues [6], relationships [7], lifestyle [8], and other social and economic difficulties [9]. Consequently, attitudes appear to be changing regarding the possibility of someone recovering from substance dependence, as previous conceptions of dependence as a chronic relapsing condition are being challenged [10-13].

In addition to non-opiate and crack-using individuals, women are another demographic that though may be presenting at services more frequently, albeit still underrepresented. Women may find it particularly difficult to access services, often due to the risks associated with the involvement of social services when dependent children are involved [14]. Despite this underrepresentation, recent estimates demonstrate that much higher proportions of women may be dependent on substances than come into contact with services [4]. As well, the number of women in the United Kingdom consuming alcohol to hazardous, harmful or dependent levels, and the rate of alcohol-related deaths among women is also increasing [15,16]. One major barrier to such individuals accessing substance misuse services may be the stigma associated with seeking support for substance dependence, especially when such services may still viewed by some as being opiate and crack-centric.

Thus, there appears to be changes in both substance use patterns and the types of groups requiring support from substance misuse services. Therefore, there is a need for evidence-based, clinically effective intervention approaches that are appropriate for addressing a wide variety of forms of substance misuse and dependence, and the mental health and wider psychosocial issues that underpin them.

Although there have been interventions developed to address substance misuse and co-occurring mental health issues [17], few address mild to moderate mental health difficulties, and instead target more severe mental health issues. Despite being reported in the literature as effective, there are challenges in getting these interventions commissioned within the substance misuse sector and making them available to the general public [18]. However, such intervention approaches may be beneficial as they provide support in such a way that helps to overcome barriers to access due to stigma, especially if individuals requiring support from such services may not necessarily identify themselves as having a mental illness (ie, mild or temporary mental health issues).

One potential solution to these requirements, in terms of reducing barriers to accessing services for individuals with mild to moderate mental health difficulties, is Breaking Free Online (BFO), an online psychosocial intervention that has been developed and commissioned in over 60 local authorities in UK-based substance misuse services [19-22].

BFO can be delivered as computer-assisted therapy (CAT), or as self-help, and is designed to support people in their recovery from substance misuse. BFO delivers evidence-based psychosocial intervention strategies that are compliant with NICE guidance around interventions for substance misuse [23], offers a range of different multimedia formats, and specifically targets 36 different substances (ie, substitute medications, legal highs, and prescribed medications of abuse).

The BFO program provides access to 22 interactive, evidence-based intervention strategies taken from cognitive-behavioral therapy (CBT) [24] and mindfulness approaches [25]. Audio and visual technology is used to deliver intervention content that has traditionally been delivered with service users via face to face interactions with a practitioner or paper-based documents. The content of the program was developed through consultation with substance misuse and mental health professionals and service users, and a review of the literature around evidence-based approaches for substance misuse. All intervention content is structured around a six domain model that conceptualizes various aspects of biopsychosocial functioning associated with substance misuse and any comorbid mental health difficulties. The model, the Lifestyle Balance Model (LBM) [22], was developed by the authors of this study and is based on the commonly used five-factor model used in mental health case formulation [26,27]. The LBM offers a guided node-link map structure for understanding an individual's substance use and associated difficulties, irrespective of the type of substance difficulty [28,29]. The domains of functioning contained within the LBM are depicted in Figure 1.

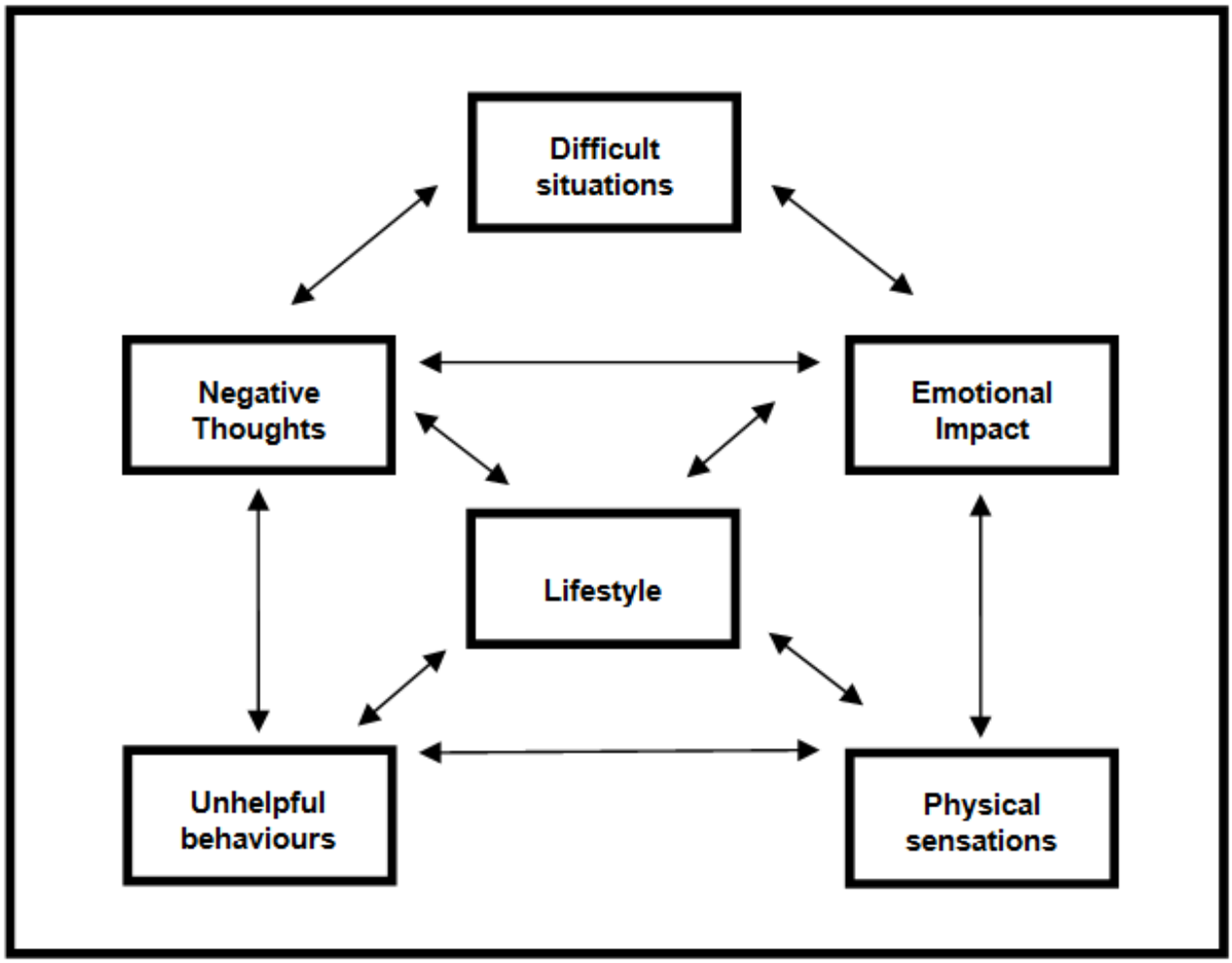
An assessment within the program, the Recovery Progression Measure (RPM) [submitted], was developed by the authors of this study to measure functioning in the six LBM domains. The RPM is an online-based series of assessments that can be completed by service users with support from a professional or alone. The assessments are completed at intervals to provide follow-up data. Data from the program are stored on a secure, offsite server, and security protocol conforms to Caldicott guidance [30], and other relevant data protection and legal requirements. All stored data are also completely confidential and contain no identifiable service user information.

Initial evaluations indicated that the BFO program significantly reduced substance misuse, and improved mental health and quality of life [20,21]. As well, there is evidence to support the modality of the delivery and therapeutic components within the program [25,31,32]. Accessing online interventions, like BFO, may help to overcome barriers such as the shame and stigma sometimes associated with accessing more visible, traditional opiate and crack- centric drug and alcohol services [33]. It can also ensure that access is both confidential and anonymous.

XSL•FO
RenderX

This study was an outcomes evaluation of the effectiveness of BFO on a heterogeneous sample of 393 service users accessing support for substance dependence via substance misuse services in the United Kingdom. In addition to assessing the effectiveness of BFO for the group as a whole, subgroup analyses were reported around the effectiveness of previously less represented groups in services, such as women, and service users seeking support for dependence on substances other than opiates and crack, including alcohol.

Figure 1. Domains of functioning contained within the LBM.



Methods

Design

This is a quantitative, repeated-measures, psychometric outcomes study that reports data from 393 service users presenting to drug and alcohol services accessing support for substance dependence via BFO.

Participants

Participants were from a group of 785 service users accessing an updated version of BFO released in January 2013, who had all completed a baseline assessment, and logged onto the program to access at least one intervention strategy during the period from January-December, 2013. Of this group of 785 service users, 393 completed a follow-up post-intervention assessment providing data to the study, resulting in a response rate of 50.0% (393/785). All participants had either self-referred or been referred into a specialist substance misuse service by a health or social care professional (eg, general practitioner or social worker), or via the criminal justice system (eg, police or probation). All participants were provided free access to BFO from the referring service. Ethical approval was previously granted for using the Breaking Free Online research database containing service user assessment data (application 'Breaking Free Online Research Database'; Research Ethics Committee Reference 12/LO/0076). Participants were not given any incentive to either use the BFO program or provide data for the study.

A total of 224 (57.0%, 224/393) participants were male, and the mean age was 42.4 years (range 15-73 years, SD 11.3). The group was predominantly White-British/White-Irish (95.9%, 377/393), with the rest being Asian/Asian-British (1.3%, 5/393), Black/Black-British (0.3%, 1/393), mixed-race (2.5%, 10/393), or other unspecified ethnicities (0.3%, 1/393). With respect to the severity of the drug dependence at baseline, of those reporting an illicit substance as their principal substance of dependence, a total of 169 participants (92.9%, 169/182) reached the cut-off for clinically significant drug dependence by scoring ≤3 on the Severity of Dependence Scale (SDS). For those 299 reporting alcohol as being one of their substances of dependence

XSL•FO
RenderX

(with 89 of these also reporting being dependent on an illicit drug as well as alcohol), 259 participants (86.6%, 259/299) reached cut-off for clinically significant alcohol dependence. A total of 275 participants (70.0%, 275/393) reached clinical cut-off for clinically significant depression by scoring ≤4 on the Patient Health Questionnaire (PHQ), and of those who reported some difficulties with anxiety, 89 (69.0%, 89/129) reached cut-off for clinically significant anxiety by scoring ≤4 on the General Anxiety Disorder (GAD) scale.

A wide range of substances were reported by participating service users as being the main problem substance, from prescribed medications, to opiate substitute medications, and synthetic psychoactive substances, however, alcohol was the most common (53.4%, 210/393), followed by heroin (13.2%, 52/393), and cannabis (11.7%, 46/393). In total, approximately a quarter (24.2%, 95/393) cited non-opiate and crack substances (excluding alcohol) as being their main substance of dependence, and just over one fifth (22.4%, 88/393) cited their main substance of dependence being an opiate (including opiate substitutes) or crack. Full details of the substances used by the sample are provided in Table 1.

Table 1. The main substances of dependence (N=19).

Substance		n (%)
Alcohol		210 (53.4)
Non-opiate (and opiate substitute) or crack		
	Amphetamines	11 (11.0)
	Cannabis	46 (11.7)
	Cocaine	16 (4.1)
	Diazepam	7 (7.0)
	Dihydrocodeine	2 (0.5)
	Etizolam	1 (0.3)
	GBL	1 (0.3)
	Ketamine	2 (0.5)
	Khat	2 (0.5)
	Mephedrone	3 (0.8)
	Temazepam	1 (0.3)
	Tramadol	2 (0.5)
	Zopiclon	1 (0.3)
	Total	95 (24.2)
Opiate (and opiate substitute) or crack		
	Buprenorphine	5 (1.3)
	Crack	12 (3.1)
	Heroin	52 (1.2)
	Methadone	18 (4.6)
	Suboxone	1 (0.3)
	Total	88 (22.4)

Procedure

Following initial referral to services, all participants were supported in setting up a BFO login by a practitioner or peer mentor, and, where required, assisted as they completed an online baseline assessment battery contained within the BFO program. The assessments are shown in Textbox 1.

XSL•FO
RenderX

Textbox 1. Baseline assessments of the BFO program.

Assessments

- Recovery Progression Measure (RPM) [submitted]
  - A 36-item measure comprising 6 Likert scale items, each with 11 points (ie, 0-10) that determines the impact of difficult situations, negative thoughts, emotions, unhelpful behaviors, physical sensations, and lifestyle on substance use.
  - Contains 30 dichotomous 'yes/no' response items that measure the presence or absence of specific psychosocial issues within each of the 6 Likert scale items.
- World Health Organization Quality of Life Measure (WHOQOL-BREF) [34]
  - A total of 5 items (1, 2, 17, 18, and 20) from the WHOQOL-BREF were selected for measuring general quality of life.
- Patient Health Questionnaire (PHQ-9) [35]
  - A 9-item scale that measures the levels of depression, and also contains validated clinical norms.
- General Anxiety Disorder Scale (GAD-7) [36]
  - A 7-item scale that measures the levels of anxiety, and also contains validated clinical norms.
- Severity of Dependence Scale (SDS) [37]
  - A 5-item scale that measures the severity of alcohol dependence.

The RPM was specifically developed by the authors of this study as a tool to measure the degree of 'recovery progression' an individual achieves in each domain. The internal reliability of the RPM measure was excellent (alpha>.70), with item-total correlations revealing moderate to excellent reliability of individual items. As well, the convergent validity was excellent, with the RPM measure correlating significantly with scores on standardized psychometric measures of related constructs, such as the SDS [37], PHQ [35], and GAD [36]. Exploratory factor analyses (EFA) revealed the RPM contained an underlying factor structure consisting of eight components.

Upon completion of the baseline assessment, individuals were provided with full access to BFO. Most individuals accessed BFO both within services with support from a practitioner or peer mentor, and at home or in community settings with internet access, such as local libraries. Time periods of engagement with the program varied, reflecting the program's ability to be tailored to the needs of the individual; some participants engaged for longer periods than others, depending on their perception of need. The amount of engagement with BFO depended on the individual's perception of how much they felt they needed to use the program in order to address the specific type and severity of difficulties they were experiencing. At the end of each individual's period of engagement with the program, the same battery of assessments was completed online.

Analysis

As data were not normally distributed, non-parametric Wilcoxon signed-ranks tests were conducted to examine the changes in psychometric scores from baseline to post-intervention follow up. Effect sizes were also calculated for these, and multiple linear regressions were conducted to examine the association between the time in weeks and changes in psychometric scores.

Results

Engagement with the BFO program varied by the period of weeks, the total amount of time spent online, and the number of intervention strategies accessed. The mean engagement period was 4.6 weeks (range from 1-12 weeks, SD 3.4), the mean time spent online was 4.7 hours (range 18 minutes-109 hours, SD 7.8), and the mean number of strategies accessed was 6.8 (range 1-12, SD 3.7).

As Shapiro-Wilk tests revealed data to be non-normally distributed (P<.05), non-parametric Wilcoxon signed-ranks tests were run to examine possible changes in psychometric scores from baseline to follow-up. Analyses revealed that a number of statistically significant changes were identified both in terms of psychometric scores, and that these were evident in each of the subgroups included in the study. In addition to exploring changes in scores from baseline to post-intervention follow-up, effect sizes were calculated in order to examine the strength of any identified changes in scores. Linear regressions were run to ascertain whether changes in scores were a function of time elapsed in weeks between baseline and follow-up assessments. Outcomes for each of the following groups are reported in Table 2.

XSL•FO
RenderX

Table 2. Psychometric outcomes from baseline to post-treatment follow-up (N=393).

Category		Changes in psychometric scores					Linear regression[a]		
		Baseline, mean (SD)	Follow-up, mean (SD)	*z* score	*P* value	Effect size, *r*	*R*	R^2	*P* value
All data, n=393									
	Quality of life	8.9 (4.5)	11.0 (4.5)	8.417	<.001	0.43	.065	.004	.198
	RPM	38.2 (14.2)	32.7 (15.7)	-6.864	<.001	0.35	.023	.001	.654
	SDS-alcohol	8.5 (4.5)	5.4 (4.2)	-9.963	<.001	0.58	.042	.002	.474
	SDS-drugs	8.4 (3.3)	5.6 (4.1)	-7.258	<.001	0.55	.081	.007	.285
	PHQ	8.4 (5.8)	6.1 (5.5)	-6.380	<.001	0.59	.129	.017	.168
	GAD	9.6 (6.5)	5.7 (6.0)	-4.400	<.001	0.69	.162	.026	.311
Females, n=169									
	Quality of life	8.6 (4.3)	10.4 (4.3)	4.590	<.001	0.35	.098	.010	.206
	RPM	38.9 (13.4)	34.3 (14.5)	-3.741	<.001	0.29	.114	.013	.144
	SDS-alcohol	8.7 (4.5)	5.6 (3.9)	-6.968	<.001	0.61	.003	.001	.969
	SDS-drugs	8.8 (4.0)	6.4 (3.7)	-3.994	<.001	0.47	.127	.016	.288
	PHQ	8.4 (5.3)	5.6 (4.5)	-4.673	<.001	0.68	.098	.010	.206
	GAD	8.9 (6.4)	4.8 (3.3)	-3.419	.001	0.85	.114	.013	.144
Males, n=223									
	Quality of life	9.0 (4.67)	11.4 (4.5)	7.119	<.001	0.48	.057	.003	.401
	RPM	37.6 (14.8)	31.5 (16.5)	-5.851	<.001	0.40	.109	.012	.114
	SDS-alcohol	8.3 (4.6)	5.3 (4.4)	-7.148	<.001	0.55	.068	.005	.384
	SDS-drugs	8.2 (3.9)	5.0 (4.2)	-6.014	<.001	0.60	.041	.002	.679
	PHQ	8.5 (6.1)	6.5 (6.1)	-4.501	<.001	0.54	.089	.008	.468
	GAD	10.1 (6.5)	6.2 (7.3)	-2.900	.004	0.58	.180	.032	.389
Alcohol, n=210									
	Quality of life	8.9 (4.3)	10.4 (4.3)	4.851	<.001	0.34	.030	.001	.670
	RPM	38.1 (13.8)	34.6 (14.7)	-3.741	<.001	0.26	.025	.001	.728
	SDS-alcohol	8.4 (4.6)	5.8 (4.3)	-7.064	<.001	0.56	.067	.005	.397
	SDS-drugs	8.4 (4.0)	5.5 (1.1)	-5.155	<.001	0.53	.104	.011	.313
	PHQ	6.6 (3.5)	5.2 (3.7)	-3.153	.002	0.48	.220	.048	.157
	GAD								
Non-opiate and crack, n=95									
	Quality of life	8.3 (4.6)	11.4 (4.7)	5.778	<.001	0.59	.196	.039	.056
	RPM	39.6 (14.7)	31.8 (15.7)	-4.846	<.001	0.50	.041	.002	.693
	SDS-alcohol	9.1 (4.3)	5.5 (3.9)	-5.585	<.001	0.62	.081	.007	.474
	SDS-drugs	8.3 (3.6)	5.3 (3.8)	-3.223	.001	0.62	.206	.042	.302
	PHQ	10.9 (6.9)	7.2 (6.8)	-4.819	<.001	0.70	.146	.021	.326
	GAD	10.0 (6.3)	6.2 (6.3)	-3.669	<.001	0.71	.080	.006	.693
Opiate and crack, n=88									
	Quality of life	9.7 (4.8)	11.8 (4.5)	4.291	<.001	0.46	.011	.001	.922
	RPM	36.8 (14.7)	29.3 (17.4)	-3.462	.001	0.38	.005	.001	.965
	SDS-alcohol	7.7 (4.8)	4.4 (4.3)	-4.173	<.001	0.57	.135	.018	.322
	SDS-drugs	8.5 (4.0)	5.8 (4.2)	-4.053	<.001	0.56	.059	.004	.677

XSL•FO
RenderX

Category		Changes in psychometric scores					Linear regression[a]		
		Baseline, mean (SD)	Follow-up, mean (SD)	*z* score	*P* value	Effect size, *r*	*R*	R^2	*P* value
	PHQ	10.3 (7.2)	5.7 (5.2)	-2.861	.004	0.56	.034	.001	.869
	GAD	9.2 (6.6)	4.7 (5.4)	-2.450	.014	0.65	.280	.079	.332

[a]Changes in scores with time in weeks

The Whole Sample

A significant increase in quality of life was found, along with significant decreases in both alcohol and drug dependence, depression, anxiety, and other areas of psychosocial impairment, as measured by the RPM (all *P*<.001). Proportions of all participant's reaching cut-off scores decreased, including for; clinically significant drug dependence (92.9%-76.6%), alcohol dependence (86.6%-71.4%), depression (70.2%-50.0%), and anxiety (69.0%-46.3%). Effect sizes were large for drug dependence (*r*=0.55), depression (r=0.59), and anxiety (*r*=0.69), while medium effect sizes were found for quality of life (*r*=0.43) and RPM scores (*r*=0.35). Linear regressions revealed that time elapsed between baseline and follow-up assessment was not predictive of changes in scores.

Females

A significant increase in quality of life was found for females, along with significant decreases in both alcohol and drug dependence, depression, and other areas of psychosocial impairment measured by the RPM (all *P*<.001). A significant decrease in anxiety (*P*=.001) was also observed. Proportions of female's reaching cut-off scores decreased, including for; clinically significant alcohol dependence (85.4%-76.7%), drug dependence (93.3%-87.5%), depression (72.8%-53.2%), and anxiety (64.7%-50.0%). The effect size was very large for anxiety (*r*=0.85), large for alcohol dependence (*r*=0.61) and depression (*r*=0.68), medium for quality of life (*r*=0.35) and drug dependence (*r*=0.47), and small for RPM scores (*r*=0.29). Linear regressions revealed that time elapsed between baseline and follow-up assessment was not predictive of changes in scores.

Males

A significant increase in quality of life was found for males, along with significant decreases in both alcohol and drug dependence, depression, and other areas of psychosocial impairment measured by the RPM (all *P*<.001). A significant decrease was also found in anxiety (*P*=.004). Proportions of male's reaching cut-off scores decreased, including for; clinically significant alcohol dependence (87.6%-67.3%), drug dependence (92.5% to 68.9%), depression (68.2%-47.8%), and anxiety (71.8%-44.0%). Large effect sizes were found for alcohol (*r*=0.55), drug dependence (*r*=0.60), depression (*r*=0.54), and anxiety (*r*=0.58), and medium for quality of life (*r*=0.48) and RPM scores (*r*=0.40). Linear regressions revealed that time elapsed between baseline and follow-up assessment was not predictive of changes in scores.

Alcohol as the Main Substance of Misuse

A significant increase in quality of life was found for service users citing alcohol as their main substance of misuse, along with significant decreases in both alcohol and drug dependence, and other areas of psychosocial impairment measured by the RPM (all *P*<.001). A significant decrease was observed in depression (*P*=.002), however, no data were available for this subgroup for anxiety. Proportions of alcohol misusers reaching cut-off scores decreased, including for; clinically significant alcohol dependence (85.8%-72.0%), drug dependence (92.9%-74.0%), and depression (65.0%-48.8%). Large effect sizes were found for alcohol (*r*=0.56) and drug dependence (*r*=0.53), medium for quality of life (*r*=0.34) and depression (*r*=0.44), and small for RPM scores (*r*=0.26). Linear regressions revealed that time elapsed between baseline and follow-up assessment was not predictive of changes in scores.

Non-Opiate and Crack Users

A significant increase in quality of life was found for non-opiate and crack users, along with significant decreases in RPM scores, alcohol dependence, depression, and other areas of psychosocial impairment measured by the RPM (all *P*<.001). A significant decrease in drug dependence (*P*=.001) was also found. Proportions of non-opiate and crack users' reaching cut-off scores decreased, including for; clinically significant alcohol dependence (87.5%-77.5%), drug dependence (96.4%-88.9%), depression (78.9%-51.0%), and for anxiety (70.8%-48.1%). Large effect sizes were found for quality of life (*r*=0.59), RPM scores (*r*=0.50), alcohol (*r*=0.62), and drug dependence (*r*=0.62), and very large effect sizes for depression (*r*=0.70) and anxiety (*r*=0.71). Linear regressions revealed that time elapsed between baseline and follow-up assessment was not predictive of changes in scores.

Opiate and Crack Users

A significant increase in quality of life was found for opiate (and opiate substitute medications) and crack users, along with significant decreases in alcohol and drug dependence (all *P*<.001), RPM scores (*P*=.001), depression (*P*=.004), and anxiety (*P*=.014). Proportions of opiate and crack users reaching cut-off scores decreased, including for; clinically significant alcohol dependence (87.7%-57.1%), drug dependence (91.1%-75.0%), depression (72.7%-50.0%), and anxiety (67.2%-42.9%). Large effect sizes were found for alcohol (*r*=0.57), drug dependence (*r*=0.56), depression (*r*=0.56), and anxiety (*r*=0.65), and medium effect sizes for quality of life (*r*=0.46) and RPM (*r*=0.38). Linear regressions revealed that time elapsed between baseline and follow-up assessment was not predictive of changes in scores.

XSL•FO
RenderX

Discussion

Principal Findings

This study sought to explore clinical outcomes in a heterogeneous group of service users accessing Breaking Free Online (BFO), an online CAT program for substance misuse. Due to the heterogeneous nature of the sample, in addition to looking at changes in the sample as a whole, subgroup analyses were also conducted to ascertain whether specific groups of service users benefited from using the program.

When the sample as a whole was included in the analyses, significant improvements in all aspects of psychosocial functioning were identified, and relatively robust effect sizes were found. Similar findings were obtained in all subgroups including females, males, service users reporting their primary problem substance as being alcohol, non-opiate and crack users, and opiate (and opiate substitute medications) users. Improvements were also seen in the severity of dependence on alcohol and drugs, reductions in depression and anxiety, improvements in quality of life, and reductions in the six domains of biopsychosocial functioning measured by the Recovery Progression Measure (RPM). The findings related to RPM outcomes are not surprising given the BFO program contains evidence-based techniques specifically included because they support improvements in functioning in the domains measured by the RPM. It is possible that the improvements in substance dependence severity, mental health, and quality of life reported could result from these improvements in biopsychosocial functioning, and also contribute to improvements in the domains of biopsychosocial functioning measured by the RPM. Further work is needed to more fully understand the interrelationships between the techniques contained within the BFO program and the various psychometric outcomes reported.

Additionally, linear regressions revealed that time elapsed between baseline and post-intervention follow-up assessments in weeks was not predictive of the degree of change in psychometric scores. This would indicate that the length of time engaged with the BFO program (in weeks) was not associated with the degree of improvement across the psychometric outcomes measured. This may reflect the fact that service users engaged with the program for the length of time they felt necessary, according to their self-perceived level of need in terms of intervention intensity requirement, or that only a brief period of engagement was sufficient for improvement, which is consistent with findings from studies with other brief interventions for alcohol [38,39]. Alternatively, this could also indicate that BFHJ is not subject to the 'dosage' effect that other interventions based on cognitive-behavioral therapy (CBT) principles are [40,41], which may be explained by the fact that BFO is modular rather than linear in nature. This means that users of the program can access the sections of the program that are most relevant to them, without having to work sequentially through content that may not be relevant. In other words, the program can be tailored to the needs of the individual. The literature suggests that the more capable a complex behavioral change intervention is of being tailored, the more likely it is to be effective [42].

"However, despite the encouraging findings when psychometrics scores were used to examine changes from baseline to post-intervention follow-up in a number of subgroups of service users, the findings were less impressive for the proportions of service users reaching cut-offs scores for clinically significant alcohol and drug dependence, depression and anxiety. Although for all subgroups the proportion of service users reaching cut-offs for clinically significant alcohol and drug dependence reduced between the baseline and post-intervention assessments, these reductions were not particularly large. The largest reduction was seen in the opiate and crack-using group who went from 87.7% (77/88) at baseline to 57.1% (50/88) at the post-intervention assessment, a reduction in 30.6% of service users in this group. The smallest decrease was seen in the female service user group with respect to the proportion reaching cut-off scores for clinically significant drug dependence (93.3%-87.5%), a reduction of just 5.8%.

The findings generated by this study have provided initial outcomes data to support the effectiveness of BFO with groups that have historically been less well represented in traditional substance misuse services, such as female service users. Women are a group of individuals who have been found to face a unique set of barriers to accessing support for substance misuse [14]. Many of these are associated with issues around finding and funding childcare, as many services have been described as being environments that are not particularly child-friendly [43]. Many women also cite concerns over social services involvement as a key barrier to approaching substance misuse services, and the additional stigma many women feel around the fact that their substance misuse may cause people to question their role as an effective caregiver, a role that has traditionally been seen as a female one [44,45]. Despite these additional barriers that women face when accessing support for substance misuse, and the underrepresentation of women in standard services, the sample in this study was quite evenly split, with 43.0% (169/393) of the sample being female. This is over twice the proportion seen in traditional substance misuse services, which is usually around 20%, with some variation according to the specific substance [46]. This may demonstrate that BFO, a completely confidential intervention that can be accessed privately at home, can provide a solution to some of the barriers women who misuse substances face when attempting to overcome their difficulties.

Another group who may wish to access support from the privacy of their own home is alcohol consumers who feel that accessing support via traditional services is associated with stigma [47]. As with women, this group of individuals may prefer to access support in private, rather than having to admit in a more open way that they have difficulties with alcohol, particularly as many people have concerns about the impact this may have on their professional and personal lives. This may also be the case for service users who, at first, used alcohol in a recreational manner, and yet may find that, in time, they became dependent. Although these groups may prefer interventions that can be used privately at home, it does not negate the potential of the kinds of

XSL•FO
RenderX

interventions typically delivered in service environments to also be effective in the home environment [48,49].

With respect to the individuals that traditionally present at substance misuse services, which were those reporting they were dependent on opiates, opiate substitutes or crack, this subgroup had the same positive outcomes as in the other subgroups. However, the significance levels and effect sizes were not quite as strong as in the other groups, specifically in relation to quality of life and severity of depression and anxiety. This makes sense given this particular subgroup may have more deeply entrenched difficulties due to the substances on which they are dependent, and the additional complexity of addressing a physical and psychological dependence, as opposed to other substances that are primarily psychologically addictive. Additionally, opiate and crack-using individuals may be leading more chaotic lifestyles than other groups of substance dependent individuals, particularly given the criminal behaviors many need to engage in whilst attempting to financially support their habit [50].

Limitations

The promising outcomes obtained in this study highlight an opportunity for traditional substance misuse services to market CAT as a targeted intervention that meets the needs of hidden and emerging drug using populations. Equally, the outcomes indicate that CAT may be a clinically effective intervention approach for a range of groups of substance dependent individuals that are not usually well represented in traditional substance misuse services. However, although the research reported here includes a relatively substantial sample size (N=393), the response rate was relatively low at 50.0% (393/785), and the work is still preliminary and exploratory, so there were some limitations that deserve consideration when drawing conclusions.

Firstly, the sample in the study were self-selecting and so could conceivably be assumed to be relatively motivated individuals, and hence more likely to reduce their alcohol and drug intake than service users who were not participating in the study. Therefore the potential impact of motivation in determining the outcomes obtained is not fully understood. It is also not known why non-participating service users did not provide follow-up assessment data; the reasons for this can only be speculative, although internet access could be one potential avenue for future research into why some services accessed BFO successfully and provide follow-up data, and others did not. Additionally, all data reported were based on service users' self-reports of psychosocial functioning, and so may not have been entirely reliable.

It is also not known how relevant the findings from the study are to a wide range of ethnic groups, as the vast majority of participants in the study were White-British or White-Irish (95.9%, 377/393). However, this is representative of the demographic of service users accessing substance misuse services in the United Kingdom, in which approximately only 10% of service users are from black and minority ethnic (BME) communities [51]. BME communities are often described as hidden populations because of their underrepresentation within treatment services. There are multiple barriers to such BME communities accessing support via traditional substance misuse services such as lack of cultural sensitivity by the service, distrust of confidentiality, language barriers, stigma, and the failure of drug services to target minority ethnic drug users [52].

The BFO program could be used to engage members of BME communities as it can be delivered in any community setting away from traditional drug and alcohol services, such as places of residence, community halls, and religious buildings. As such, it can overcome the stigma of attending a traditional service. There is also a supporters section within BFO, which provides guidance to those who are supporting an individual using the programs to address their substance use. This supporter's guidance section could be used by individuals who may act as culturally appropriate supporters not associated with traditional drug and alcohol services, such as community elders, outreach workers, or other members of a community.

Another limitation lies in the fact that there was no control group included in the study. However, work is already underway that includes comparison controls, randomization to study groups, and a follow-up element to examine the long-term impact of BFO. Other work currently being conducted includes examining how different service users access the program in terms of time spent online and the specific outcomes that might be expected from accessing specific techniques within the program. Additionally, the possible additive effect of BFO when used in conjunction with 'treatment as usual' will also be examined as it is not known whether the outcomes obtained can be attributed to BFO, or whether the service users included in the present study may have improved due to the combined effect of BFO and other sources of support and treatment.

Conclusions

This study demonstrated that BFO provides, at least in the short-term, a clinically effective intervention option for a wide range of service users accessing support for issues around dependence to a range of different substances. Additionally, the outcomes reported here are a first in terms of CAT approaches for substance misuse as they come from a sample of service users accessing support in real world substance misuse services, not a sample of participants in a highly controlled research study with limited ecological validity. Furthermore, as BFO provides confidential support that can be accessed outside of standard opiate and crack-centric services, it may enable some groups of individuals to overcome barriers that may prevent them from approaching services for support via more traditional interventions, with stigma being a significant obstacle for many. A comprehensive research program is currently underway to evaluate the program further, and it is hoped that as the evidence base for BFO increases, the program will be made even more widely available, enabling more individuals to access and use it in their recovery from substance dependence.

XSL•FO
RenderX

Conflicts of Interest

The authors of this study are all employed by Breaking Free Group where the Breaking Free Online program was developed.

References

1. Department of Health (England) the devolved administrations Drug misusedependence UK guidelines on clinical management. London. 2007. In: Department of Health (England), Welsh Assembly Government and Northern Ireland Executive, ed URL: http://www.nta.nhs.uk/uploads/clinical_guidelines_2007.pdf [accessed 2015-04-16] [WebCite Cache ID 6XpYy2xO0]
2. NICE. National Clinical Practice Guideline Number 51. London: NICE Drug Misuse- Psychosocial Interventions URL: http://www.nice.org.uk/guidance/CG51 [accessed 2015-04-16] [WebCite Cache ID 6XpZ0IkZR]
3. National Treatment Agency for Substance Misuse. NTA. Addiction to medicine: An investigation into the configuration and commissioning of treatment services to support those who develop problems with prescriptiononly or over-the-counter medicine URL: http://www.appgita.com/wp-content/uploads/2011/05/Report%202%20NTA%20Addiction%20to%20Medicines.pdf [accessed 2015-04-16] [WebCite Cache ID 6XpZ7nnx0]
4. National Treatment Agency for Substance Misuse. NTA. Models of care for treatment of adult drug misusers: Update 2006; URL: http://www.nta.nhs.uk/uploads/nta_modelsofcare_update_2006_moc3.pdf [accessed 2015-04-16] [WebCite Cache ID 6XpZFFsRM]
5. Radcliffe P, Stevens A. Are drug treatment services only for 'thieving junkie scumbags'? Drug users and the management of stigmatised identities. Soc Sci Med 2008 Oct;67(7):1065-1073. [doi: 10.1016/j.socscimed.2008.06.004] [Medline: 18640760]
6. Delgadillo J. Depression and anxiety symptoms: measuring reliable change in alcohol and drug users. Advances in Dual Diagnosis 2012;5(3):102-114.
7. Stuart GL, O'Farrell TJ, Temple JR. Review of the association between treatment for substance misuse and reductions in intimate partner violence. Subst Use Misuse 2009;44(9-10):1298-1317 [FREE Full text] [doi: 10.1080/10826080902961385] [Medline: 19938919]
8. Pluck G, Lee KH, Lauder HE, Fox JM, Spence SA, Parks RW. Time perspective, depression, and substance misuse among the homeless. J Psychol 2008 Mar;142(2):159-168. [doi: 10.3200/JRLP.142.2.159-168] [Medline: 18447073]
9. Keyes KM, Hasin DS. Socio-economic status and problem alcohol use: the positive relationship between income and the DSM-IV alcohol abuse diagnosis. Addiction 2008 Jul;103(7):1120-1130 [FREE Full text] [doi: 10.1111/j.1360-0443.2008.02218.x] [Medline: 18494841]
10. Leshner AI. Addiction is a brain disease, and it matters. Addiction 2003;1(2):190-193.
11. Koob GF. The neurobiology of addiction: a neuroadaptational view relevant for diagnosis. Addiction 2006 Sep;101 Suppl 1:23-30. [doi: 10.1111/j.1360-0443.2006.01586.x] [Medline: 16930158]
12. White WL, Boyle M, Loveland D. Alcoholism/addiction as a chronic disease: From rhetoric to clinical reality. Alcoholism Treatment Quarterly 2002;20(3-4):107-129.
13. Best D, Laudet A. London: RSA. The potential of recovery capital URL: https://www.thersa.org/globalassets/pdfs/blogs/a4-recovery-capital-230710-v5.pdf [accessed 2015-04-16] [WebCite Cache ID 6Xpa8dhAr]
14. Tuchman E. Women and addiction: the importance of gender issues in substance abuse research. J Addict Dis 2010 Apr;29(2):127-138. [doi: 10.1080/10550881003684582] [Medline: 20407972]
15. Shipton D, Whyte B, Walsh D. Alcohol-related mortality in deprived UK cities: worrying trends in young women challenge recent national downward trends. Journal of Epidemiology and Community Health 2013:2574-2582.
16. Emslie C, Lewars H, Batty GD, Hunt K. Are there gender differences in levels of heavy, binge and problem drinking? Evidence from three generations in the west of Scotland. Public Health 2009 Jan;123(1):12-14 [FREE Full text] [doi: 10.1016/j.puhe.2008.06.001] [Medline: 19049841]
17. Drake RE, O'Neal EL, Wallach MA. A systematic review of psychosocial research on psychosocial interventions for people with co-occurring severe mental and substance use disorders. J Subst Abuse Treat 2008 Jan;34(1):123-138. [doi: 10.1016/j.jsat.2007.01.011] [Medline: 17574803]
18. Scottish Advisory Committee on Drug Misuse. The Scottish Government. Edinburgh: Scottish Executive; 2003. Mind the gaps: Meeting the needs of people with co-occurring substance misuse and mental health problems URL: http://www.gov.scot/resource/doc/47063/0013752.pdf [accessed 2015-04-16] [WebCite Cache ID 6XptyzIHW]
19. Elison S, Ward J, Davies G, Moody M. Implementation of computer-assisted therapy for substance misuse: a qualitative study of Breaking Free Online using Roger's diffusion of innovation theory. Drugs and Alcohol Today 2014;14(4):207-218.
20. Elison S, Ward J, Davies G, Lidbetter N, Dagley M, Hulme D. An outcomes study of eTherapy for dual diagnosis using Breaking Free Online. Advances in Dual Diagnosis 2014;7(2):52-62.
21. Elison S, Humphreys L, Ward J, Davies G. A Pilot Outcomes Evaluation for Computer Assisted Therapy for Substance Misuse- An Evaluation of Breaking Free Online. Journal of Substance Use 2013;19(4):1-6.
22. Davies G, Elison S, Ward J, Laudet A. The role of lifestyle in perpetuating substance use disorder: the Lifestyle Balance Model. Subst Abuse Treat Prev Policy 2015;10:2 [FREE Full text] [doi: 10.1186/1747-597X-10-2] [Medline: 25595205]

XSL•FO
RenderX

23. National Institute for Health and Care Excellence. London: National Institute for Health and Clinical Excellence; 2007. Drug Misuse: Psychosocial Interventions URL: http://www.nice.org.uk/guidance/cg51/resources/guidance-drug-misuse-psychosocial-interventions-pdf [accessed 2015-04-16] [WebCite Cache ID 6Xpu7Fved]
24. Beck AT, Wright FD, Newman CF, Liese BS. Cognitive therapy of substance abuse. New York: Guilford Press; 2011.
25. Marlatt GA, Bowen S, Chawla N, Witkiewitz K. Mindfulness-based relapse prevention for substance abusers: Therapist training and therapeutic relationships. In: Mindfulness and the therapeutic relationship. New York: Gilford Press; 2008:107.
26. Greenberger D, Padesky CA. Mind Over Mood: Change How You Feel by Changing the Way You Think. New York: Guilford Press; 1995.
27. Williams C, Garland A. A cognitive-behavioural therapy assessment model for use in everyday clinical practice. Advances in Psychiatric Treatment 2002;8(3):172-179.
28. Dansereau DF, Simpson DD. A Picture Is Worth a Thousand Words: The Case for Graphic Representations. Professional Psychology: Research and Practice 2009;40(1):104-110.
29. Czuchry M, Dansereau DF. A model of the effects of node-link mapping on drug abuse counseling. Addict Behav 2003 Apr;28(3):537-549. [Medline: 12628625]
30. The CC. Report on the Review of Patient-Identifiable Information. London: Department of Health; 1997. The Caldicott Committee URL: http://webarchive.nationalarchives.gov.uk/20130107105354/http:/www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4068404.pdf [accessed 2015-04-16] [WebCite Cache ID 6XpuoSepm]
31. Gollwitzer PM, Sheeran P. Implementation intentions and goal achievement: A meta‐analysis of effects and processes. Advances in Experimental Social Psychology 2006;38:69-119.
32. Carroll KM, Ball SA, Martino S, Nich C, Babuscio TA, Nuro KF, et al. Computer-assisted delivery of cognitive-behavioral therapy for addiction: a randomized trial of CBT4CBT. Am J Psychiatry 2008 Jul;165(7):881-888 [FREE Full text] [doi: 10.1176/appi.ajp.2008.07111835] [Medline: 18450927]
33. Marks IM, Cavanagh K, Gega L. Computer-aided psychotherapy: revolution or bubble? Br J Psychiatry 2007 Dec;191:471-473 [FREE Full text] [doi: 10.1192/bjp.bp.107.041152] [Medline: 18055948]
34. Skevington SM, Lotfy M, O'Connell KA. The World Health Organization's WHOQOL-BREF quality of life assessment: psychometric properties and results of the international field trial. A report from the WHOQOL group. Qual Life Res 2004 Mar;13(2):299-310. [Medline: 15085902]
35. Kroenke K, Spitzer RL, Williams JBW. The PHQ-9. Journal of General Internal Medicine 2001;16(9):606-613.
36. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med 2006 May 22;166(10):1092-1097. [doi: 10.1001/archinte.166.10.1092] [Medline: 16717171]
37. Gossop M, Darke S, Griffiths P, Hando J, Powis B, Hall W, et al. The Severity of Dependence Scale (SDS): psychometric properties of the SDS in English and Australian samples of heroin, cocaine and amphetamine users. Addiction 1995 May;90(5):607-614. [Medline: 7795497]
38. Kypri K, Langley JD, Saunders JB, Cashell-Smith ML, Herbison P. Randomized controlled trial of web-based alcohol screening and brief intervention in primary care. Arch Intern Med 2008 Mar 10;168(5):530-536. [doi: 10.1001/archinternmed.2007.109] [Medline: 18332300]
39. Guth S, Lindberg SA, Badger GJ, Thomas CS, Rose GL, Helzer JE. Brief intervention in alcohol-dependent versus nondependent individuals. J Stud Alcohol Drugs 2008 Mar;69(2):243-250. [Medline: 18299765]
40. Glenn D, Golinelli D, Rose RD, Roy-Byrne P, Stein MB, Sullivan G, et al. Who gets the most out of cognitive behavioral therapy for anxiety disorders? The role of treatment dose and patient engagement. J Consult Clin Psychol 2013 Aug;81(4):639-649 [FREE Full text] [doi: 10.1037/a0033403] [Medline: 23750465]
41. Carroll KM, Kiluk BD, Nich C, Babuscio TA, Brewer JA, Potenza MN, et al. Cognitive function and treatment response in a randomized clinical trial of computer-based training in cognitive-behavioral therapy. Subst Use Misuse 2011;46(1):23-34 [FREE Full text] [doi: 10.3109/10826084.2011.521069] [Medline: 21190403]
42. Baker R, Camosso-Stefinovic J, Gillies C, Shaw EJ, Cheater F, Flottorp S, et al. Tailored interventions to overcome identified barriers to change: effects on professional practice and health care outcomes. Cochrane Database Syst Rev 2010(3):CD005470 [FREE Full text] [doi: 10.1002/14651858.CD005470.pub2] [Medline: 20238340]
43. Moore S. Providing drug and alcohol services to female offenders: exploring gender and geographical barriers. Drugs and Alcohol Today 2011;11(3):130-136.
44. Lai DWL, Luk PKF, Andruske CL. Gender differences in caregiving: a case in Chinese Canadian caregivers. J Women Aging 2007;19(3-4):161-178. [doi: 10.1300/J074v19n03_11] [Medline: 18032259]
45. Gaunt R. Biological Essentialism, Gender Ideologies, and Role Attitudes: What Determines Parents Involvement in Child Care. Sex Roles 2006;55(7-8):523-533.
46. European Monitoring Centre for Drugsand Drug Addiction. European Drug Report 2013: Trends and developments. European Drug Report 2013: Luxembourg Publication Office of the European Union; 2013. URL: http://www.emcdda.europa.eu/publications/edr/trends-developments/2013 [accessed 2015-04-16] [WebCite Cache ID 6XpvQ9iwu]

XSL•FO
RenderX

47. Ling J, Smith KE, Wilson GB, Brierley-Jones L, Crosland A, Kaner EFS, et al. The 'other' in patterns of drinking: a qualitative study of attitudes towards alcohol use among professional, managerial and clerical workers. BMC Public Health 2012;12:892 [FREE Full text] [doi: 10.1186/1471-2458-12-892] [Medline: 23092252]
48. National Treatment Agency for Substance Misuse. 2013. Alcohol Treatment in England 2011 - 2012 URL: http://www.nta.nhs.uk/uploads/alcoholcommentary2013final.pdf [accessed 2015-04-16] [WebCite Cache ID 6Xpcvhido]
49. National Treatment Agency for Substance Misuse. Women in drug treatment: what the latest figures reveal. London: NHS; 2010. URL: http://www.nta.nhs.uk/uploads/ntawomenintreatment22march2010.pdf [accessed 2015-04-16] [WebCite Cache ID 6XpvWcyAZ]
50. Bennett T, Holloway K, Farrington D. The statistical association between drug misuse and crime: A meta-analysis. Aggression and Violent Behavior 2008;13(2):107-118.
51. Health Social Care Information Centre. Statistics on Drug Misuse: England. London: Health Social Care Information Centre; 2014. URL: http://www.hscic.gov.uk/catalogue/PUB15943/drug-misu-eng-2014-rep.pdf [accessed 2015-04-16] [WebCite Cache ID 6XpvYMkWG]
52. Fountain J, Bashford J, Underwood S. Drug use amongst Black and minority ethnic communities in the European Union and Norway. Probation Journal 2004;51(4):362-378.

Abbreviations

BFO: Breaking Free Online
CAT: computer-assisted therapy
CBT: cognitive behavioral therapy
GAD: General Anxiety Disorder Scale
LBM: Lifestyle Balance Model
NICE: National Institute for Health and Care Excellence
PHQ: Patient Health Questionnaire
RPM: Recovery Progression Measure
SDS: Severity of Dependence Scale
WHOQOL-BREF: World Health Organization Quality of Life assessment

Edited by G Eysenbach; submitted 17.02.15; peer-reviewed by L Hogan, L Dubrow-Marshall; comments to author 04.03.15; revised version received 11.03.15; accepted 14.03.15; published 23.04.15.

Please cite as:
Elison S, Davies G, Ward J
Effectiveness of Computer-Assisted Therapy for Substance Dependence Using Breaking Free Online: Subgroup Analyses of a Heterogeneous Sample of Service Users
JMIR Mental Health 2015;2(2):e13
URL: http://mental.jmir.org/2015/2/e13/
doi:10.2196/mental.4355
PMID:26543918

©Sarah Elison, Glyn Davies, Jonathan Ward. Originally published in JMIR Mental Health (http://mental.jmir.org), 23.04.2015. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

An Online, Moderated Peer-to-Peer Support Bulletin Board for Depression: User-Perceived Advantages and Disadvantages

Kathleen Margaret Griffiths[1], PhD; Julia Reynolds[1], BA (Hons), MPsych (Clin); Sara Vassallo[2], BA, BPsych

[1]National Institute for Mental Health Research (NIMHR), Research School of Population Health, Australian National University, Canberra, Australia

[2]National Institute for Mental Health Research (NIMHR, former affiliation), Research School of Population Health, The Australian National University, Canberra, Australia

Corresponding Author:
Kathleen Margaret Griffiths, PhD
National Institute for Mental Health Research (NIMHR)
Research School of Population Health
Australian National University
Eggleston Road
Acton
Canberra, 2601
Australia
Phone: 61 2 6125 9723
Fax: 61 2 6125 0733
Email: kathy.griffiths@anu.edu.au

Abstract

Background: Online, peer-to-peer support groups for depression are common on the World Wide Web and there is some evidence of their effectiveness. However, little is known about the mechanisms by which Internet support groups (ISGs) might work.

Objective: This study aimed to investigate consumer perceptions of the benefits and disadvantages of online peer-to-peer support by undertaking a content analysis of the spontaneous posts on BlueBoard, a well-established, moderated, online depression bulletin board.

Methods: The research set comprised all posts on the board (n=3645) for each of 3 months selected at 4 monthly intervals over 2011. The data were analyzed using content analysis and multiple coders.

Results: A total of 586 relevant posts were identified, 453 (77.3%) reporting advantages and 133 (22.7%) reporting disadvantages. Positive personal change (335/453, 74.0%) and valued social interactions and support (296/453, 65.3%) emerged as perceived advantages. Other identified benefits were valued opportunities to disclose/express feelings or views (29/453, 6.4%) and advantages of the BlueBoard environment (45/453, 9.9%). Disadvantages were negative personal change (50/133, 37.6%), perceived disadvantages of board rules/moderation (42/133, 31.6%), unhelpful social interactions/contact with other members (40/133, 30.1%), and technical obstacles to using the board (14/133, 10.5%).

Conclusions: Consumers value the opportunity to participate in an online mutual support group for mental health concerns. Further research is required to better understand how and if these perceived advantages translate into positive outcomes for consumers, and whether the perceived disadvantages of such boards can be addressed without compromising the safety and positive outcomes of the board.

(JMIR Mental Health 2015;2(2):e14) doi:10.2196/mental.4266

KEYWORDS

Internet; support group; mental health; depression

Introduction

Peer-to-peer Internet support groups (ISGs) are an accessible source of support and advice for health conditions. According to the Pew Internet study, 18% of Internet users have searched online for others “with health concerns similar to theirs” [1]. ISGs which enable users to communicate with their peers anonymously may be particularly attractive to those with

XSL•FO
RenderX

stigmatized conditions such as depression [2], particularly where the condition is treatment resistant.

A recent randomized controlled trial found that a depression ISG was effective relative to an attention control group in reducing clinically significant depressive symptoms over a 6-month period [3]. Little is known about the mechanisms by which such an improvement might occur. However, some insight into these processes might be provided by a consideration of consumer perceptions of the benefits and disadvantages of Internet support groups.

Several studies have provided data on consumer-perceived benefits of depression ISGs using data from quantitative surveys [4-7]. Reported advantages from these survey studies included emotional support [5,6], an outlet for expression [5], the opportunity to talk about matters that could otherwise not be discussed [4], reduced isolation [4,5], and information about medication [4]. Other documented advantages included improved symptoms [6] and increased formal help seeking [4]. In a study which undertook a factor analysis of items measuring perceived advantages of ISGs, Nimrod and colleagues [7] reported two factors: offline improvements, in the form of improved daily functioning, and online advantages, in the form of social support. However, several items cross-loaded on both factors.

A limitation of the studies used by authors to date to investigate perceived advantages of ISGs is that they have all relied on the subset of ISG users who are prepared to respond to a survey. In addition, such surveys are devised by researchers rather than based on users' reports of benefits. It is, therefore, possible that they do not encompass all user-perceived advantages of ISGs. An alternative approach to investigating the benefits or problems with ISGs is to analyze statements of benefit and disadvantage in spontaneous posts (ie, user messages) on a support forum. To our knowledge, only one published study has reported the results of such an investigation. Horgan, McCarthy, and Sweeney [8] investigated the posts of university students who posted anonymously on a researcher-moderated depression ISG. Reported benefits based on a qualitative analysis included "sharing their feelings," "a sense of not being alone," a "shared understanding," and "anonymity." However, the analysis was based on only 56 posts by 13 participants on an experimental ISG. Further, the support group was generated for the purpose of the research and was available for only two university terms. Thus, it is unclear if the findings would be applicable to a more heterogeneous target group with respect to age and background, and to a more established, open, and publicly available support group. In addition, the study did not explicitly investigate the potential disadvantages of the support group.

This study aimed to investigate the perceived advantages and disadvantages of an online ISG by analyzing posts sampled over several periods of time from a well-established, moderately large, publicly available mental health ISG [9]. This ISG was open to any person aged 18 years or older, regardless of demographic background or country of residence.

Methods

Overview

BlueBoard is a moderated, online peer-to-peer support group for mental health problems, including depression, bipolar disorder, anxiety, and borderline personality disorders [9] (see Figure 1). The majority of the posts are made on the depression forum. BlueBoard is run as a service by the National Institute for Mental Health Research with funding from the Australian Department of Health. BlueBoard moderators—known as the Mod Squad—are trained consumers overseen by an experienced registered clinical psychologist (JR). The moderators are tasked with ensuring that users interact in a respectful manner and adhere to the rules of BlueBoard. The rules do not allow members to post potentially identifying details or explicit references to suicide, self-injury, and harm to others. The moderators monitor posts and, where necessary, remove material that does not comply with the rules. The moderators do not, however, participate as members. Although BlueBoard is a formal service rather than a research intervention, all members of BlueBoard consent to the analysis of BlueBoard posts at the time of joining the board, according to an Australian National University (ANU) Human Research Ethics approved protocol. Members learn of the board primarily through online searches and links from other websites.

The data were extracted in the form of board posts and reported posts, the latter being messages sent to the board moderators, for example, when reporting a post the member perceived to be inappropriate. These data were analyzed using content analysis, multiple coders, and an inductive approach [10].

XSL•FO
RenderX

Figure 1. BlueBoard homepage.



Coding

Overview

The coding categories for positive and negative experiences on BlueBoard were developed using a multiple-step approach commencing with a set of data not used in the final coding.

Training Sets

In the first instance, three researchers (JR, SV, LB) independently identified and developed a classification system for posts in Training Set 1, which consisted of all posts for October 2010 (n=1220). The coders subsequently discussed and agreed on a common classification system for the positive and negative posts. Two of the coders (JR, LB) then recoded the posts according to the common classification system, further refining it by consensus.

Training Set 2 consisted of all posts for September 2010 (n=1295). The refined system was then tested by the two raters on approximately half of the posts in Training Set 2 (599/1295, 46.25%). Since the raters found it difficult to code the large dataset directly into categories, a two-stage coding process was developed and tested on the second half of the posts in Training Set 2 (696/1295, 53.75%). In phase 1, posts were first scored as relevant or not relevant. Discrepancies were resolved by consensus. In phase 2, the posts were allocated into categories and subcategories with final coding by consensus.

Research Set

The research set comprised all posts (n=3645) for each of 3 months selected at 4 monthly intervals over 2011. Posts not relevant to the study, such as duplicate posts and material posted by moderators and spammers, were removed (n=1928). This left 1717 posts—April (n=598), August (n=509), and December (n=610). These posts were rated for relevance by two raters (JR, SV) with discrepancies resolved by consensus. The latter were imported into NVivo 10 and rated according to the categories and subcategories developed in the training sets. The resulting coding scheme was subsequently refined at the category and subcategory level by the first author (KG) using NVivo 10. For the purposes of reporting, each post was rated according to whether it contained content relevant to each of the categories. The final coding was checked by a second rater (JR) with the few discrepancies (n=5) resolved by consensus. Quantitative data (ie, frequencies) are reported for the top-level categories.

Results

Overview

A total of 586 posts were found by rater consensus to refer to the advantages and/or disadvantages of participating in

XSL•FO
RenderX

BlueBoard. A total of 212 members contributed to the total research set and 103 to the posts rated as relevant. Of the latter, 97 provided demographic data at registration of whom the majority were women (74/97, 76%), lived in Australia (89/97, 92%), and resided in a city (77/97, 79%). Participant age was recorded in age bands. Using midpoints of the bands for the purpose of calculation, mean participant age was found to be 37.2 (range 18-19 years to 60-65 years). Of the 94 participants who provided information about their clinical status, 69 (73%) were consumers.

Findings

Of the 586 relevant posts, the majority (453, 77.3%) reported advantages of the board. The remainder were concerned with disadvantages. The themes extracted for each are reported in Table 1, are summarized in turn below, and are illustrated by quotations derived from user posts. Spelling errors in these quotations have been edited for ease of reading.

Table 1. Broad themes emerging from the content analysis of BlueBoard: advantages and disadvantages.

Broad themes	Number of posts, n (%)
Advantages (n=453)	
Positive personal change	335 (74.0)
Valued social interactions and support	296 (65.3)
Valued opportunities to disclose/express feelings or views	29 (6.4)
Other advantages of the BlueBoard environment	45 (9.9)
Disadvantages (n=133)	
Negative personal change	50 (37.6)
Perceived disadvantages of board rules/moderation	42 (31.6)
Unhelpful social interactions/contact with other members	40 (30.1)
Technical and perceived obstacles in using the board	14 (10.5)

Advantages

As noted in Table 1, four broad themes emerged from the posts about the advantages of the support group, including positive personal change, valued social interactions and support, valued opportunities to disclose/express feelings or views, and other advantages of the BlueBoard environment. Since many posts contained multiple elements and were, therefore, coded into more than one broad theme, the values in Table 1 do not sum to 100%.

Advantage 1: Positive Personal Change

Overview

One of the top two themes to emerge from members' posts related to positive personal changes associated with participating in the board. Many members reported nonspecific positive changes—"I post a lot on this site and it helps that's why I do it" [Participant #1]. In addition, specific emotional, cognitive, and behavioral changes were reported as illustrated in the following sections.

Emotional Effects

Members reported positive emotional changes in the course of using the board. Sometimes these referred to a nonspecific effect of the board—"Thanks so much, believe it or not that make me feel a lot better" [Participant #2]. However, often the emotional change reported was specific and involved a current, rather than long-term, effect.

The most common specific emotional effect was gratitude or appreciation with the posts incorporating either general—"Thanks heaps" [Participant #3]—or targeted expressions of appreciation—"Thanks guys, appreciate your kind words…" [Participant #4].

Members also expressed happiness or feeling glad:

> *Thank you for your support and caring words...brought a smile to my face :).* [Participant #1]
>
> *Glad to hear you're feeling better.* [Participant #5]

Members also derived amusement from posts on the board:

> *LOL, very funny! Thanks for sharing that :).* [Participant #6]
>
> *hahahhah hilarious!* [Participant #7]

rofl. (rolling on the floor laughing) [Participant #8]

Some participants expressed a sense of relief—"Phew, it's normal" [Participant #9]. Typically members did not discuss the effect of the board on their mood specifically, although one member noted, "I don't know if it's related, but my mood has actually been better these past few days I've been on here!" [Participant #10].

Members also reported experiencing hope as a result of hearing the stories of others:

> *I love to hear about others healing and journey, helps me keep on track and gives me hope. :)* [Participant #7]
>
> *Ever seen the movie Pay It Forward? Hope given, and hope returned... I think that's what this place is all about.* [Participant #11]
>
> *It is good to know that there is some sort of light at the end of the tunnel.* [Participant #12]

XSL•FO
RenderX

Further, the board was seen as a source of inspiration:

> *I want to thank you as well, I felt very inspired by your story.* [Participant #13]
>
> *Maybe we need our own little motto in here... inspired by you...* [Participant #12]
>
> *WOW!!!!! That is inspirational, I'm in the right head space to be able to truly believe that.* [Participant #9]

Cognitive Effects

Use of the board was also associated with self-reported change in members' thinking.

Posts on the board were seen as enhancing the member's knowledge—"Have read this forum since I was diagnosed in Jan and have found the information very informative" [Participant #14]. Through exposure to the board, members also developed the knowledge that they were not alone—"It's nice to know that I'm not alone, and that other people are going through the same thing" [Participant #10]. Further, posts were perceived as thought provoking—"Your post really made me think" [Participant #15]—and as providing insight and a different perspective:

> *Both of you have [made] me realize opening up to him is a good idea and is not weak :).* [Participant #1]
>
> *I did not think of it in this way but maybe you are right.* [Participant #16]

Supporting others was seen by some members as a way of helping themselves by thinking through their own circumstance or reminding and motivating themselves to maintain their aims:

> *I really thank you for your post on this forum as my post has been as much about thinking about why I'm feeling good, as it is providing advice to you.* [Participant #17]
>
> *Actually, helping others helps me, because it reinforces all of the things I have learned, keeps them fresh in my mind so I don't slip.* [Participant #15]

Finally, participants reported that the posts on the board stimulated the intention to act:

> *That's a really good idea...thanks for your advice, I'm going to try that the next time I'm in the situation :).* [Participant #18]
>
> *Thanks for the advice I'll try writing a journal and see how I go.* [Participant #19]

Behavioral Effects

The use of BlueBoard was also associated with behavioral changes. This included evidence of acting on a suggestion on the board. In some cases, the suggestion involved a recommendation that the member consult an information source:

> *You recommended a book called "The Happiness Trap" by Dr Russ Harris. I've just started reading it and am finding it extremely helpful and informative, thank you so much for passing on your knowledge :).* [Participant #7]

In other cases, the behavioral effect involved a help-seeking act:

> *Thanks for your help guys. I went and saw a new doctor this morning with my mum's support as I couldn't face seeing someone new on my own. He has prescribed me medication and swapped me to a different counsellor and has told me if I don't connect with him he will change me to someone else. I guess I just needed that little push to make a change as I am not good with change.* [Participant #30]

Advantage 2: Valued Social Interactions and Support

Overview

> *I don't know why I didn't join one of these discussion boards earlier, the support from everyone is amazing!* [Participant #10]

The second major theme to emerge from the analysis was that BlueBoard provided an opportunity for its users to engage in valued social interactions and to receive support. Members frequently expressed gratitude to other BlueBoard members for their responsiveness—"Thank you again for replying, you have no idea how much it means to me that someone did because I was so scared" [Participant #20]. Other posts referred to the culture of the board as one of support or mutual support:

> *You need to understand that you will have support from everyone on here.* [Participant #21]
>
> *We are all here to support each other.* [Participant #22]

Although there were generic references to support—"A big thank you to everyone who posts on BlueBoard. Your thoughts and words are valuable and appreciated" [Participant #11]—many of the posts pointed to specific types of social support, such as shared understanding, a nonjudgmental environment, advice, information, and emotional support and companionship. Each is described separately below, but we acknowledge that these concepts are interlinked, rather than mutually exclusive.

Shared Understanding

A key perceived advantage of the board was that it provided its members with the opportunity to interact with others with a shared understanding of living with mental health problems:

> *It's like a breath of fresh air to find this BlueBoard and communicate with people who truly understand and relate to what I'm going through.* [Participant #9]
>
> *You can talk openly...and the members really get what you're going through because they have been there themselves.* [Participant #23]

This shared understanding in turn was perceived as validating, reducing the sense of isolation, and enhancing a sense of belonging:

> *Depression makes you feel totally isolated and detached, and you convince yourself that no-one else could possibly understand how you feel. This site has been a blessing for me because it gives me a sense of "belonging"...I guess it makes me feel a little less like a freak, knowing there's a whole bunch of other "freaky" people out there, who are just as normal as*

XSL•FO
RenderX

> *everyone else, aside from our common illnesses. In a society that judges, and bullies, and stigmatises everything, this place is a welcome reprieve.* [Participant #11]

Nonjudgmental Environment

As illustrated in the latter and other posts, this shared understanding was also seen as fostering a nonjudgmental (ie, nonstigmatizing) and, therefore, emotionally safe space:

> *I'm so glad that you've found this forum, it's a great way to express how you feel without judgment and get support from others who know the place you're in.* [Participant #24]

> *It's a lovely safe haven where no one judges you.* [Participant #11]

Emotional and Companionship Support

In addition to the support arising from the validation and sense of belonging associated with a shared understanding of depression/mental health problems, the analysis of posts indicated that participants valued a number of other forms of emotional and/or companionship support on the board, including kindness, caring, comfort, warmth, understanding, encouragement, self-esteem support, and friendship (see Table 2).

Table 2. Examples of emotional and companionship support reported on BlueBoard.

Emotional and companionship support	Examples
Kindness	"You're another special presence on this board. A silent sufferer who has so many kind and loving words for others" [Participant #11].
Caring	"Just as [BlueBoard member] above said, we care what happens to you, even though it's virtual friendship doesn't mean we don't care because we do" [Participant #8].
Comfort	"Thanks for that - yes I did read (and take comfort) from the other post you mentioned" [Participant #25]. "oooo nice cuddle, thanks :)" [Participant #7].
Warmth	"I love that when there is a new member intro, [Member 1] and [Member 2] are usually the first to welcome them in the warmest way!" [Participant #7].
Understanding	"Welcome [new BlueBoard member], you'll find lots of support understanding and advice on offer here on BB. Good on you for joining, not always an easy thing to do, but you've done it" [Participant #8].
Encouragement	"You have no idea how much encouragement you've given me, as a sufferer of this stinking illness" [Participant #11].
Self-esteem support	"Dear [lists 7 BlueBoard members] and all who read this post, I thank you for all the compliments and or feedback" [Participant #26].
Friendship	"You will find virtual friendships of all kinds here, you sound like you have plenty of experience to draw from. You will find plenty of us with similar stories of isolating ourselves for all sorts of reasons" [Participant #8].

Advice, Wisdom, and Informational Support

A frequently cited specific form of social support was advice from other members:

> *That's a really good idea...thanks for your advice, I'm going to try that the next time I'm in the situation :).* [Participant #18]

> *Your advice is always so good... :o.* [Participant #27]

A minority of these posts specifically concerned advice designed to promote help seeking from a professional—"Thanks for the tips about seeing a counsellor" [Participant #28]—from another service—"Thanks also for the idea of disability employment organisations and a case worker:)" [Participant #13]—related to the provision of information—"Thanks for letting us know of the program, I didn't manage to catch the insight program but will see if I can access it online" [Participant #29]—or by sharing wisdom—"Thank you so much for the wisdom that you have shared, I really appreciate it.:)" [Participant #9].

Advantage 3: Valued Opportunities to Disclose/Express Feelings or Views

The board was seen as a safe place for members to express their feelings or views. In particular, a number of members referred to the importance of the board as a place to vent—"It's amazing how it helps to vent, just getting things said and off your chest can make the burden seem just that much lighter" [Participant #8]. This was not necessarily in expectation that others would respond or solve the problem:

> *I don't want words of sympathy or support or comfort. I just need to say this.* [Participant #30]

> *I'm not really expecting anyone to reply...I just really need to get some things out otherwise I feel like it will completely take over me.* [Participant #19]

Posting was seen as an alternative to ruminating on the problem—"I guess I'm just sick of not talking about it and being stuck inside my head" [Participant #17]—and as a place where members would not be judged for their words—"I personally have found this forum to be good to express how I feel at any given time [without] ridicule or judgment" [Participant #9]. The board also provided a means of venting safely without burdening

XSL•FO
RenderX

those closest to them—"I need a place to vent so my poor husband doesn't have to listen to my constant whining and misery I just need someone to hear me and not call me crazy!" [Participant #31].

Advantage 4: Other Advantages of the Board Environment

Other posts related to the advantages of the BlueBoard environment and its accessibility. The board was seen as providing a means for people to reach out—"This is a good site. I too never thought of reaching out on something like this but its been fantastic" [Participant #32]. Some members saw the board as an accessible alternative when face-to-face or therapist support was unavailable:

> *We are here for you so you can always count on us. I know that it is not the same as seeing someone face to face and particularly your psychologist but it is better than having no support at all.* [Participant #16]
>
> *I was supposed to be seeing my psychologist this afternoon but now I can't so I thought I'd talk to you instead lol :rolleyes:.* [Participant #33]
>
> *The psychologist forgot to turn up today, a day when I could really do with seeing him...hence why I am on here.* [Participant #34]

Some felt the board had advantages over face-to-face interactions, providing a safer and less threatening and constraining medium to communicate about their illness:

> *No forums aren't the same as talking face-to-face, but sometimes it's easier to just simply be yourself and open up without having to worry about being judged face to face.* [Participant #35]
>
> *I don't exactly tell people about this face-to-face but that is what this forum is for. And we certainly won't judge you.* [Participant #15]
>
> *I am much more talkative and eloquent in here than in real life.* [Participant #36]
>
> *Safer to vent here than actually act on it in the real world.* [Participant #15]

The Internet support group was also seen by some as a preferable format to social media sites such as Facebook:

> *[A social network]...it's not that comfortable and I find it difficult to write what I'm really thinking, feeling...* [Participant #13]
>
> *I'd rather listen to all your honest voices than the not so honest perceptions I often find on the dreaded Facebook.....lol.* [Participant #37]

One member who had been refused admission to another online website noted of the board, "So it's nice to know there are sites out there that are willing to accept anybody and listen" [Participant #20].

Some members referred to the special qualities of other members—"It's very nice to have places like this, where people are genuine, honest and understanding. I think having this condition is very leveling and makes superficial, ego boosting conversation tiresome, just doesn't feel good" [Participant #13].

The availability of the posts as a record was seen as an asset. For example, members noted that lessons could be learned from consulting both one's own and others' past posts:

> *Before you make any rash decisions about going off your meds...do yourself a favour and read through your old posts...just to remind yourself what happened last time.* [Participant #11]
>
> *All I can suggest is have a read through some of the posts on this site. It may give you a little bit more of an insight into what we go through on a daily basis and may help you to understand a little bit more about depression.* [Participant #32]

The board was seen by one member as a potential facilitator of communication with a health care professional—"Perhaps you could print out your posts and give them to your GP to read if you have trouble talking" [Participant #8].

Another advantage of the format of the board was that even when people were not well enough to participate fully on the board, they were able to follow the progress of others by monitoring the board—"Haven't been too far away...always keeping an eye on things and following your progress...just wasn't up to participating for a while" [Participant #11].

Finally, members expressed appreciation for the role of the board moderators, known as the Mod Squad:

> *Thanks for...the time and effort you put in to keep this board going, so that we have somewhere to come and share and support one another. Much appreciated.* [Participant #11]
>
> *I think the mods will continue to keep up the good work in keeping things, well, moderate here.* [Participant #8]

Disadvantages

Overview

A minority of relevant posts (133/586, 22.7%) were concerned with the disadvantages of the board. Four themes emerged: negative personal change, perceived disadvantages of board rules/moderation, unhelpful social interactions/contact with other members, and technical and perceived obstacles in using the board (see Table 1).

Disadvantage 1: Negative Personal Change

There is potential for the descriptions of negative experiences to cause distress to those reading them. A number of members indicated that they were sorry, sad, or worried to learn of another member's problems:

> *It's really upsetting to hear how you are struggling with it all...as people on this site we know it's hard.* [Participant #1]
>
> *I'm really worried about you.* [Participant #38]

It is likely that many of these posts were expressions of empathy and reassurance, rather than deep personal emotional distress—"I'm so sorry to hear things aren't improving" [Participant #11]. On the other hand, one member was clearly distressed by content posted on the board, content that was

XSL•FO
RenderX

against the BlueBoard rules. The member read the material before it was removed by moderators and experienced the post as "triggering" and asked the other member to "Stop writing about your overdoses!!!!!! I can't handle it!!!!!!" [Participant #39].

Some members expressed concerns about how their posts might be received by, or have affected, others:

> *By the way I'm having a really crappy day today so I hope I don't sound cross, I'm not.* [Participant #15]
>
> *I was worried I would be hammered by writing this :P.* [Participant #40]
>
> *Oh sorry, I obviously got confused! Sorry for any offense.* [Participant #7]
>
> *I worry a lot about being accepted (I'm even worried about posting this).* [Participant #41]

One member who used humor to cope with their depression wrote, "BTW, if anyone has a problem with the manner in which I project my depression, please do let me know, I do not wish to upset anyone!" [Participant #42].

Finally, some members expressed frustration and irritation with the BlueBoard rules—"So Mr/Mrs Moderator. I am very annoyed" [Participant #30] (see theme in the next section).

Disadvantage 2: Perceived Disadvantages of Board Rules/Moderation

For reasons of safety, BlueBoard is governed by strict rules and careful moderation according to preestablished protocols. This was seen as unnecessarily restrictive by some members—"It may take you a while to get used to the rules of this website. I have had substantial trouble given that as a creative person I tend to dislike rules and be a bit deviant at times" [Participant #16].

For example, some participants perceived the BlueBoard rule that members should not discuss self-harm or suicide as a limitation of the board—"I understand these sites have a certain 'duty of care.' Having said that, I read an article in the paper yesterday which made this comment: 'It has been found that talking about suicide does not cause it to happen, and not talking about it does not prevent it'" [Participant #11]. Moreover, a strongly felt need to speak about suicide or suicidal thoughts led to attempts by some members to circumvent the rules by expressing their thoughts indirectly—"Dark thoughts which I know we're not supposed to talk about" [Participant #43].

Other members were frustrated when sections of their posts that did not conform to the rules were edited by the moderators—"What's the point of trying to express myself if you're going to cut bits of my expression out?" [Participant #30]—or removed—"Kinda spewing. My other whole post got wiped out" [Participant #36]. Perceptions that participants were treated unequally were also an occasional source of complaint—"Why is this still here and my post deleted? It's not fair" [Participant #39]. A small number of members knowingly broke the rules to make a post that they hoped would be read by other members before the moderators removed it, which resulted in placing the members' accounts on delayed "telecast" (ie, premoderation).

To ensure members are not identifiable, BlueBoard requires registration with an alias that is not the name of any person since the moderators have no means to determine that a real name is not the name of the member. However, this rule confused some members:

> *You've disabled my account under the name [deleted]. I assure you this is a random name I saw in a book once and I like it. It isn't even close to my real name. Could you please reinstate it?* [Participant #44]

Other members were disappointed or annoyed that they were unable to easily contact the moderators, or alternatively felt that it was "Pointless trying to talk to a faceless moderator" [Participant #30].

Finally, some members felt the asynchronous nature of the board was a limitation. BlueBoard does not have a real-time chat facility due to the challenge and resources that would be required to moderate it. In addition, for the safety of the members, the board does not permit posting of links to other providers. However, some felt that BlueBoard should introduce a chat room or refer members to appropriate external providers:

> *There are very few safe places online for people like us to meet, and while BlueBoard brings us together, it's frustrating not being able to converse with other members. Surely as adults we have the right to contact other like-minded folk, and can take responsibility for the outcome of those choices without holding BlueBoard to ransom. A disclaimer on the BlueBoard site would surely cover that? Or am I being naive?"* [Participant #11]

Disadvantage 3: Unhelpful Social Interactions/Contact With Other Members

Although all members on BlueBoard register anonymously, one member reported that they believed they were being stalked on the board by a person from their offline life with whom they were in conflict. This accounted for 26 posts (65%) of a total of 40 in this theme. Unsure if the alleged stalker was restricting themselves to reading, as opposed to contributing, BlueBoard posts, the member wrote, "I hate to think this but one of you who even post to me may be the stalker" [Participant #45]. The member concluded, "I will never post again, as I don't need [them] knowing about my life. I actually felt safe here" [Participant #45].

Other concerns, each of which was identified in 1 to 3 posts, were negative debate, unanswered posts, disputed information or advice, and misinterpreted communications. Finally, a small number of members felt that they did not "belong" on BlueBoard due to their current mental health status, age, or other attributes. For example, one member who had recovered was concerned about the impact their happiness might have on others on the board, and another member was unable to identify with others on the board—"I doubt anyone walks my footsteps. I've read through the first page of posts here and I seem to be the most violent person here" [Participant #30].

XSL•FO
RenderX

Disadvantage 4: Technical and Perceived Obstacles in Using the Board

A small number of posts referred to difficulties in using the board, including not knowing how to post to the board, losing a post whilst in the process of composing it because they exceeded the board's automated time-out period, and prematurely sending a message before it was completed. Other perceived obstacles to board use each noted in 2 posts included difficulty in communicating emotional nuances online and difficulty in writing posts when depressed or anxious.

Discussion

There were both perceived advantages and perceived disadvantages of the board environment. These are discussed in turn in the following sections.

Advantages

Based on the posts by members, the board provided valued social interactions and support, including shared understanding, a nonjudgmental environment, advice, and informational, emotional, and companionship support. The latter included kindness, caring, comfort, warmth, understanding, encouragement, self-esteem support, and friendship. The board was also associated with positive personal change, including specific emotional, cognitive, and behavioral effects. Further, it enabled participants to express or vent their feelings. Other advantages of the board environment were that it provided a place for members to reach out, was an accessible alternative to face-to-face support, particularly when the latter was unavailable, was preferable to social media sites, and had members with special qualities. Further, the availability of the posts as a reference or for facilitating communication with the members' doctors were seen as strengths of the board.

To our knowledge, these findings represent the first comprehensive evaluation of the perceived benefits of a publicly available, depression-related support group based on members' posts. The results are consistent with, but considerably extend, the findings of the small pilot study of student posts on a closed experimental board reported by Horgan et al [8]. The latter reported some themes that were categorized in the current study as valued social interactions and support. However, the paper did not explicitly discuss other strengths, such as a nonjudgmental environment, advice, or informational, emotional, or companionship support. Nor did it identify personal change or most strengths of the board environment. It is possible that this reflects the limited duration and number of posts on the student board and, hence, a lack of a strong board culture or trusted social network. Horgan et al [8] did, however, report that anonymity was a valued attribute of the student board. This factor did not emerge in this study. It is possible that such anonymity is more highly valued by younger than older people, or by members who share a physical environment where they may be known by, or physically encounter, other members.

The advantages of the board documented in this study point to possible mechanisms for observed improvement in depression outcomes associated with an Internet support group [3]. For example, improvements in depression might be mediated by emotional support or advice, or information provided by participants which facilitate coping or help seeking. This is consistent with findings from an unpublished quality assurance survey of BlueBoard members, half of whom indicated that they were more likely to access other forms of help, such as consulting a doctor or therapist, as a result of using BlueBoard.

Our own previous research has documented an increase in perceived emotional and informational social support following the use of an online support group [11]. However, it also found an increase in perceived social support of similar magnitude among participants in the control condition [11]. It is possible that although members value social support, it does not mediate improved depression outcomes. Alternatively, the scale may have failed to measure elements of online social support that might be critical to improved mood. For example, the scale did not explicitly measure shared understanding or access to nonjudgmental support. Future outcome research on ISGs should employ items and scales that measure attributes which users consider to be positive aspects of the board and investigate their role in mediating change in mental health outcomes.

Many of the elements identified in this study are consistent with the therapeutic factors outlined by Yalom and Leszcz [12] to explain the positive effects of group psychotherapy. These include the installation of hope, universality, imparting information, altruism, the development of social skills, interpersonal learning, group cohesiveness, and catharsis. ISGs such as BlueBoard may also be conceptualized as supporting processes critical to broader mental health recovery. For example, a recent systematic review identified five key processes in recovery: connectedness, hope/optimism, identity, meaning, and empowerment [13]. The advantages of participation spontaneously expressed by BlueBoard members are highly consistent with these processes, particularly connectedness, hope/optimism, and identity. Future research should incorporate measures relevant to Yalom and Leszcz's [12] therapeutic factors and to recovery outcomes in addition to measures of specific symptoms such as mood.

Disadvantages

Members also identified some potential disadvantages associated with the board. This included negative emotional changes associated with concern for another member of the board, a member's concern about the possible impact of their posts on others, and frustration at the board rules. Some members perceived disadvantages of the board environment, including the restriction imposed by the board rules and the effects of moderation (eg, that posts on the topic of self-harm and suicide were not permitted, nonconforming posts were edited, participants were required to register using an alias that was not a name, and that the format of the board did not include a synchronous service). A small number technical obstacles were also encountered by members using the board. Finally, there were some instances of unhelpful social interactions/contact with other members, including a report of alleged stalking on the board. Other unhelpful interactions did not account for a large number of posts.

To our knowledge, this is the first systematic analysis of the disadvantages of a publicly available, depression-related support

XSL•FO
RenderX

group based on members' posts. The advantages substantially outweighed the disadvantages of the board. However, by definition the posts on the board would have been weighted toward the perspectives of members who remained active on the board and, thus, might overestimate the advantages relative to the disadvantages of the ISG. Regardless of the magnitude of the problem, the identified problems do raise important questions.

The board enforces strong rules with a view to ensuring the safety of its members. The fact that very few members complained of unsafe or negative interactions with others on the board may reflect the benefits of such rules. However, little is known empirically about the impact of such rules on members or their mental health. Research to investigate such questions would raise significant ethical considerations. For example, although it would be of academic and practical interest to compare the effects of a moderated board with those of an unmoderated board, the ethics of undertaking a trial, were it to entail randomization, would require careful consideration. A study might be undertaken with the cooperation of the owner of an unmoderated board to compare the effects of a moderated and an unmoderated board. However, the meaningfulness of such a comparison would be limited by the methodology employed.

As a first step, there may be value in analyzing the posts on an unmoderated board of a similar size to investigate the perceived advantages and disadvantages of this type of board, and to compare them with those documented here. Unfortunately, it is unlikely that registrants on an unmoderated board will have provided prior ethics releases for research on their posts and it would not be feasible to obtain retrospective permissions for all past participants. Accordingly, such research raises moral and ethical issues even in a publicly available board. If these concerns can be addressed, there may be considerable benefits to undertaking such a comparative study.

Limitations

For pragmatic reasons, this study analyzed only a small subset of the messages posted on the Internet support group over a relatively short time frame. The primary themes emerged strongly, suggesting that the sample was adequate for the time period targeted. However, themes or their relative importance may change over time as a support group matures. Therefore, this data may not be applicable to all the developmental phases of an online support group. Further, as noted above, the posts of those who have left the board—both those who were dissatisfied with the board and those who have recovered—may be underrepresented in the data. Finally, the analysis focused on one ISG only, with a preponderance of members from one country—it is unclear if the findings will generalize to other online depression support groups comprising citizens from other countries or cultures.

Conclusions

Consumers value the opportunity to participate in an online mutual support group for depression. Further research is required to better understand how and if these perceived advantages translate into positive outcomes for consumers, and whether the perceived disadvantages of such boards can be addressed without compromising the safety and positive outcomes of the board.

Acknowledgments

Professor Kathleen Griffiths is supported by National Health and Medical Research Council (NHMRC) Fellowship 1059620.

Authors' Contributions

KG conceived the study. KG, JR, and SV designed the study. JR, SV, and KG undertook the thematic analysis and rated the posts. JR extracted the demographic data. KG drafted the paper. JR and SV critically reviewed the paper. Other contributors to the paper included Dr Lisa Barney who provided methodological advice and engaged in coding the papers, Ms Jennie Walker who provided editorial assistance, and Ms Kylie Bennett and Mr Anthony Bennett (AB) who manage the IT and related aspects of the board. AB downloaded the BlueBoard posts in a form that could be analyzed by the authors.

Conflicts of Interest

Professor Kathleen Griffiths established BlueBoard, which is the subject of the research. She is Director of e-hub Mental Health services and the National Institute for Mental Health Research, which delivers the board to the public. Julia Reynolds is responsible for the day-to-day management and decision making for BlueBoard and is a member of the e-hub management group. She and Professor Griffiths are employed by the ANU, however, neither they nor Sara Vassallo receive any additional personal financial benefit from BlueBoard.

References

1. Fox S. Pew Research Center. Washington, DC: Pew Internet & American Life Project; 2011 Feb 28. Peer-to-peer healthcare URL: http://www.pewinternet.org/files/old-media//Files/Reports/2011/Pew_P2PHealthcare_2011.pdf [accessed 2015-01-16] [WebCite Cache ID 6VcBYNtxD]
2. Davison KP, Pennebaker JW, Dickerson SS. Who talks? The social psychology of illness support groups. Am Psychol 2000 Feb;55(2):205-217. [Medline: 10717968]

XSL•FO
RenderX

3. Griffiths KM, Mackinnon AJ, Crisp DA, Christensen H, Bennett K, Farrer L. The effectiveness of an online support group for members of the community with depression: a randomised controlled trial. PLoS One 2012;7(12):e53244 [FREE Full text] [doi: 10.1371/journal.pone.0053244] [Medline: 23285271]
4. Powell J, McCarthy N, Eysenbach G. Cross-sectional survey of users of Internet depression communities. BMC Psychiatry 2003 Dec 10;3:19 [FREE Full text] [doi: 10.1186/1471-244X-3-19] [Medline: 14664725]
5. Alexander SC, Peterson JL, Hollingshead AB. Help is at your keyboard: Support groups on the Internet. In: Frey LR, editor. Group Communication in Context: Studies of Bona Fide Groups. 2nd edition. Mahwah, NJ: Lawrence Erlbaum Associates; 2003:309-334.
6. Houston TK, Cooper LA, Ford DE. Internet support groups for depression: a 1-year prospective cohort study. Am J Psychiatry 2002 Dec;159(12):2062-2068. [Medline: 12450957]
7. Nimrod G. Online depression communities: members' interests and perceived benefits. Health Commun 2013;28(5):425-434. [doi: 10.1080/10410236.2012.691068] [Medline: 22809441]
8. Horgan A, McCarthy G, Sweeney J. An evaluation of an online peer support forum for university students with depressive symptoms. Arch Psychiatr Nurs 2013 Apr;27(2):84-89. [doi: 10.1016/j.apnu.2012.12.005] [Medline: 23540518]
9. Australian National University. Canberra, Australia: National Institute for Mental Health Research - The Australian National University; 2015. BlueBoard URL: https://blueboard.anu.edu.au/ [accessed 2015-01-16] [WebCite Cache ID 6Vc8RFCrp]
10. Pistrang N, Barker C. Varieties of qualitative research: A pragmatic approach to selecting methods. In: Cooper H, Camic PM, Long DL, Panter AT, Rindskopf D, Sher KJ, editors. APA Handbook of Research Methods in Psychology, Volume 2. Research Designs: Quantitative, Qualitative, Neuropsychological, and Biological. Washington, DC: American Psychological Association; 2012:5-18.
11. Crisp D, Griffiths K, Mackinnon A, Bennett K, Christensen H. An online intervention for reducing depressive symptoms: secondary benefits for self-esteem, empowerment and quality of life. Psychiatry Res 2014 Apr 30;216(1):60-66. [doi: 10.1016/j.psychres.2014.01.041] [Medline: 24534125]
12. Yalom ID, Leszcz M. The Theory and Practice of Group Psychotherapy. 5th edition. New York, NY: Basic Books; 2005.
13. Leamy M, Bird V, Le Boutillier C, Williams J, Slade M. Conceptual framework for personal recovery in mental health: systematic review and narrative synthesis. Br J Psychiatry 2011 Dec;199(6):445-452 [FREE Full text] [doi: 10.1192/bjp.bp.110.083733] [Medline: 22130746]

Abbreviations

ANU: Australian National University
ISG: Internet support group
NHMRC: National Health and Medical Research Council

Edited by G Eysenbach; submitted 20.01.15; peer-reviewed by C Barker; comments to author 06.02.15; revised version received 09.03.15; accepted 23.03.15; published 24.04.15.

Please cite as:
Griffiths KM, Reynolds J, Vassallo S
An Online, Moderated Peer-to-Peer Support Bulletin Board for Depression: User-Perceived Advantages and Disadvantages
JMIR Mental Health 2015;2(2):e14
URL: http://mental.jmir.org/2015/2/e14/
doi:10.2196/mental.4266
PMID:26543919

©Kathleen Margaret Griffiths, Julia Reynolds, Sara Vassallo. Originally published in JMIR Mental Health (http://mental.jmir.org), 24.04.2015. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Viewpoint

Adjustment Disorders Are Uniquely Suited for eHealth Interventions: Concept and Case Study

Andreas Maercker[1], MD, PhD; Rahel C Bachem[1], MSc; Louisa Lorenz[1], BSc; Christian T Moser[1], BSc; Thomas Berger[2], PhD

[1]Division of Psychopathology and Clinical Intervention, Department of Psychology, University of Zurich, Zurich, Switzerland

[2]Institute of Psychology, Department of Clinical Psychology and Psychotherapy, University of Berne, Berne, Switzerland

Corresponding Author:
Andreas Maercker, MD, PhD
Division of Psychopathology and Clinical Intervention
Department of Psychology
University of Zurich
Institute of Psychology
Binzmühlestrasse 14/17
Zurich, 8050
Switzerland
Phone: 41 44 635 7310
Fax: 41 44 635 73
Email: maercker@psychologie.uzh.ch

Abstract

Background: Adjustment disorders (also known as mental distress in response to a stressor) are among the most frequently diagnosed mental disorders in psychiatry and clinical psychology worldwide. They are also commonly diagnosed in clients engaging in deliberate self-harm and in those consulting general practitioners. However, their reputation in research-oriented mental health remains weak since they are largely underresearched. This may change when the International Statistical Classification of Diseases-11 (ICD-11) by the World Health Organization is introduced, including a new conceptualization of adjustment disorders as a stress-response disorder with positively defined core symptoms.

Objective: This paper provides an overview of evidence-based interventions for adjustment disorders.

Methods: We reviewed the new ICD-11 concept of adjustment disorder and discuss the the rationale and case study of an unguided self-help protocol for burglary victims with adjustment disorder, and its possible implementation as an eHealth intervention.

Results: Overall, the treatment with the self-help manual reduced symptoms of adjustment disorder, namely preoccupation and failure to adapt, as well as symptoms of depression, anxiety, and stress.

Conclusions: E-mental health options are considered uniquely suited for offering early intervention after the experiences of stressful life events that potentially trigger adjustment disorders.

(JMIR Mental Health 2015;2(2):e15) doi:10.2196/mental.4157

KEYWORDS

adjustment disorders; intervention; e-mental health; unguided self-help; depression

Introduction

E-Mental Health Interventions and Adjustment Disorder

E-mental health interventions have made significant progress in the treatment of common mental disorders. For all major groups of mental disorders, effective e-interventions are available, for example [1,2]. There is one diagnostic group, however, that received little attention in e-mental health, the adjustment disorders (also known as mental distress in response to a stressor). Studies scarcely included the adjustment disorder diagnosis in e-mental health or telemedicine contexts [3-5]. In population-based epidemiological studies, prevalence rates of adjustment disorder range from 0.5%-2% [6]. Comprehensive mental service utilization studies indicate that high proportions

XSL•FO
RenderX

of clients have an adjustment disorder diagnosis, with up to 30% in general practice or consultant-liaison services [7]. Experts agree that the previous definition of adjustment disorders in the psychiatric classification systems is still problematic for various reasons [8]. However, the two most recent revisions of the Diagnostic and Statistical Manual of Mental Disorders, 5th Version (DSM-5) and the Beta-Version of the International Classification of Diseases 11 (ICD-11) provide several improvements with respect to the previous issues [7]. Consequently, it seems reasonable for the field of e-mental health to consider its potential contributions to the progress and satisfaction of clinical needs in the area of adjustment disorders.

The current paper is organized in two main parts: first, it provides an overview about general aspects of adjustment disorders, its distinction from other disorders, current diagnostic and treatment approaches, as well as the current empirical status; and second, it presents the outline of a new structured cognitive behavior therapy (CBT) program for treating adjustment disorders (sample case, burglary victims) and illustrates this by intervention results of a single case.

Distinctiveness From Other Disorders

An adjustment disorder is defined as a transient maladaptive or pathological reaction to identifiable stressors or changes in life circumstances with symptoms emerging within three months of the onset of the stressor, according to DSM-5 and ICD-10. Adjustment disorders are triggered by serious, but nontraumatic, stressors, which usually are severe life events. The stressors can either be acute (eg, loss of work, break-up of a romantic relationship, trouble with a family member), chronic (eg, financial burden, family or occupational problems), recurring (eg, seasonal economic slumps), or continuous (eg, living in a criminal environment) [6]. Adjustment disorder symptom criteria include various types of behavioral symptoms, which are manifested far beyond the expected magnitude when confronted with such a burdensome event. In addition, the disorder leads to significant social, occupational, or academic performance-related impairments.

The distinction between adjustment disorder and a normal stress response is based on the severity and the duration of symptoms or impairments. The criterion of impact on personal functioning takes into account the nature of the stressor, the personal and interpersonal context in which it has occurred, and cultural norms with regard to such responses. A recent study in preparation of the new classification system ICD-11 shows sufficient discriminant validity for distinguishing nondisorders (ie, normal stress response) from adjustment disorder with more than 2500 practitioners across several countries [9].

In contrast, posttraumatic stress disorder (PTSD) and acute stress reaction require the presence of a stressor of a magnitude that would be traumatic for almost everybody, as well as a specific symptom constellation [10]. Since not everybody exposed to traumatic events responds by developing PTSD, but may nevertheless develop significant symptoms and/or functional impairment, adjustment disorder should be considered as an alternative diagnosis [7,11].

Distinguishing adjustment disorder from a depressive episode (DE) may be particularly challenging. The appearance of a DE is often linked to a precipitating severe life event. If all diagnostic requirements of a DE are fulfilled (low mood, loss of interests and energy for at least two weeks), often along with a known history of a previous depressive disorder, the diagnosis of depression is appropriate. Adjustment disorder differs from DE by the fact that it remits when the stressor is removed or a new level of adaptation is reached. Doherty et al [12] empirically distinguished adjustment disorder and DE in a liaison psychiatry sample by means of a precipitating events list, the Beck Depression Inventory (BDI), and a social support scale. They identified 5 criteria distinguishing between both diagnoses with excellent predictive values: (1) a relationship breakdown event (for adjustment disorder), (2) high BDI (for DE), versus (3) low BDI (for adjustment disorder), (4) higher self-blame (for adjustment disorder), and (5) higher self-reported social support (for adjustment disorder). Doherty et al [12] published a similar study, additionally including a personality measure. Fewer depressive symptoms, fewer problems with social contacts, and a less pronounced disposition toward perfectionism were predictors of adjustment disorder. It is expected more studies will soon appear in this area cf. [7].

Recent Diagnostic Approaches, Diagnostic and Statistical Manual of Mental Disorders, 5th Version and International Classification of Diseases 11

Until now adjustment disorder symptom criteria are rather ill defined, for example, including various kinds of behavioral symptoms, which are manifested far beyond the expected magnitude when confronted with such a burdensome event. The lack of explicitly defined or pathognomonic symptoms had been widely criticized since it impedes empirical research [8].

To overcome these general deficits, it was decided to enhance the disorder's characterization in the two common classification systems for mental disorders, DSM and ICD, by defining adjustment disorder as a stress-response disorder. This relates adjustment disorder to a useful biological context within the framework proposed by Selye [13] and to the key role of the hypothalamic-pituitary-adrenocortical (HPA) system in the human stress response. Such work has been updated by a current, more sophisticated understanding of the neurocircuitry and the psychobiological systems that mediate and moderate this response [14]. This understanding also provides a perspective for the development of therapeutic approaches that effectively produce clinical remission in PTSD, acute stress disorder, or prolonged grief disorder, and thus may inform new therapeutic innovations for adjustment disorder.

The current DSM-5 describes five criteria for the diagnosis of an adjustment disorder: (1) emotional or behavioral symptoms arising within 3 months of exposure to an identifiable stressor; (2) the symptoms must be clinically significant or exceeding what is expected in response to the stressor and/or existing significant impairment in social or occupational functioning; (3) the stress-related disturbance is not due to another mental disorder or merely exacerbating a preexistent mental disorder; (4) the symptoms do not represent normal bereavement; and (5) the symptoms do not persist for more than 6 months once

XSL•FO
RenderX

the stressor or its consequence have terminated [15]. However, criticism has been raised regarding the failure of DSM-5 to provide specific diagnostic criteria and to guide clinicians in distinguishing problematic responses from normal adaptive reactions to stress [12].

The Beta-Version of ICD-11 introduces two core symptom groups for adjustment disorder, preoccupations with the stressor and failure to adapt [10]. Consistent with DSM-5, these symptoms have to arise within 3 months of exposure to an identifiable stressor or multiple stressors. The symptoms typically resolve within 6 months, unless the stressor persists for a longer period. Examples include divorce, illness or disability, socioeconomic problems, and conflicts at home or at work. The reaction to the stressor is characterized by preoccupation with the stressor or its consequences, including excessive worry, recurrent and distressing thoughts about the stressor, or constant rumination about its implications. Failure to adapt to the stressor causes significant impairment in personal, family, social, educational, occupational, or other important areas of functioning. If functioning is maintained only through significant additional effort, or is significantly impaired compared to the individual's prior functioning or what would be expected, he or she would be considered impaired due to the disturbance.

A previous version of the ICD-11 proposal had been empirically validated in different samples in epidemiological, etiological, and, initially, in treatment research [6]. Direct comparisons between the DSM-5 and ICD-11 Beta approach are still lacking. Nevertheless, both new classificatory approaches will stimulate further innovations and investigations for this underresearched area.

Approaches to Adjustment Disorder Treatment

Despite high prevalence rates and the fact that adjustment disorder patients suffer from a significant decrease in quality of life and an increased risk of suicidal behavior [16], only few specific treatment approaches are available to date. This may partly be due to the paucity of explicit theoretical models of adjustment disorder on which intervention programs could be based. Additionally, the indistinctive diagnostic criteria of ICD-10 and DSM-IV may account for the lack of empirically validated interventions. However, two theoretical models are suitable to explain the development of an adjustment disorder.

In the “crisis model”, Caplan [17] postulated typical trajectories that occur after extreme stress and that destabilize the individual. A personal crisis is defined as a problem that is unsolvable for the individual concerned. Caplan proposed that psychopathological symptoms develop when an individual has insufficient or inflexible defense mechanisms to handle the problem. Such personal crises are comparable to critical life events that might cause adjustment disorder. A second useful model for psychotherapy of adjustment disorder is the stress-response model by Horowitz [18]. Within this model, adjustment disorder, along with PTSD and complicated grief disorder, were conceptualized as stress response syndrome for the first time to our knowledge [19]. The model proposes that a severe stressor evokes intensive negative emotions, which in a next step are reacted upon through suppression, avoidance, or dysfunctional behavior. Consequently, a vicious circle of intrusive memories and avoidance responses arises in the individual. This discontinues as soon as a phase of working-through takes place within a person and cognitive processing enables a decrease of preoccupation and maladaptive behaviors.

Regarding treatment, several authors recommended a modular approach including therapeutic elements from a variety of approaches specific for other disorders [20,21]. An individual etiological model, the analysis of competences and resources, and the integration of the critical life event are considered important parts in adjustment disorder therapy. Maercker [22] suggests to put a special focus on techniques adapted from PTSD treatment, such as exposition (eg, imaginative or narrative exposition, writing assignments), cognitive restructuring (eg, for recurrent distressing thoughts about the stressor, blaming oneself or others), and eye movement desensitization and reprocessing (EMDR).

Empirical Evaluations

A limited amount of controlled clinical trials have been conducted in order to evaluate the different treatment approaches for adjustment disorder in a face-to-face context. They comprise a wide variety of therapeutic techniques depending on the clinical focus of the authors and range from problem solving training, grief work, and anxiety coping strategies to client-centered psychotherapy.

The first study, to our knowledge, to apply techniques from the treatment of PTSD was conducted by Cvetek [23], who introduced EMDR as a technique for reducing anxiety due to intrusions of stressful memories. Three hours of EMDR were compared with a placebo control condition (active listening for three hours) and a wait list condition. The EMDR group showed significantly lower anxiety scores and less intrusive and avoidance symptoms compared to the control groups.

In a randomized controlled design, Van der Klink et al [24] evaluated an adjustment disorder intervention aiming at the development of problem solving strategies for daily life problems including elements of time management, stress inoculation, and cognitive restructuring. Compared to treatment as usual, the intervention led to shorter sickness leave and lower recurrence rates, but showed no difference in psychopathological symptoms.

Taking a different theoretical approach, Altenhöfer et al [25] evaluated 12 sessions of client-centered psychotherapy in comparison to a wait list in a nonrandomized trial and found lasting treatment effects for adjustment disorder symptoms, life satisfaction, and general functioning. The effects were maintained in a 2 year catamnesis study [26]. Furthermore, brief dynamic therapy and brief supportive therapy were both applied in a sample of patients suffering from minor depression and adjustment disorder. Both approaches produced significant improvement in depression and anxiety compared to a nontreated control group [27]. In a later study by Ben-Itzhak et al [28], short-term psychodynamic treatment (3 months) was found to be as effective as intermediate dynamic therapy (12

XSL•FO
RenderX

months). The authors concluded that brief interventions seem thereby well suited for the treatment of adjustment disorder.

There are two intervention studies in a group setting that are available. First, a German cognitive-behavioral manual provides a modular approach to facilitate adjustment to significant life stressors in 10 sessions [21]. In a clinical trial, a significant decrease in anxiety, anger, and an increase in mood were achieved compared to the wait list condition. Second, a recent study by Hsiao et al [29] randomly assigned adjustment disorder patients to an eight week body-mind-spirit group psychotherapy and a control group (one session psychoeducation). The intervention focused on enhancing patients' resilience to cope with stress. With regard to depression and anxiety symptoms, no differential change was achieved, however, suicidal ideation was significantly reduced and HPA axis hyperactivity was reduced in the intervention group.

It is generally agreed that psychotherapy is the treatment of choice in adjustment disorder and very few pharmacotherapy studies are available to date [20]. Surprisingly, a recent study found that 37% of the patients with adjustment disorder are prescribed a psychotropic drug [30].

In conclusion, there are several promising psychotherapeutic approaches that have shown to be effective in treating adjustment disorder in several empirical studies. However, these studies are based on very heterogeneous theoretical foundations and a need for replication of the results is apparent.

E-Mental Health-Based Interventions

A vast amount of research in the area of e-mental health has been conducted during the past decade, particularly for the highly prevalent disorders such as depression and anxiety disorders, for example [31]. Common e-mental health interventions are based on CBT approaches and include unguided self-help treatments and guided self-help interventions with varying degrees of therapist contact [32]. E-interventions are effective approaches in the treatment of various psychological health problems, for example [2], and they have several advantages such as easy accessibility, easy use independent of time and place at a self-determined pace, and low cost of delivery to large populations [32].

As adjustment disorder is highly prevalent and frequently diagnosed by general practitioners, a large part of patients go without psychotherapeutic treatment [33]. Consequently, there is a need for widely available treatment options. E-mental health interventions have the capacity to meet this demand. In addition, the transient character of adjustment disorder makes it uniquely suited for low-threshold interventions.

Previous E-Mental-Health-Based Approach

Until now, only one intervention in the domain of e-mental health was devoted to adjustment disorder [3] using the virtual reality (VR) program "EMMA's world" [34]. The authors developed this VR program aiming at the introduction of positive mood and joy. Furthermore, the virtual world enables patients to confront, accept, and handle difficult emotions and cognitions connected to a stressful life event. Strategies are deduced from positive psychology attempting to increase the natural ability of humans to resist and grow in adverse circumstances [35]. The VR program is meant to be used in combination with face-to-face elements, and comprises six weekly sessions. In every session, patients are introduced to new concepts by the therapist and then proceed to the emotional processing part of the "EMMA's world" program [34]. This program is used to activate and process emotions and cognitions associated with the stressful event and to provide exposure to anything that is avoided in order to allow processing of all emotional aspects of the event. The approach builds upon a rationale developed by Foa and Kozak [36] for the treatment of PTSD, which proposes that pathological fear structures need to be activated as completely as possible to enable processing. The stressful event is represented by three-dimensional objects, images, sounds, and music, which can be enriched by the patient with significant personal items such as photographs. Moreover, the therapist has options to modify landscapes and environments in real time (eg, daytime, weather) if he wants to enhance or reflect on the patient's emotions. Such interactive components are a noteworthy feature of computer-assisted mental health interventions. Results from a preliminary case study point to favorable effects of the 6 week VR supported program with regard to posttraumatic growth, depression, and negative affect [3]. However, the program has not yet been evaluated in a large-scale empirical study.

Methods

Own Self-Help Approach

In order to bridge the gap between demands and supply in adjustment disorder interventions, a bibliotherapeutic self-help manual for adjustment disorder was developed [37], based on the diagnostic concept of the ICD-11 [10]. The manual targets the specific group of burglary victims as a first sample case into the field of adjustment disorders. Burglary is experienced as a severe violation or intrusion of the victim's privacy, possibly resulting in adjustment disorder. Symptoms range from anxiety, anger, and hostility to sadness, for example [38]. Moreover, victims frequently report avoidance of thoughts and feelings related to the event and concern for the future in combination with hypervigilance [39].

In an ongoing project by our team, the unguided self-help manual is implemented as a stand-alone eHealth intervention. The self-help manual in its bibliotherapeutic as well as its Internet presentation follows cognitive-behavioral principles, integrating formerly validated exercises from treatment approaches for PTSD, anxiety disorders, or depression. The exercises were specifically selected to work on symptoms of preoccupation (eg, constant rumination, excessive worry about the stressor) and failure to adapt (eg, difficulties concentrating, sleep disturbance, loss of interest in previously enjoyable activities). The manual is meant to be worked on during approximately 4 weeks, guided by a timetable outlining the course of the treatment,

- In the first part of the manual, a self-test for the screening of adjustment disorder symptoms (Adjustment Disorders - New Module-6) [40] is provided, which helps the reader to evaluate the psychological burden caused by a stressful

XSL•FO
RenderX

life event. Subsequently, the individual stress reaction is acknowledged as a normal response to an unusual situation, and it is pointed out that strong feelings are not uncommon in such situations. Sympathy and understanding is shown to further increase the therapy motivation.

- The second part of the manual is predominantly psychoeducational and provides information about the causes of adjustment disorder, its symptoms, and its possible consequences. It also contains a checklist for evaluating if face-to-face contact with a trained mental health professional was more appropriate than the self-help approach.
- The third part constitutes the main body of the self-help manual and comprises the CBT exercises that are structured in four chapters: (1) *Sense of self*, the aim is to guide the person to understand the origin of his or her stress-response, including an analysis of risk and protective factors. Psychoeducation on the mutual influence of feelings and cognitions is provided and an evaluation of previous coping strategies performed. (2) *Coping*, a variety of cognitive strategies such as the recognition and correction of mental biases are presented as well as thought-stopping techniques and strategies for the management of anxieties. As in the treatment of PTSD, a written narrative exposition exercise is suggested. (3) *Activation*, exercises aiming at introducing positive emotions, activation of personal resources, and forming and expressing positive and realistic personal goals. The reader receives information on the advantages of physical activities and is motivated to engage in athletic activities. And (4) *Recovery*, the importance of balancing activity on the one hand and relaxation on the other hand is stressed, and psychoeducation on the physical processes during relaxation is provided. A variety of relaxation exercises and a focus on sleep hygiene complete the intervention. While readers are advised to work through exercises in chapters 1 and 2 successively, exercises in chapters 3 and 4 are advised to be chosen according to personal priority and taste.

The following case description illustrates the course and outcomes of the self-help manual for a person suffering from burglary-triggered adjustment disorder. The sample case stems from an evaluative study of the bibliotherapeutic form of the manual.

Case Illustration

Mr. H. is a 68-year-old married man who lives together with his wife in a house owned by the family. They have no children and he is retired. Approximately eight months ago, he was victimized in a burglary. During the night, when he and his wife were asleep, the burglars entered his home, and with it his most private space. Although Mr. H. was in his bedroom, the burglars did not hesitate to steal most of his electronic devices and caused financial damage of over 10'000 Swiss Francs (approx. US $10,000). Mr. and Mrs. H. did not hear or see the burglars. They became aware of the situation the next morning when they woke up and called the police immediately.

Mr. H. was horrified over what had happened to him and his wife. He reacted with physical arousal such as sweating, shivering, and heart palpitation. In addition, he felt very angry and frustrated that he could not do more himself. Although the burglars did not vandalize the house, Mr. H. himself felt tainted. Mr. H. has suffered from “psychological pain”, as he called it ever since, despite the fact that his insurance covered the material damage. It was the first time somebody had entered his home without permission, and for him it was very difficult to adapt to this thought. The weeks passed and Mr. H. could not really adapt to the situation. It all did not make sense to him, and he asked himself, “Why me?” “Why now?” “Will it happen again?”. He spent much time ruminating about the burglary and felt that he could not control his thoughts in this respect.

Therefore, he signed up for participation in the study investigating the psychological effects of burglary on the victims and evaluating the above mentioned self-help manual. Although Mr. H. had never worked with a self-help guide before, he was confident that it could help him. During his participation, Mr. H. had to fill out questionnaires at three points in time.

Course of Study

The first survey contained questions about the burglary, and Mr. H.s’ emotional reactions. According to the new ICD-11 concept of adjustment disorder [10], Mr. H fulfilled the criteria of an adjustment disorder with strong symptoms of preoccupation and failure to adapt. In the Adjustment Disorder New Module-20 [41], a questionnaire specially designed for the ICD-11 adjustment disorder, Mr. H reached a total score of 48 (scale range 20-60). He scored 15 (range 4-16) points on the preoccupation subscale and 11 (range 4-16) points on the failure to adapt subscale. Figures 1 and 2 show graphic illustrations of the symptom decrease between pre and posttreatment. On the Depression Anxiety Stress Scales (DASS-21) [42], an established instrument to measure low positive affectivity (depression), physiological arousal (anxiety), and stress, he scored a total of 24 out of 63 points. He reached higher scores on the depression and anxiety subscales than on the stress subscale.

Mr. H. worked with the treatment manual for four weeks. He took out the book several times a week for two or three hours, so that in the end of the fourth week, he had worked on all the exercises in the book. In total, he worked for about 15 hours during the four weeks. He thought that almost all the exercises helped him to deal with his pain over the incident.

XSL•FO
RenderX

Figure 1. Adjustment Disorder – New Module-20 (ADNM) total symptom improvement (the total score comprises preoccupation, failure-to-adapt, and accessory symptoms).

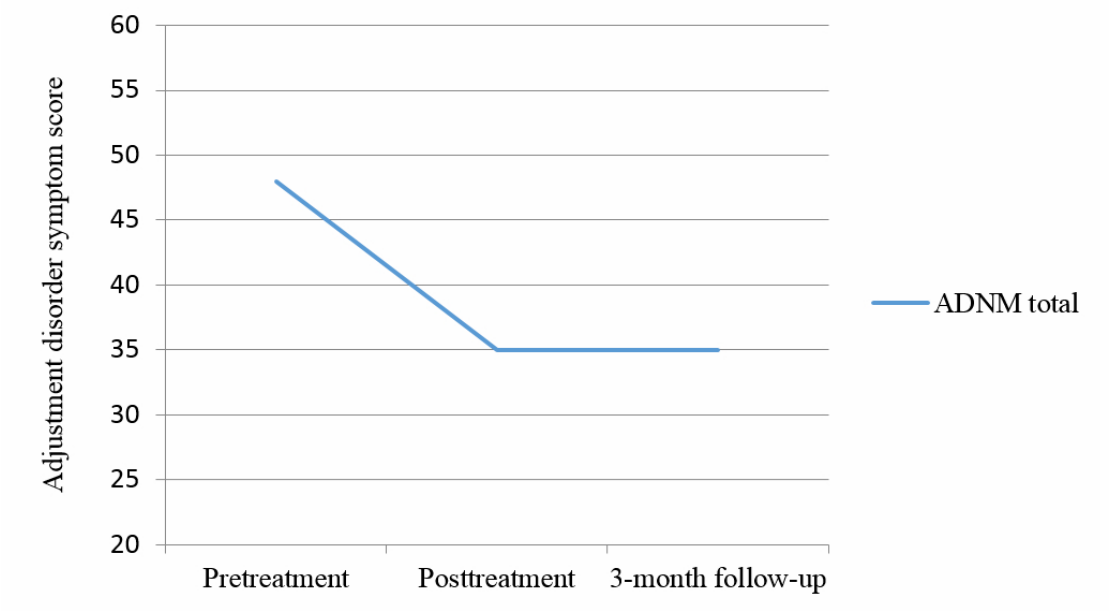
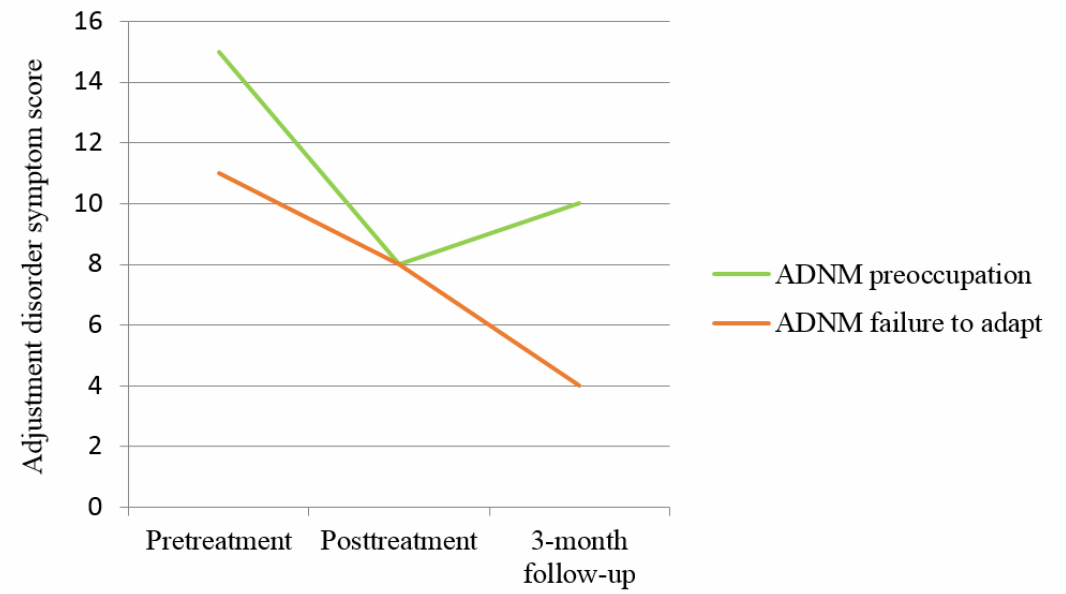


Figure 2. Symptom improvement on subscales of the Adjustment Disorder – New Module-20 (ADNM).



Results

Second Measurement

At the time of the second measurement, the symptoms of adjustment disorder had decreased for Mr. H., and he did not fulfill the adjustment disorder criteria anymore. He scored a total of 35 points with 8 points on the preoccupation subscale and 8 points on the failure to adapt subscale (Figure 2). Thus, he showed a substantial decrease in preoccupation (7 points) and failure to adapt symptoms (3 points). It seems as if the manual had helped him to learn to manage his thoughts and feelings about the burglary. He still thought about it from time to time, but not as often as before. Moreover, the thoughts did not cause as much pain as before, and it was easier now to adapt to the situation. In the DASS-21, he showed a slight decrease of symptoms on the depression (two points) and anxiety subscales (one point). The measures for the DASS-21 are presented in Table 1.

XSL•FO
RenderX

Table 1. Pretreatment and posttreatment measures on DASS-21.

	Pretreatment	Posttreatment	3-month follow-up
DASS-21 total[a]	24	21	16
Depression[b]	4	2	0
Anxiety[c]	4	3	1
Stress[d]	2	2	0

[a]DASS-21 total range 0-63

[b]depression range 0-21

[c]anxiety range 0-21

[d]stress range 0-21

Three Month Follow-Up Assessment

After three months, Mr. H. participated in the follow-up assessment of the study. He still did not fulfill the criteria of an adjustment disorder. Some symptoms of preoccupation had reoccurred, but apparently he learned how to handle them. Adaption to the situation did not seem to be a problem anymore as is represented by the lowest possible scoring on the failure to adapt subscale. In the long run, the manual seemed to help Mr. H. by showing him ways on how to deal with the difficult thoughts and feelings that occurred when being reminded of the burglary. The depression, anxiety, and the stress symptoms were also reduced during the three months after Mr. H. had worked with the self-help manual.

Overall, the treatment with the self-help manual reduced symptoms of adjustment disorder, namely preoccupation and failure to adapt, as well as symptoms of depression, anxiety, and stress. Mr. H. reports that it was the combination of exercises, rather than an individual exercise, which helped him in dealing with difficult feelings and thoughts related to the burglary. He did not point out any specific exercise that helped the most. Before treatment, he suffered from recurrent thoughts about the loss and problems to adapt to the altered circumstances for eight months. After one month of working on the subject with cognitive behavioral methods as captured in the self-help manual, a significant change in symptom severity was reached. After three months of treatment, symptom severity was still considerably lower than before treatment. Although some preoccupation symptoms recurred, symptoms of failure to adapt declined again. This case demonstrates that a low intensity intervention on the basis of a cognitive behavioral approach, can achieve positive short- and long-term results in the treatment of adjustment disorder.

Discussion

eHealth Treatments for Adjustment Disorders

With this paper, we want to show promising prospects for eHealth treatments for adjustment disorder. Adjustment disorder is a highly prevalent, but underresearched, mental disorder that is undergoing significant changes with regard to its diagnostic conceptualization in the course of DSM-5 and ICD-11 revisions.

Discriminant validity is improved by the definition of specific diagnostic criteria, especially with regard to depressive episodes. This development lays the foundation for future research on development and upholding factors as well as intervention methods. Despite the extensive need for interventions, the current empirical evidence for the effectiveness of psychotherapeutic treatment is sparse in the face-to-face context, as well as in e-mental health. Individual studies based on various theoretical backgrounds such as CBT, dynamic approaches, or client-centered psychotherapy are encouraging, but need to be replicated. The domain of e-mental health has only reluctantly begun to expand to adjustment disorder, with one approach demonstrating the usage of VR to complement face-to-face therapeutic contact [3]. A new unguided self-help manual, which is currently adapted to an Internet-based format, shows positive preliminary results, as was illustrated in a case study.

The structure, distribution, and costs of mental health care make it unavailable for a significant part of the world's population [43]. Taking into account the high occurrence rates of serious nontraumatic stressors and the considerable prevalence rates of adjustment disorder, a continued promotion of eHealth approaches might prove uniquely capable for meeting the diverse demands for scientists, health care practitioners, and patients. From a research perspective, an e-intervention enables the researcher to get information on user behavior (eg, which components are accessed most often) and to evaluate the effectiveness of different modules. Furthermore, it can contribute to reduce the translational lag in transfer of research findings into regular mental health care [44]. From the perspective of health care providers, it could be used to propose a cost-effective prevention measure as soon as a disturbing critical life event gets reported. This way, it could be possible to reduce the chance for subsequent disorders for which adjustment disorder can act as a precursor [45,46]. From the viewpoint of those possibly affected by adjustment disorder, a website can potentially reach out to persons that otherwise would not be able to access mental health care. Even more important, it initially eliminates the stigma connected to seeing a therapist [32], while providing clear guidelines to assess whether professional care is indicated.

Preliminary results from the evaluation study of the new unguided self-help manual for burglary victims with adjustment disorder are encouraging. The manual is one of the few interventions specifically targeting the symptoms of adjustment disorder, and the first one to be based on the revised concept according to ICD-11, to our knowledge. In a case study, symptoms of preoccupation and failure to adapt, as well as

XSL•FO
RenderX

symptoms of depression, anxiety, and stress were successfully reduced. This case demonstrates that with low intensity intervention on the basis of a cognitive behavioral approach, positive short- and long-term results in treatment of adjustment disorder can be achieved. CBT interventions are particularly well suited for adaption to a computer format due to their structured, mostly modular organization and the focus on behavior and cognition [47].

A second domain of e-mental health interventions comprises interventions with varying degrees of therapist contact, namely, guided self-help approaches. Such treatments seem promising for the future as their effectiveness was shown for various mental disorders, for example [2]. Even though some advantages of unguided self-help are attenuated, such as cost-effectiveness or accessibility, guided self-help interventions are usually associated with higher adherence to treatment, less dropouts, and higher effects than unguided interventions. Furthermore, the approach by Andreu-Mateu et al [3] that combined VR and face-to-face elements was evaluated positively in two case studies. Its inspiration from PTSD therapy focusing on activation of pathological fear structures makes it a particularly interesting procedure [36]. This is in line with the new conceptualization of adjustment disorder as a stress-response syndrome. More traditionally, however, therapist contact in guided self-help interventions takes place via the Internet, for example, in the form of weekly emails [48]. Such approaches would be desirable for adjustment disorder treatment, but have not yet been identified in the literature.

To date, the Spanish VR study [3], as well as our own self-help approach for treating burglary victims with adjustment disorder [37], have been evaluated only in case studies. We are currently testing the bibliotherapeutic manual in a randomized controlled clinical trial. With regard to the Internet version, a pilot study for its implementation into clinical practice is in preparation. Additionally, the manual is currently constricted to treating the specific group of burglary victims suffering from adjustment disorder symptoms. In the future, the intervention should be adapted to more diverse populations suffering from different stressful life events such as job loss, relationship break-up, or financial problems.

Conclusions

Present classificatory changes pave the way for further research on conceptual theories of adjustment disorder, upholding factors as well as therapeutic interventions. On the one hand, this concerns interventions in the domain of traditional face-to-face psychotherapy; on the other hand, the sector of e-mental health is encouraged to expand to this currently underrepresented territory. E-mental health options are considered uniquely suited for offering early intervention after the experiences of stressful life events that potentially trigger adjustment disorder.

Conflicts of Interest

None declared.

References

1. Andrews G, Cuijpers P, Craske MG, McEvoy P, Titov N. Computer therapy for the anxiety and depressive disorders is effective, acceptable and practical health care: A meta-analysis. PLoS One 2010;5(10):e13196 [FREE Full text] [doi: 10.1371/journal.pone.0013196] [Medline: 20967242]
2. Andersson G, Cuijpers P, Carlbring P, Riper H, Hedman E. Guided Internet-based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: A systematic review and meta-analysis. World Psychiatry 2014 Oct;13(3):288-295 [FREE Full text] [doi: 10.1002/wps.20151] [Medline: 25273302]
3. Andreu-Mateu S. Behavioral Psychology/Psicología Conductal. 2012. La utilización de la realidad virtual y estrategias de psicología positiva en el tratamiento de los trastornos adaptativos URL: http://repositori.uji.es/xmlui/handle/10234/67199 [accessed 2015-04-29] [WebCite Cache ID 6Y9GL66Tl]
4. Callahan EJ, Hilty DM, Nesbitt TS. Patient satisfaction with telemedicine consultation in primary care: Comparison of ratings of medical and mental health applications. Telemed J 1998;4(4):363-369. [Medline: 10220477]
5. Baños Rivera RM, Guillén Botella V, Botella Arbona C, García Palacios A, Jorquera Rodero M, Quero Castellano S. Apuntes de Psicología. 2008. Un programa de tratamiento para los trastornos adaptativos. Un estudio de caso URL: http://apuntesdepsicologia.es/index.php/revista/article/view/263 [accessed 2015-04-29] [WebCite Cache ID 6Y9GXKbU0]
6. Maercker A, Bachem RC, Simmen Jannevska K. Encyclopedia of clinical psychology, Vol. X. Chichester: Wiley; 2015. Adjustment disorders URL: http://www.wiley.com/WileyCDA/WileyTitle/productCd-0470671270.html [accessed 2015-04-29] [WebCite Cache ID 6Y9GdKsYM]
7. Casey P. Adjustment disorder: New developments. Curr Psychiatry Rep 2014 Jun;16(6):451. [doi: 10.1007/s11920-014-0451-2] [Medline: 24748555]
8. Strain JJ, Diefenbacher A. The adjustment disorders: The conundrums of the diagnoses. Compr Psychiatry 2008;49(2):121-130. [doi: 10.1016/j.comppsych.2007.10.002] [Medline: 18243883]
9. Keeley JW, Reed GM, Roberts MC, Evans SC, Robles R, Matsumoto C, et al. Disorders specifically associated with stress: A case controlled field study for ICD-11 mental and behavioral disorders. (forthcoming 2015) 2015 (forthcoming).
10. Maercker A. Lancet. 2013. Proposals for mental disorders specifically associated with stress in the International Classification of Diseases-11 URL: http://www.thelancet.com/journals/lancet/article/PIIS0140-6736(12)62191-6/abstract [accessed 2015-04-29] [WebCite Cache ID 6Y9GldvpZ]

XSL•FO
RenderX

11. Baumeister H, Kufner K. It is time to adjust the adjustment disorder category. Curr Opin Psychiatry 2009 Jul;22(4):409-412. [doi: 10.1097/YCO.0b013e32832cae5e] [Medline: 19436201]
12. Doherty AM, Jabbar F, Kelly BD, Casey P. Distinguishing between adjustment disorder and depressive episode in clinical practice: The role of personality disorder. J Affect Disord 2014 Oct 15;168:78-85. [doi: 10.1016/j.jad.2014.06.034] [Medline: 25043318]
13. Selye H. The stress of life. New York: McGraw-Hill; 1978.
14. Strain JJ, Friedman MJ. Considering adjustment disorders as stress response syndromes for DSM-5. Depress Anxiety 2011 Sep;28(9):818-823. [doi: 10.1002/da.20782] [Medline: 21254314]
15. American Psychiatric Association. Diagnostic and statistical manual of mental disorders (5th ed.), Washington, DC. Diagnostic and statistical manual of mental disorders (5th ed.), Washington, DC 2013. [doi: 10.1176/appi.books.9780890425596.dsm07]
16. Gradus JL, Qin P, Lincoln AK, Miller M, Lawler E, Lash TL. The association between adjustment disorder diagnosed at psychiatric treatment facilities and completed suicide. Clin Epidemiol 2010;2:23-28 [FREE Full text] [Medline: 20865099]
17. Caplan G. Principles of preventive psychiatry. New York: Basic books; 1964. URL: http://psycnet.apa.org/psycinfo/1965-02239-000 [accessed 2015-04-29] [WebCite Cache ID 6Y9Hjxj10]
18. Horowitz MJ. Stress-response syndromes: A review of posttraumatic and adjustment disorders. Hosp Community Psychiatry 1986 Mar;37(3):241-249. [Medline: 3957267]
19. Maercker A, Einsle F, Kollner V. Adjustment disorders as stress response syndromes: A new diagnostic concept and its exploration in a medical sample. Psychopathology 2007;40(3):135-146. [doi: 10.1159/000099290] [Medline: 17284941]
20. O'Connor B, Cartwright H. in Handbook of evidence-based practice in clinical psychology, P. Sturmey M. Hersen, Editors, John Wiley & Sons Hoboken, NJ. 2012. Adjustment disorder URL: http://onlinelibrary.wiley.com/doi/10.1002/9781118156391.ebcp002021/abstract [accessed 2015-04-26] [WebCite Cache ID 6Y9Hr2pXz]
21. Reschke K, Teichmann K. Entwicklung und Evaluation eines kognitiv-behavioralen Therapieprogramms für Patienten mit Anpassungsstörung. Psychosom Konsiliarpsychiatr 2008 Jul 28;2(2):98-103. [doi: 10.1007/s11800-008-0099-z]
22. Maercker A. Posttraumatische belastungsstörungen. Aufl 2009. [doi: 10.1007/978-3-540-88489-7]
23. Cvetek R. EMDR treatment of distressful experiences that fail to meet the criteria for PTSD. J EMDR Prac Res 2008 Mar 01;2(1):2-14. [doi: 10.1891/1933-3196.2.1.2]
24. van der Klink J J L, Blonk R W B, Schene AH, van Dijk F J H. Reducing long term sickness absence by an activating intervention in adjustment disorders: A cluster randomised controlled design. Occup Environ Med 2003 Jun;60(6):429-437 [FREE Full text] [Medline: 12771395]
25. Altenhöfer A, Schulz W, Schwab R, Eckert J. Psychotherapie von Anpassungsstörungen. Psychotherapeut 2007 Jan;52(1):24-34. [doi: 10.1007/s00278-006-0520-z]
26. Gorschenek N. , R. Schwab, and J. Eckert, Psychotherapie von Anpassungsstörungen. Psychotherapie, psychosomatik, medizinische psychologie 2008;58(05):200-207. [doi: 10.1055/s-2007-986180]
27. Maina G, Forner F, Bogetto F. Randomized controlled trial comparing brief dynamic and supportive therapy with waiting list condition in minor depressive disorders. Psychother Psychosom 2005;74(1):43-50. [doi: 10.1159/000082026] [Medline: 15627856]
28. Ben-Itzhak S, Bluvstein I, Schreiber S, Aharonov-Zaig I, Maor M, Lipnik R, et al. The effectiveness of brief versus intermediate duration psychodynamic psychotherapy in the treatment of adjustment disorder. J Contemp Psychother 2012 Mar 31;42(4):249-256. [doi: 10.1007/s10879-012-9208-6]
29. Hsiao FH, Lai YM, Chen YT, Yang TT, Liao SC, Ho Rainbow T H, Chan Cecilia L W, et al. Efficacy of psychotherapy on diurnal cortisol patterns and suicidal ideation in adjustment disorder with depressed mood. Gen Hosp Psychiatry 2014;36(2):214-219. [doi: 10.1016/j.genhosppsych.2013.10.019] [Medline: 24342115]
30. Fernández A. Adjustment disorders in primary care: Prevalence, recognition and use of services. The British Journal of Psychiatry 2012;201(2):137-142. [doi: 10.1192/bjp.bp.111.096305]
31. Titov N, Dear B, Rapee R. New horizons for e-mental health research and practice. Aust N Z J Psychiatry 2013 Feb;47(2):107-108. [doi: 10.1177/0004867412470011] [Medline: 23382505]
32. Cuijpers P, van SA, Andersson G. Internet-administered cognitive behavior therapy for health problems: A systematic review. J Behav Med 2008 Apr;31(2):169-177 [FREE Full text] [doi: 10.1007/s10865-007-9144-1] [Medline: 18165893]
33. Carta M, Balestrieri M, Murru A, Hardoy MC. Adjustment disorder: Epidemiology, diagnosis and treatment. 2009. URL: http://www.cpementalhealth.com/content/5/1/15 [accessed 2015-04-22] [WebCite Cache ID 6XyrALvJ5]
34. Botella C. Using an adaptive display for the treatment of emotional disorders: A preliminary analysis of effectiveness. in CHI 2006:2006. [doi: 10.1145/1125451.1125574]
35. Seligman Martin E P, Steen TA, Park N, Peterson C. Positive psychology progress: Empirical validation of interventions. Am Psychol 2005;60(5):410-421. [doi: 10.1037/0003-066X.60.5.410] [Medline: 16045394]
36. Foa EB, Kozak MJ. Emotional processing of fear: Exposure to corrective information. Psychol Bull 1986 Jan;99(1):20-35. [Medline: 2871574]

XSL•FO
RenderX

37. Bachem R. Ratgeber um nach einem Einbruch zurück in das gewohnte Leben zu finden. 2013. and A. Maercker, Ist es noch mein Zuahuse? URL: http://www.mobi.ch/static/mobi.ch/content/documents/SV_Ratgeber/einbruch_meinzuhause_ratgeber_de.pdf [accessed 2015-04-29] [WebCite Cache ID 6Y9I0MXyh]
38. Beaton A, Cook M, Kavanagh M, Herrington C. The psychological impact of burglary. Psychology, Crime & Law 2000 Jan;6(1):33-43. [doi: 10.1080/10683160008410830]
39. Caballero M, Ramos L, Saltijeral MT. Salud Mental. 2000. El trastorno por estres postraumatico y otras reacciones en las victimas del robo a casa URL: http://bibliosedia.bi-digital.com:8080/handle/123456789/1814 [accessed 2015-04-29] [WebCite Cache ID 6Y9I7XaZS]
40. Komper CJ, Brähler E, Zenger M. Medizinisch Wissenschaftliche Verlagsgesellschaft. Berlin: Medizinisch Wissenschaftliche Verlagsgesellschaft; 2014. ADNM-6: Anpassungsstorungs-Screeningskala, in Psychologische und sozialwissenschaftliche Kurzskalen. Standardisierte Erhebungsinstrumente fur Wissenschaft und Praxis URL: http://www.mwv-berlin.de/buecher-bestellen/product_info.php?info=p614_Psychologische-und-sozialwissenschaftliche-Kurzskalen.html&XTCsid=682bc791f578895511e042421a9ceac7 [WebCite Cache ID 6Y9IGlElV]
41. Einsle F, Köllner V, Dannemann S, Maercker A. Development and validation of a self-report for the assessment of adjustment disorders. Psychol Health Med 2010 Oct;15(5):584-595. [doi: 10.1080/13548506.2010.487107] [Medline: 20835968]
42. Antony M. Psychological assessment. 1998. Psychometric properties of the 42-item and 21-item versions of the Depression Anxiety Stress Scales in clinical groups and a community sample URL: http://psycnet.apa.org/journals/pas/10/2/176/ [accessed 2015-04-30] [WebCite Cache ID 6YBLtSS4L]
43. Christensen H, Hickie IB. E-mental health: A new era in delivery of mental health services. Med J Aust 2010 Jun 7;192(11 Suppl):S2-S3. [Medline: 20528702]
44. Christensen H, Petrie K. Information technology as the key to accelerating advances in mental health care. Aust N Z J Psychiatry 2013 Feb;47(2):114-116. [doi: 10.1177/0004867412471088] [Medline: 23382508]
45. Strain JJ, Smith GC, Hammer JS, McKenzie DP, Blumenfield M, Muskin P, et al. Adjustment disorder: A multisite study of its utilization and interventions in the consultation-liaison psychiatry setting. Gen Hosp Psychiatry 1998 May;20(3):139-149. [Medline: 9650031]
46. Kocalevent R, Mierke A, Danzer G, Klapp BF. Adjustment disorders as a stress-related disorder: A longitudinal study of the associations among stress, resources, and mental health. PLOS ONE 2014;9(5):7303. [doi: 10.1371/journal.pone.0097303]
47. Spek V, Cuijpers P, Nyklícek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: A meta-analysis. Psychol Med 2007 Mar;37(3):319-328. [doi: 10.1017/S0033291706008944] [Medline: 17112400]
48. Berger T, Caspar F, Richardson R, Kneubühler B, Sutter D, Andersson G. Internet-based treatment of social phobia: A randomized controlled trial comparing unguided with two types of guided self-help. Behav Res Ther 2011 Mar;49(3):158-169. [doi: 10.1016/j.brat.2010.12.007] [Medline: 21255767]

Abbreviations

BDI: Beck Depression Inventory
CBT: cognitive behavior therapy
DASS-21: Depression Anxiety Stress Scales
DE: depressive episode
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, 5th Version
EMDR: eye movement desensitization and reprocessing
HPA: hypothalamic-pituitary-adrenocortical
ICD-11: International Classification of Diseases 11
PTSD: posttraumatic stress disorder
VR: virtual reality

Edited by G Eysenbach; submitted 18.12.14; peer-reviewed by RD Kocalevent; accepted 05.03.15; published 08.05.15.

Please cite as:
Maercker A, Bachem RC, Lorenz L, Moser CT, Berger T
Adjustment Disorders Are Uniquely Suited for eHealth Interventions: Concept and Case Study
JMIR Mental Health 2015;2(2):e15
URL: http://mental.jmir.org/2015/2/e15/
doi:10.2196/mental.4157
PMID:26543920

XSL•FO
RenderX

©Andreas Maercker, Rahel C Bachem, Louisa Lorenz, Christian T Moser, Thomas Berger. Originally published in JMIR Mental Health (http://mental.jmir.org), 08.05.2015. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Integrating Health Behavior Theory and Design Elements in Serious Games

Colleen Cheek[1], RN, BSc, MIS; Theresa Fleming[2], PhD; Mathijs FG Lucassen[2], NZROT, PhD; Heather Bridgman[3], BPsych, DClinHlthPsych; Karolina Stasiak[2], PhD; Matthew Shepherd[4], DClinPsy; Peter Orpin[3], PhD

[1]Rural Clinical School, School of Medicine, University of Tasmania, Burnie, Australia

[2]Werry Centre for Child and Adolescent Mental health, Department of Psychological Medicine, University of Auckland, Auckland, New Zealand

[3]Centre for Rural Health, University of Tasmania, Launceston, Australia

[4]School of Counselling, Human Services and Social Work, University of Auckland, Auckland, New Zealand

Corresponding Author:
Colleen Cheek, RN, BSc, MIS
Rural Clinical School
School of Medicine
University of Tasmania
Private Bag 3513
Burnie,
Australia
Phone: 61 3 6430 5922
Fax: 61 3 6431 5670
Email: Colleen.Cheek@utas.edu.au

Abstract

Background: Internet interventions for improving health and well-being have the potential to reach many people and fill gaps in service provision. Serious gaming interfaces provide opportunities to optimize user adherence and impact. Health interventions based in theory and evidence and tailored to psychological constructs have been found to be more effective to promote behavior change. Defining the design elements which engage users and help them to meet their goals can contribute to better informed serious games.

Objective: To elucidate design elements important in SPARX, a serious game for adolescents with depression, from a user-centered perspective.

Methods: We proposed a model based on an established theory of health behavior change and practical features of serious game design to organize ideas and rationale. We analyzed data from 5 studies comprising a total of 22 focus groups and 66 semistructured interviews conducted with youth and families in New Zealand and Australia who had viewed or used SPARX. User perceptions of the game were applied to this framework.

Results: A coherent framework was established using the three constructs of self-determination theory (SDT), autonomy, competence, and relatedness, to organize user perceptions and design elements within four areas important in design: computer game, accessibility, working alliance, and learning in immersion. User perceptions mapped well to the framework, which may assist developers in understanding the context of user needs. By mapping these elements against the constructs of SDT, we were able to propose a sound theoretical base for the model.

Conclusions: This study's method allowed for the articulation of design elements in a serious game from a user-centered perspective within a coherent overarching framework. The framework can be used to deliberately incorporate serious game design elements that support a user's sense of autonomy, competence, and relatedness, key constructs which have been found to mediate motivation at all stages of the change process. The resulting model introduces promising avenues for future exploration. Involving users in program design remains an imperative if serious games are to be fit for purpose.

(JMIR Mental Health 2015;2(2):e11) doi:10.2196/mental.4133

KEYWORDS

depression; adolescent; cognitive behavior therapy; online intervention; user-centered; self-efficacy; motivation; adherence; SPARX

XSL•FO
RenderX

Introduction

Background

Mental health conditions account for 13% of the global disease burden, with depression being the largest single cause of disability worldwide [1]. The large gap between treatment need and service provision is a global issue: in high income countries, 35% to 50% of people receive no treatment for severe mental disorders and in low income countries this fraction is much greater, with 76% to 85% not getting treatment [1]. The investment needed to provide trained therapists to fill these gaps is unfeasible [2], and not all potential patients favor or can access existing modes of treatment delivery. Integrating a range of accessible user-driven options into general community-level settings is one of the strategies promoted in the World Health Organization Mental Health Action Plan 2013-2020 [1]. There is an emphasis on early intervention, respecting the autonomy of individuals with mental health issues, and nonpharmacological therapies promoted particularly for young people [1].

The ubiquity of the Internet provides an opportunity for online computerized tools to extend the reach of psychotherapies such as cognitive behavioral therapy (CBT). While computerized therapies have been shown to be effective in alleviating depression and anxiety symptoms in adults, adolescents, and children [3,4], issues with user engagement and high attrition are noted in efficacy studies of the currently available text-based computerized cognitive behavioral therapy (cCBT) programs [5]. The opportunity for a more interactive and graphically rich experience via computer-based gaming technology has fueled development of serious games for mental health.

Gaming strategies that increase positive outcomes for mental health issues are only beginning to be explored; nevertheless, there have been a number of serious games developed to enhance mental well-being. These range from publically available but not necessarily clinically tested tools, such as Depression Quest, to programs which have been evaluated and reported in the peer-reviewed literature, such as gNats Island [6], SPARX [7], Camp Cope-A-Lot [8], Reach Out Central [9], and Virtual Iraq [10]. Most serious games for mental health issues that have been clinically tested are not publically available, and few describe the features of the game in any depth [11]. The gaming elements differ markedly among these: serious games for mental health include exercise programs with biofeedback, virtual reality simulations, word or number puzzles, and fantasy adventures. The program delivery processes also differ; some serious games (such as gNats Island) require a facilitator or clinician while others (such as SPARX) can be completed independently.

Bringing Health Behavior Theory to Game Design

A model which incorporates elements that promote engagement and adherence and help users meet their goals could inform development and evaluation of serious games. Theories based on existing knowledge can provide a coherent framework to organize ideas and rationale clearly, facilitating communication among stakeholders [12]. Health interventions to promote behavior change that are based in theory and evidence and tailored to psychological constructs have been found to contribute to increased adherence and effectiveness [13]. In this instance, the psychological constructs of most interest are those that help explain which features of a serious game will support the individual to engage with the program, maintain interest in completing the tasks, and incorporate therapeutic concepts into everyday life.

Appealing to a user's sense of self or agency and connectedness with others has been suggested to improve uptake and support engagement of computer-delivered therapies for depression and/or anxiety [14,15]. User-centered healthcare has gained momentum over recent decades and is very relevant with the expansion of consumer-based online resources and health technologies. More patient-led approaches to delivering services have also been recognized as a means of directing more efficient and effective use of health budgets [16]. Delivering health care in the 21st century requires personal access to tools that empower and allow citizens to pursue the best health strategies for themselves and their families [17]. This compels us to consider within the development process: "How can researchers construct participative health environments to support a patient's sense of autonomy (personal control over health decisions), competency (mastery over self-management skills), and connectedness (social support from relevant others)—all factors implicated by psychological research to influence the intrinsic motivation of individuals? [18]"

These are the tenets of self-determination theory (SDT)[19]. Self-determination is defined by Deci and Ryan [20] as "a quality of human functioning that involves the experience of choice;" the options to choose are the determinants of an individual's actions. SDT and other theoretical models of health behavior change, such as social-cognitive theory [21] and protection motivation theory [22], are considered continuum models; they predict that fulfillment of important psychological constructs will move an individual along a continuum toward behavior change and increase the likelihood of sustained outcomes. While continuum models are considered useful for explanation and prediction, stage models reflect the relative importance of different constructs at different stages of the behavior change process. From precontemplation through intention and action, interventions more specifically targeted to the needs of individuals at these different stages of health behavior change are considered more likely to improve recruitment, retention, and progress [23]. Stage theories of behavior change include the transtheoretical model [24] and the health action process approach (HAPA)[25]. HAPA demonstrates how understanding and designing to motivate users to access help, adhere to therapy, and sustain behavior change are critical in developing effective health interventions and include post-intentional volitional processes that lead to behavioral change [23]. In this model, self-efficacy and social support are important mediating factors at each stage of change, from contemplation through to actioning and sustaining change. Self-efficacy is defined as the confidence individuals must have in their ability to perform the desired action [25].

To provide a tool that can easily be communicated across stakeholder groups involved in serious game development, user perceptions based in psychological constructs need to be

XSL•FO
RenderX

translated into practical design features that enable or support the desired user perceptions. Due to its relative immaturity as a formal discipline, the underlying theories and elements of design for serious games for mental health issues must be inferred from each of the stakeholder groups involved in development. Marne and colleagues [26] described the creation of serious games as a collaboration of two broad stakeholder groups: pedagogical experts and game experts. Each brings strengths to the task, ensuring that the games are both engaging and educationally strong. There is also a body of literature relating to effective elements of commercial computer games and serious games for online learning. Links have been made between gaming and learning mechanics to emphasize instructional value [27]. We propose that the playful platform of computer gaming, the accessibility of the program online, the therapy and manner in which the therapy is delivered, and the way the content is structured to maximize learning all contribute to effective serious games for mental health.

By understanding how an intervention supports a user, we hypothesized we could identify serious game design elements that contribute to increased user engagement and adherence to therapy. We sought to do this by exploring user experiences with SPARX (smart, positive, active, realistic, X-factor thoughts), a self-help tool developed by authors of this paper (TF, ML, KS, MS) and others. SPARX was designed to deliver cCBT to adolescents aged 12 to 19 years old experiencing mild-to-moderate symptoms of depression using an engaging computerized platform.

The development and testing of SPARX has been described previously [28-32]. In brief, SPARX was developed using CBT and learning theory, with input on game design from youth and stakeholders. It uses a bicentric frame of reference [33]. In each module, users are explicitly introduced to therapeutic content using a virtual therapist or guide (Figure 1) and then transition to a fantasy setting to undertake CBT-based challenges and develop CBT-based skills within an overall narrative of restoring balance to the fantasy world (Figures 2 and 3). Following this exploratory learning, users return to the guide at the end of each level to reflect on the tasks and how they might be applied in their own lives. SPARX was shown to be at least as good as usual care (primarily counseling delivered face-to-face by a mental health clinician) for young people seeking help for low mood or depression [7]. It showed promising results in exploratory trials with Māori youth and students in alternative education programs, and a modified version (Rainbow SPARX) showed promise for sexual minority youth [29-31]. Youth trialing SPARX have also reported a high level of satisfaction and engagement with the program [28-32]. In the course of design and testing the program, a considerable amount of qualitative data was collected from these user-participants (Table 1). The user experiences are those of young people in community settings, thus targeted individuals in the pre-intentional to intentional stages according to HAPA. While data from these studies concerning young people's satisfaction with the program and opportunities to improve it have been reported in previous publications, the data have not been combined across these different groups, and no systematic exploration of design elements has been undertaken.

In this study, we sought to define design elements from a user-centered perspective, specifically for youth aged 12 to 19 years, by extracting from all the SPARX focus group and interview data those features that users perceived as being most important in meeting their goals. We hypothesized important elements of the four areas—a playful platform of computer gaming, the accessibility of the program online, the therapy and manner in which the therapy is delivered, and the way the content is structured to maximize learning—could be identified within SPARX and linked by the results with evidence of supporting the user experience. This analysis is useful because serious games for mental health are seldom described in depth and there is little research to elucidate components of serious games that might be useful or appealing. Articulation of a framework of critical design elements could facilitate theory development and testing in this new field.

XSL•FO
RenderX

Figure 1. SPARX: the guide.



XSL•FO
RenderX

Figure 2. SPARX: canyon dwellers.



Figure 3. SPARX: user avatar flying on the eagle Te Hokioi.



XSL•FO
RenderX

Methods

Qualitative Data

Approval for this study was granted by the Tasmanian Social Sciences Higher Research Ethics Committee in Australia (H14061). The qualitative data used were gathered during five separate studies of five different user groups to explore the acceptability of SPARX. The selection of participants, approval and consent processes, methods of data collection, and transcription of the interviews are described in the respective papers for which the data were gathered [28-32].

All checked and corrected interview and focus group transcripts from these studies were used in this analysis.

The interviews were conducted with participants from groups that had three different experiences of the SPARX program: youth and family members, practitioners, and community elders during the program design stages (design); youth shown components of the program once it had been finalized (preview); and youth who had used the finalized program as part of a formal research trial (users). Each brought a particular perspective and experience, constituting a form of triangulation (Table 1).

Table 1. Interview participant data.

Number of groups or interviews	N	Age range (years)	Characteristics	Participant code
Design—viewed a prototype				
6 groups, 1 interview	26	16-18	19 Māori 16-18, 7 parents/caregivers (15 female, 11 male) MHS NA[a]	Māori 1-26
3 groups	10[b]	16-27	Lesbian, gay and bisexual young people (5 female, 4 male) MHS NA[a]	Rainbow 1-9
Preview—viewed or tested one module				
4 groups, 1 interview	16	13-18	Rural Australian youth (4 female, 12 female) MHS NA[a]	Aus 1-16
9 groups	39	13-16	Youth excluded from mainstream education (10 female, 29 male) MHS NA[a]	AE 1-39
Users—participated in a trial/program				
25 interviews	25	13-19	Young people attracted to the same sex, both sexes, or not sure (13 female, 12 male) MHS NA[a]	Rainbow (User) 1-25
39 interviews	39	13-16	Youth excluded from mainstream education with symptoms of possible depression using the Children's Depression Rating Scale Revised (CDRS-R), 9 without symptoms. (15 female, 24 male).	AE (User) 1-39
5 interviews	5	14-16	Youth attending mainstream school presenting with symptoms of mild-moderate depression using CDRS-R. (4 female, 1 male)	Māori (User) 1-5

[a]Mental health status not assessed in sample.

[b]One participant participated in two focus groups (ie, there are nine unique individuals).

Full transcripts of all interviews and focus groups were obtained and were subjected to iterative thematic analysis using NVivo version 10 software (QSR International) for organization. This study proceeded in two main stages: a hypothetical model was derived from the literature for the purpose of testing against the available interview data and confirmatory testing of the model was performed through the use of qualitative data.

Creation of the Model

The data and publications describing the views of participants of SPARX were initially reviewed to assimilate emerging themes. From this an existing theoretical model of behavior change was selected which was relevant to the psychological constructs being described in the data. We then sought theories and features which pertained to the four areas: playful platform of computer gaming, the accessibility of the program online,

XSL•FO
RenderX

the therapy and manner in which the therapy is delivered, and the way the content is structured to maximize learning.

Confirmatory Testing of the Model Through the Use of Qualitative Data

Coding User Perceptions

An initial coding tree was established using the psychological constructs of the selected continuum model. Perceptions that did not fit within these constructs were coded as divergent themes. Initial coding was conducted by one member of the research team (CC) and validated at frequent intervals by a second member of the research team (HB). Neither of these authors was involved in the development of SPARX. Codes were then checked and refined with three other members of the research team (TF, KS, ML). Perceptions did not always neatly fit within one construct. Where a block of text appeared relevant to two constructs it was coded to both. The focus was not to quantify views relating to each construct but to capture important perceptions of the game according to participants. Within each construct, subnodes were established to organize perceptions into the four proposed design areas. This was either directly observable in the information or latent in the underlying experience.

Mapping User Perceptions to Design Elements

Within each construct, similar perceptions were then grouped into a common thread and entered into an Excel (2010) spreadsheet. The common threads were then mapped to the design element which was most explicit in the user perception.

Results

Creation of the Model

Themes reported in the various qualitative studies showed users valued the choices and control SPARX offered; the game was accessible to them when they wanted it, it protected their privacy, and it existed in a medium with which they were familiar. Engagement with the program arose from the playful medium, customizing their own character and, for those users who had completed the program, the sense they had benefited from the program and that the characters cared about them and gave them hope.

Psychological Constructs

SDT proposes that when people perceive they have more control over their treatment, a sense of competence in the activities and tasks required of them, and a sense of being cared for and connected with another, they will be more likely to integrate learning and behavior change [22]. SDT has been applied to psychotherapy, education, online learning, human motivation, and health [23-29]. The three constructs, autonomy, competence, and relatedness, are seen as central to an individual's sense of self and well-being and key motivators toward changing behavior. Supporting strategies such as positive regard, feedback, and structure facilitate motivation [34-40]. These three constructs also align well with important features in HAPA—self-efficacy (having confidence in performing tasks), social resources, and risk aversion.

As an established theory incorporating a motivational basis for effective change, SDT was likely to be a useful tool in understanding the perspectives of users and explaining the importance of supporting these with appropriate game design.

Serious Game Design Elements

Overview

The theories pertaining to the four areas we proposed as contributing to serious games were selected for their relevance to the strategies that were adopted intentionally by the developers of SPARX and to classification of potential design elements (Table 2).

Elements Contributing to Enjoyable Computer Games

Computer game play is a worldwide phenomenon with an increasingly diverse participant base. Players are engaged using a mix of rich graphics and audio to undertake challenging quests or explore alternative worlds. The key design features that influence player enjoyment have been explored [41,42], and taxonomy of six design elements has been proposed and empirically tested [43,44].

Elements Relating to the Accessibility of Online Content

In the 1990s, the World Wide Web Consortium launched the Web Accessibility Initiative to advise web builders on strategies to enable equitable access to web sites. The Web Content Accessibility Guidelines 2.0 are the current standard and emphasizes testable principles embracing dynamic, rich environments to ensure Web content is more accessible [45]. While specifically applicable to web content, it provides a useful standard for design or evaluation of accessibility.

Elements of the Therapeutic Relationship

In examining the evidence for a link between a positive therapeutic alliance and effective outcomes for patients in community mental health services, Howgego and colleagues summarized historical development and contemporary theory constructs [46]. Bordin's working alliance describes the relationship between a person seeking change and the change agent as one of the keys to the change process [47]. Fundamental to success is the active role clients play and the degree to which they value and believe in the purpose of the intervention and the process by which to attain change. While Bordin proposed generalizability of the working alliance model, it has been adopted predominantly in mental health, with a number of studies linking effective patient outcomes to a positive therapeutic alliance [46].

Elements Contributing To Learning Through Immersion

Existing theories of learning have been drawn upon to support the emerging pedagogy of online learning, supplemented with factors unique to this medium. Situated learning is a constructivist social learning theory [48] based on participatory knowledge acquisition from authentic contexts requiring activity, expert guidance, modeling of behavior, and a community of practice. Newcomers conduct simple, low-risk tasks, becoming familiar with language and organizing principles, and mature through a more active central role. Within an immersive

XSL•FO
RenderX

interface, interaction of the participant's avatar with other virtual characters can simulate a problem-solving community.

Immersion refers to the impression the user is participating in a real experience despite the user inherently knowing that some of the situations are not just unreal but impossible. Sensory, actional, and symbolic factors strengthen the degree of immersion [49]. Sensory immersion replicates digitally the experience of being in a three-dimensional space, using different camera angles, surround sound, motion, or vibration. Actional immersion allows the participant to initiate actions impossible in the real world (for example, flying on a bird), which intensifies the degree of user concentration. Symbolic immersion involves triggering powerful semantic psychological associations via the content of an experience.

Other ways which have been shown to enhance educational outcomes in an immersive environment are multiple perspectives and simulation of the real world [33]. Changing a participant's view, or frame of reference, from the exocentric (providing an external view of an object or thing) to the egocentric (providing an internal view of the object or thing) strengthens actional immersion and motivation through embodied, concrete learning [33]. Transfer refers to the ability to conceptualize knowledge learned in a way that allows an individual to use it in practical, real-life environments. Simulated learning environments which reflect the real world are thought to aid in transfer of knowledge more than, for example, passive face-to-face counseling sessions or classroom activities.

XSL•FO
RenderX

Table 2. Design elements of computer game, accessibility, working alliance, and learning in immersion.

Area	Elements	Description
Computer Game		
	Challenge	Overcoming obstacles and challenges to master and beat games
	Companionship	Socializing and cooperating with others
	Exploration	Unfamiliar places, hidden things, different strategies, discovery
	Fantasy	Enjoyment of fantasy worlds, fictional characters, unreal abilities, imaginary creatures, role-playing
	Fidelity	Realistic 3D graphics, animation, sounds
Accessibility		
	Perceivable	Content is easy to see, hear; presented in different ways
	Operable and navigable	Function available from keyboard; users can find content and navigate easily; creates no well-being issues
	Understandable	Text is readable and understandable; game operates in predictable ways
	Robust and reliable	Interpretation clear, minimal errors, secure, compatible
Working alliance		
	Goal	The outcome the person values and perceives will improve their circumstances
	Tasks	A series of relevant and efficacious tasks, which the person perceives will assist in attaining the goal
	Bond	Positive personal attachments such as trustworthiness, confidence, expertness, attractiveness, acceptance, empathy, nonjudgmental, and sincerity
Learning in immersion		
	Activity	Through tasks, puzzles, movement, and feedback
	Expert guidance	Builds on and adjusts existing knowledge
	Modeling of behavior	Behaviors learned by observation and modeling
	Community of practice	Where newcomers conduct simple, low-risk tasks, becoming familiar with language and organizing
	Sensory, actional, and symbolic factors	Replicates the experience of being in a 3D space; participant able to initiate actions; content triggers semantic responses
	Multiple perspectives	Changing a participant's view of an object from external to internal
	Simulation of the real world	Aids transfer of knowledge from conceptual to real-life

Confirmatory Testing of the Model Using Qualitative Data

Coding User Perceptions

User perceptions mapped well to the broad constructs of autonomy, competence, and relatedness and to the four groups of computer game, accessibility, working alliance, and learning in immersion. Overall, the content across the participant groups and interviews was analogous with very similar quotes evident in the data, suggesting both a level of saturation and a high level of agreement regarding important design elements.

Perceptions of Autonomy

All groups appreciated accessing a recommended treatment without having to be referred to a counselor or other mental health service. Reasons varied from the cost or unavailability

XSL•FO
RenderX

of services, stigma associated with getting help for mental health issues, having to tell someone else about problems, not having the confidence to talk about things, not being able to make sense of or find words to describe their feelings, or feeling embarrassed having to tell others. Users liked being able to talk or reveal their feelings to the computer without fear of being judged as well as learning skills and activities they could choose from and apply in their own lives.

The utility of having SPARX on their own computer, in their own space to play at a time that was convenient for them was central to their sense of control and protecting their privacy. Options to choose and personalize a character or avatar allowed individuals to express themselves as someone other than their own actual identity in a world other than their own. Several participants did not favor a computer game to access help, preferring face-to-face counseling instead. Furthermore, some participants did not have ready access to a computer. Examples of mapped participant quotes are provided in Textbox 1.

Textbox 1. Example quotes of user perceptions of autonomy.

Computer Game:

I think it is a fairly good idea to have it so that you can customize it. It is a way of expressing yourself. [Rainbow 2/9]

It's a sort of a Warcraft thing where you pick your characters and go into a whole different world. I love fantasy. [AE (User) 4/39]

Accessibility:

It was good because I felt that I had the control in that it was something that I could just do by myself.

Researcher then asks,

How could you ensure that you had control over it?

Well, seeing as it was on the computer I could just put it in my own file and no one else would go there and they would leave it to just me. It was good. [Rainbow (User) 25/25]

Working Alliance:

So you are learning it from a computer game but it is still a really recommended thing. [Rainbow (User) 3/25]

They don't really have to talk to an actual person about it, and that way they don't have to worry about getting judged with the feedback and stuff like that. [Aus 2/16]

You have more control [with SPARX compared to counselor]. You can't just leave a counselor that you don't like. [AE (User) 1/39]

Learning in Immersion:

I wanted to do it this way. I wouldn't have liked being told what to do. [AE (User) 2/39]

Perceptions of Competence

Quests to unknown worlds to accomplish tasks and collect gems as rewards were perceived as fun. While a sense of accomplishment was reported when puzzles and challenges were completed, the level of challenge difficulty experienced by participants varied greatly—from too easy to too hard.

Ease of operation, predictability of controls and actions, and content that was easy to see and hear supported a feeling of competence. Conversely, users were critical and frustrated when, for example, content was incompatible with their web browser or operating system. Users valued and gave examples of learning skills they could and had used in real life, conferring a real sense of achievement. Modeling of behavior was evident in the way participants described their learning experiences using SPARX.

Observing game characters going through problems similar to their own and helping the characters to overcome these in the game externalized the problem for users and instilled confidence to try these strategies in the real world. Where participants found the challenges too easy they felt a loss of engagement and subsequently thought that the program should be for a younger audience. More interaction was suggested as a strategy to combat this. Examples of mapped participant quotes are provided in Textbox 2.

Textbox 2. Example quotes of user perceptions of competence.

Computer Game:

Cass [a character in SPARX] was cool because I felt cool helping her, you felt good. [Rainbow (User) 16/25]

Accessibility:

The one thing I really, really hated about it was—you know how when you go to move your character you have to click. I reckon it would be better if you could use the arrow keys. [Rainbow (User) 3/25]

XSL•FO
RenderX

When I was in the Ice-land level and I spent five minutes looking for the Yeti and it was right in front of me but I didn't notice it because it was the same sort of white as the rest. [Rainbow (User) 7/25]

Working Alliance

It gives good advice. It actually gives you real life techniques and skills that you can use and that are easy to use. And you are able to use them in everyday life. [Māori (User) 2/5]

I was not really depressed at the beginning—I was just really angry—but it was still useful, I am less angry now…I don't hate.

The researcher then asked,

Is this different from before you did SPARX?

Yes…I used the take 10 seconds, walk away…it has changed me. My Dad is happy with the changes, he is proud of me changing. [AE (User) 6/39]

I think the negotiating one to solve it or sort it, about how you have to think half-half kind of thing. I think that works really well in school. I have used it on my teacher. [Rainbow (User) 8/25]

I haven't had a fight since it started. I'm not getting into trouble since SPARX. I learn more now, concentrating in class. SPARX taught me confidence. [AE (User) 8/39]

Learning in Immersion:

When the bird comes out of the box it always speaks in a nice way—doesn't shout. You have to practice saying things in a good way. [AE (User) 7/39]

Perceptions of Relatedness

The richness of the computer game interface was evident in comments from the participants. For instance, many people commented that the characters were likeable, the 3D graphics were appealing, and the fantasy-based program was valuable. Fun was expressed as a feature of computer games in general: the settings, adventures, quests, and different characters in the game. Users related to various and different fantasy characters in SPARX. Being able to represent abstract concepts as something real within a fantasy computer game genre was also perceived as enjoyable. Users identified preferential character qualities including attractiveness, expertise, empathy, warmth, and sincerity. Immersive factors helped to engage participants by making them feel like they were part of the program. It was acknowledged the program would not suit everyone. A small number of users reported the language was too simple and that, by extension, they considered the characters patronizing. Examples of mapped participant quotes are provided in Textbox 3.

Textbox 3. Example quotes of user perceptions of relatedness.

Computer Game:

It is really cool. It is like a real life thing in an imaginary world so it is really fun and more engaging. Because if it had real people it would be a bit boring. [Māori (User) 2/5]

I really liked Hope [a talking character in the form of a bird]. I thought that was a cool idea…that hope was something tangible. [Rainbow (User) 10/25]

Accessibility:

I am not paying much attention to the words in the box for some reason. I think I am more of an audio person. [Māori 1/26]

Working Alliance:

The bird is cute—Hope. Yes, that was my little favorite thing on SPARX—Hope.

The researcher then queried,

What was it about the bird of hope that really stuck with you? Because she said "I am always here to help you." It was so cute and just the color of her. She looks so beautiful. I wish she was my pet. [Māori (User) 3/5]

When he [the guide character] asks how I've been—that was good. I like that someone cares. [AE 9/39]

And it never said you were wrong, it just said maybe try another way or that sort of thing.

The researcher then queried,

Why do you think that was important?

Well, if people keep saying you are wrong, you are going to give up. [Rainbow (User) 7/25]

XSL•FO
RenderX

> *I thought one of them was good looking plus the voice over was…but, no, it is a game.* [Rainbow (User) 20/25]
>
> Learning in Immersion:
>
> *You know the Gnats [Gloomy Negative Automatic Thoughts]? They sounded like Voldemort which I thought was helpful because they are evil.* [Rainbow (User) 1/25]
>
> *It was cool. And you can feel as if you are in the actual game itself.*
>
> The researcher then queried,
>
> *How would you describe that process of being in it?*
>
> *Well, with the guide you felt like he was talking to you and not to a character as other games do. And you are controlling the person—walking by itself and that type of thing.* [Māori (User) 2/]

Mapping User Perceptions to Design Elements

Similar user perceptions were then grouped. For example, comments relating to how participants could use the program independently had been linked to the SDT construct autonomy, and where these were a function of the utility of the program, they were organized into the subnode accessibility. Similar perceptions of this were grouped into the thread "I can use the program where I want, when I want, how I want."

While the perceptions most explicitly identified features of the serious game from within the group to which the perceiving thread was organized, at times other design elements were latent in user perceptions. For example, users perceived gaining skills, based on CBT tasks, that they could use in real life. Those user perceptions were linked to the competence construct and organized within the working alliance group. The supporting features of the serious game most evident were the practical CBT skills based within the design element of tasks of the working alliance. While this mapping is most obvious in the user perception, to teach this skill the game drew upon other elements which were less explicit in the participant data. These included working alliance factors (encouragement and feedback) learning pedagogy (initial learning of deep breathing, expert guidance about when and how to use it, modeling the activity, opportunity to practice the skill), multiple perspectives (virtual therapist) and immersive factors (actional and sensory); computer gaming (fantasy world characters, realistic animation, graphics, and sound); and accessibility (clear, perceivable content presented in both audio and text format, available widely through online or portable media). These elements were noted but not mapped directly to that perceiving thread.

We present the model of design elements in Figure 4, with the groups computer game, accessibility, competence, and relatedness color-coded orange, red, green, and blue, respectively. The design elements are presented for autonomy (Figure 5), competence (Figure 6) and relatedness (Figure 7) with the four groups color-coded to correspond to the main diagram (Figure 4) for interpretation. The perceiving threads are those expressed by users in the data. The supporting features are those identified in the four groups in the model and found in the game. The design elements that are not mapped directly to any user perceptions in that SDT construct were noted as supporting but were more latent in user perceptions such as those described in the CBT task above.

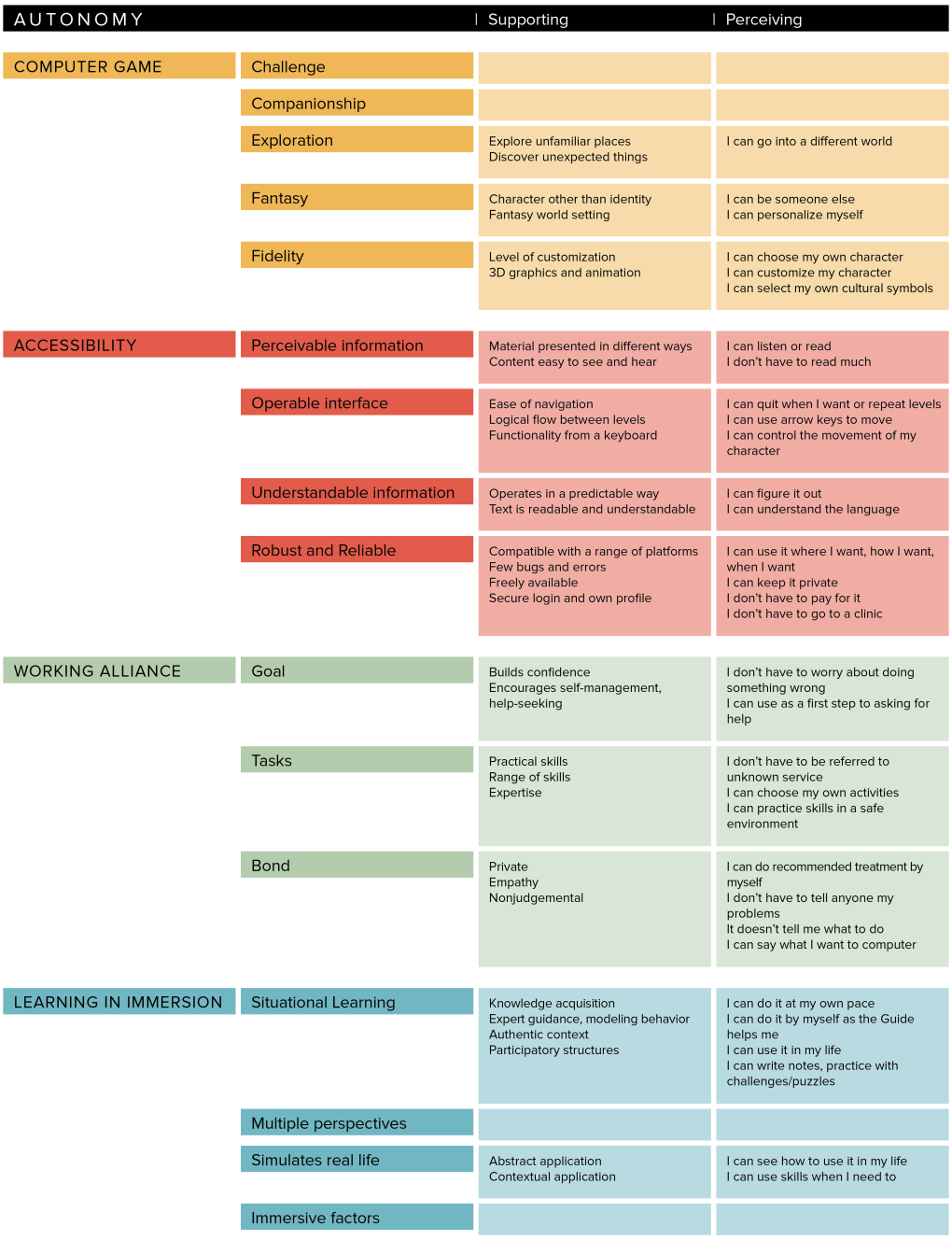
XSL•FO
RenderX

Figure 4. Serious game design elements.



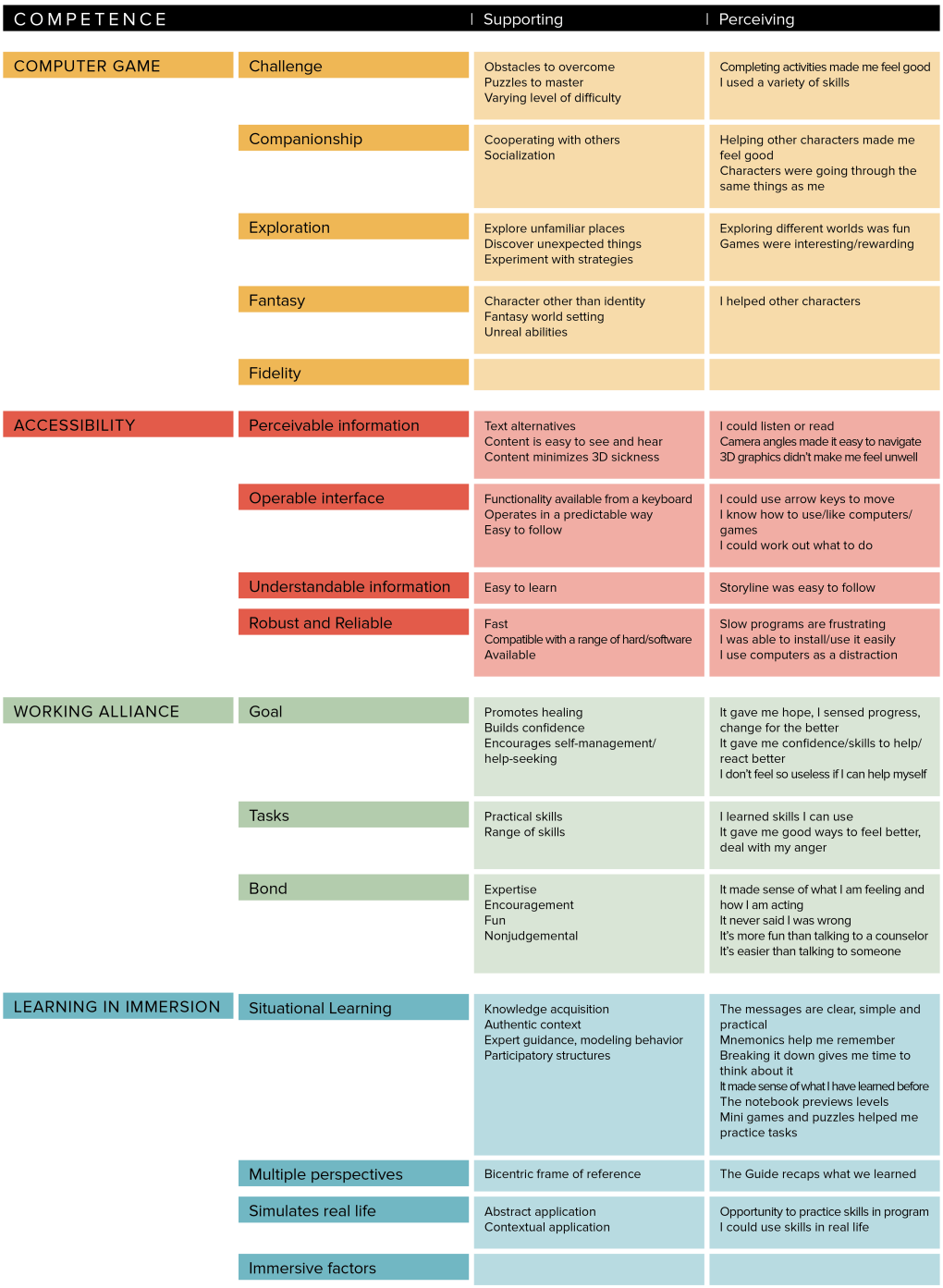
XSL•FO
RenderX

Figure 5. User perceptions of autonomy mapped to supportive features of a serious game.



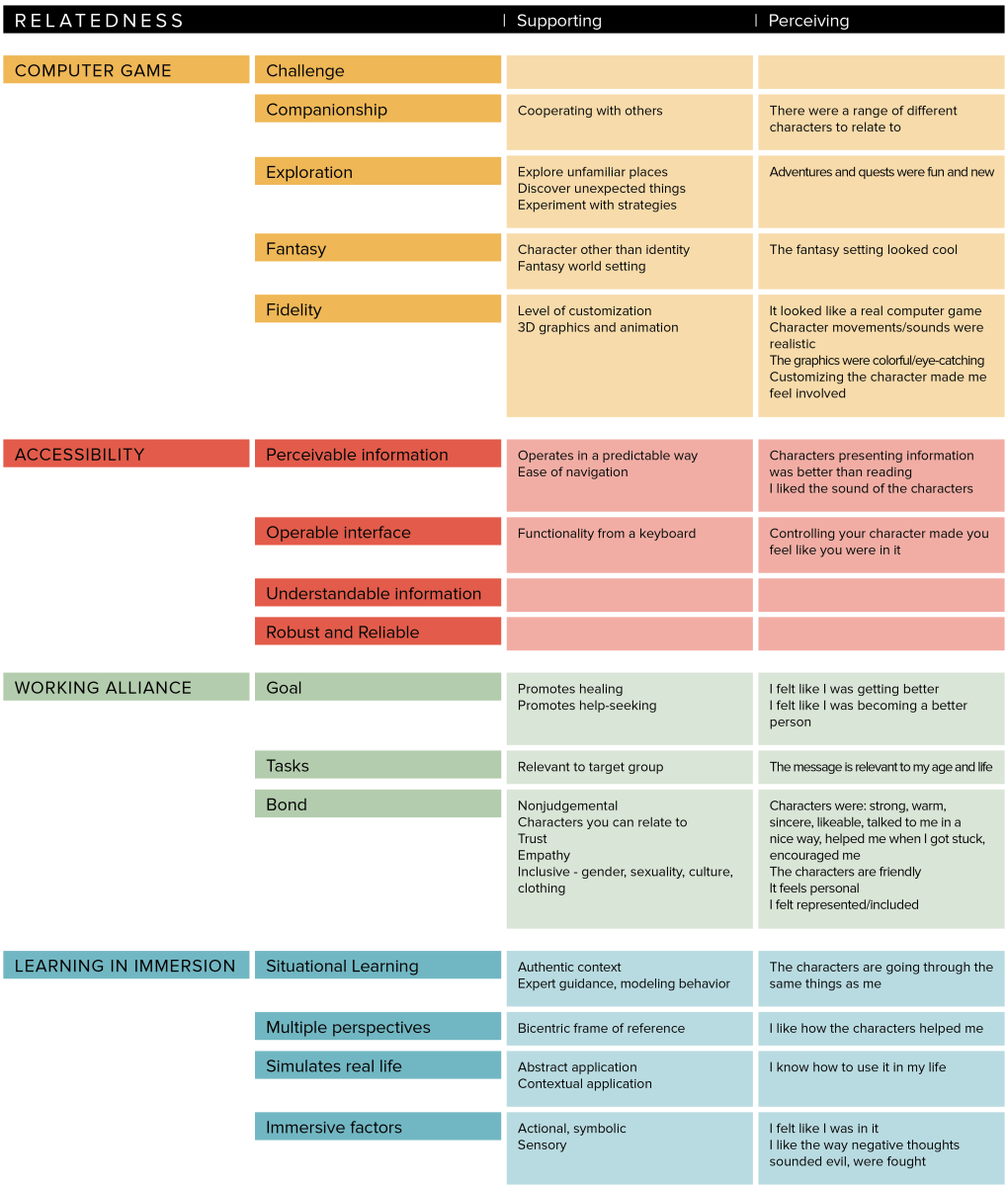
XSL•FO
RenderX

Figure 6. User perceptions of competence mapped to supportive features of a serious game.



XSL•FO
RenderX

Figure 7. User perceptions of relatedness mapped to supportive features of a serious game.



XSL•FO
RenderX

Discussion

Self-Determination Theory and Serious Game Design

The serious game designer faces a complex task in marrying gaming elements with therapeutic and learning goals without compromising either. The groups of design elements shown in Figure 4 are strongly supported by and map well against the user perceptions outlined in our findings and represent a potentially useful tool for serious game designers. By further mapping these elements against the constructs of SDT (Figures 5-7), we provide a theoretical dimension to the model that opens up promising avenues for future exploration.

A major problem facing serious game designers is that while users have strong ideas about what will work for them, they are generally unable to consciously synthesize and/or articulate these requirements for the designer. The framework and model presented in this paper articulate the context of user needs and will assist developers to bridge the cultural divide and resolve some of the tensions between themselves and the therapy and pedagogy experts whose input will be vital in designing a serious game that meets therapeutic and learning goals. Articulating young people's views of the game SPARX and matching these to supportive design features to improve the quality of community-based mental health interventions is important as over 75% of mental disorders commence before the age of 25 years [50].

HAPA distinguishes between the processes that motivate people to change (pre-intentional stage) and the processes that lead to the actual health behavior (volition stages) [25]. During the first stage, self-efficacy, outcome expectancies and risk perception affect motivation. During the second stage, people need detailed instruction on how to perform the desired action, and they must be confident that they can accomplish this. Thus, the constructs of SDT are important considerations at these stages of health intervention. The SDT framework presented contextualizes how young people perceived these factors and the serious game design elements which supported them.

Young people want help for mental health issues, but our findings reveal that there is also a strong need to be able to control how they access services or get help. This sense of autonomy at the pre-intentional stage was engendered by having recommended therapy available to them without having to be referred to or attend a clinic, having a program on their computer so they could use it where and when they wanted, having choices to personalize their avatar, and being able to select activities which were relevant to their life. These preferences regarding choice and control parallel normal developmental trajectories of exploring independence, autonomy, and identity during adolescence [51]. Given the transformative opportunity adolescence presents in biological, psychological, and social domains [52], matching any appropriate mental health intervention to the developmental concerns of this phase is crucial for positive treatment outcomes.

A sense of competence was supported by design elements from all four groups. The efficacy of the tasks and the extent to which they made the individual feel like they were improving was evident in powerful perceptions in the data of healing and gaining confidence. These facets of the working alliance seemed to be the strongest contributors to a sense of competence, yet more latent in perceptions was the deliberate application of learning pedagogy to the organization and presentation of content to impart knowledge in a more meaningful way. The exocentric virtual therapist provided observer perception and reflection, fostering more abstract, symbolic insights to help the user separate the problem from the person. An egocentric game component, by way of accomplishing a series of user tasks, enabled participants' actional immersion and motivation through embodied, concrete learning [33]. Since most people have a sense of what makes a consistent fantasy world, players already have a level of fantasy world competence, while developers have a ready platform on which to build [53]. In SPARX, fantasy was an important tool: users valued externalizing their identity and enjoyed the unreal abilities of characters. The genre enabled developers to represent abstract concepts as concrete entities and participants to practice skills using immersive factors and perspective, which have been found to enhance educational outcomes [49]. Employing the most effective ways for users to navigate and control their interactions and movements are all important aspects of accessibility. The deliberate combination of strategies within the playful medium enriched the experience for the user and is likely to contribute to motivation at both the pre-intentional and volition stages.

Nevertheless, users had diverse views about the level of challenge within the game, implying it is unlikely one serious game will suit all and personalization may be needed. This aligns with the principles of HAPA, where predictors operate differently on those in one stage group compared with those in an adjacent stage group [25].

Relatedness support is not specifically addressed in the studies of SDT and online education; Sorebo and Haehre found no association between perceived relatedness and students' levels of intrinsic motivation in educational computer games [37]. As issues of attrition challenge the efficacy of online therapies for mental health, incorporation of therapist support or active therapy ingredients is recommended [4,53,54]. Social support is considered a resource applicable at all stages in HAPA; lack of it can be a barrier to adopt or maintain health behaviors [25], and this relationship between the individual and the change agent is emphasized in Bordin's working alliance [47]. The SPARX program was designed to be used independent of therapists but incorporated a guide as a virtual therapist. The text was chosen carefully to be therapeutic; the image of a powerful and supportive character (Figure 1) and an actor with a warm encouraging voice were specifically selected for the guide. User perceptions that the guide and other characters made them feel supported and cared for suggests that relatedness support can be achieved in serious games and it is perceived as helpful by users.

Whether serious games maintain the positive effects thus far described in efficacy studies has yet to be determined as most of the peer-reviewed literature describes open or randomized controlled studies where a short window of engagement exists with participants. Incorporating ways to get further help with a positive user experience (one in which the goal, sense of control,

XSL•FO
RenderX

ability, and healing are valued) might promote further help-seeking behavior.

Opportunities for Further Research

Users compared the experience of the serious game with commercial computer games. This may be an unrealistic expectation for serious game developers given the differences in development funding and the business models supporting the use of commercial games. While computer game users play for fun, serious games for mental well-being are targeting a specific and personal health-related goal. If the serious game sufficiently motivates the user to work toward a valued goal, the reliance on the serious game to entertain might be tempered. How the element of fun mediates user motivation is unknown in this context at this time. Tools to measure motivation have been used in online learning [34]. Measuring user motivation toward various serious games for mental health could assist in validating the critical design elements and further inform development.

It is likely that the relative importance of design elements will change as the goals of treatment, the target user audience, and the way in which a serious game is implemented vary. For example, we found participants in these groups and interviews were very keen to ensure their privacy, while Lederman and colleagues describe participants who had suffered psychosis valuing the peer support offered by an online social therapy tool [55]. This emphasizes the importance of consulting with potential users of the program during analysis and design.

Limitations

This study examines the user perceptions of youth predominantly aged 12 to 19 years for one program, SPARX. The gaming elements among serious games for mental health issues differ. While the theories selected to support the qualitative data for this study were also selected for their generalizability, it is not presumed this model will be generalizable to other age groups, populations, or interventions. While once the domain of young people, computer gaming is gaining popularity among broader age groups and different cultures (in 2014, the average age of an Australian computer gamer was 32 years; 47% are female and 19% are older than 51 years [56]). It will be interesting to test this model against other age ranges and populations.

The divergent themes reinforce that this serious game does not suit everyone. Developing interventions which explicitly target one or more stages of HAPA may help us understand whether interventions engage people at one stage rather than another, whether the medium of gaming suits everyone, or whether a different set of constructs applies to their requirements. It was also clear the level of challenge varied among participants, and when tasks or language were considered too difficult or too easy, there was a loss of engagement. Cognitive capacity is thought to moderate the impact of treatment approaches, but it is unclear whether this is a factor of the cognitive capacity of the user or the way content is presented [49]. Some people would rather get treatment via traditional modes of therapy, specifically face-to-face therapy; others may not have ready access to a computer or the Internet. In general, young people who identified as gamers indicated a higher level of engagement with the concept of SPARX; however, it is currently unclear how important the design features are in engaging young people who don't normally play computer games or who are reluctant to engage in any mode of mental health intervention or support. The focus group and interview participants represented some of the people most underserved by mental health services in the community. While their perspectives are unique, they may not be representative of the overall adolescent population.

Given four of the authors were codevelopers of SPARX, the risk of bias is inherent. The remaining authors were independent of the development of SPARX; members of this group conducted the initial thematic analysis and validation of codes.

Conclusions

Mental health issues affect a large number of people, many of whom will not access care through traditional models of care. Serious games offer a means of extending the reach of evidence-based early intervention, but they need to be well designed to deliver therapy in a way that engages users and helps them. The methods used in this study allowed articulation of design elements from a user-centered perspective in a structured framework. The framework and model may provide a guide for developers to ensure programs support important user-centered requirements. The relative importance of the various design elements is likely to vary with the purpose of the serious game and goals of treatment. Involving users in development is imperative if serious games are to be fit for purpose.

Acknowledgments

We would like to thank all the young people who participated in these studies. This study was supported by the University of Tasmania Rural Clinical School. The authors would like to acknowledge Professor Sally Merry and the team at the University of Auckland's Werry Center for Child and Adolescent Mental Health in New Zealand in developing SPARX and making it available for research. Graphic design by Christiane King, Digital Innovation Team, University of Tasmania.

Conflicts of Interest

TF, ML, KS, and MS are codevelopers of SPARX; the intellectual property for SPARX is owned by Uniservices at the University of Auckland, and the developers can benefit financially from its commercialization.

XSL•FO
RenderX

Multimedia Appendix 1

Table 2: Interview Participant data-extended.

[PDF File (Adobe PDF File), 112KB - mental_v2i2e11_app1.pdf]

References

1. World Health Organisation. Mental health action plan 2013-2020. In: Mental health action plan. In. Geneva, Switzerland: WHO Document Production Services; 2013.
2. National Institute for Clinical Excellence (NICE). Clinical guidelines CG90. Depression : The treatment and management of depression in adults (update) URL: http://aspe.hhs.gov/sp/nhii/documents/NHIIReport2001/report5.htm [accessed 2015-01-23] [WebCite Cache ID 6VmuHAWgs]
3. Calear AL, Christensen H. Review of internet-based prevention and treatment programs for anxiety and depression in children and adolescents. Med J Aust 2010 Jun 7;192(11 Suppl):S12-S14. [Medline: 20528700]
4. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev 2012 Jun;32(4):329-342. [doi: 10.1016/j.cpr.2012.02.004] [Medline: 22466510]
5. Melville KM, Casey LM, Kavanagh DJ. Dropout from Internet-based treatment for psychological disorders. Br J Clin Psychol 2010 Nov;49(Pt 4):455-471. [doi: 10.1348/014466509X472138] [Medline: 19799804]
6. Coyle D, Doherty G, Sharry J. An evaluation of a solution focused computer game in adolescent interventions. Clin Child Psychol Psychiatry 2009 Jul;14(3):345-360. [doi: 10.1177/1359104508100884] [Medline: 19515752]
7. Merry SN, Stasiak K, Shepherd M, Frampton C, Fleming T, Lucassen MFG. The effectiveness of SPARX, a computerised self help intervention for adolescents seeking help for depression: randomised controlled non-inferiority trial. BMJ 2012 Apr 19;344(apr18 3):e2598-e2598. [doi: 10.1136/bmj.e2598] [Medline: 3330131]
8. Khanna MS, Kendall PC. Computer-assisted cognitive behavioral therapy for child anxiety: results of a randomized clinical trial. J Consult Clin Psychol 2010 Oct;78(5):737-745. [doi: 10.1037/a0019739] [Medline: 20873909]
9. Shandley K, Austin D, Klein B, Kyrios M. An evaluation of 'Reach Out Central': an online gaming program for supporting the mental health of young people. Health Educ Res 2010 Aug;25(4):563-574 [FREE Full text] [doi: 10.1093/her/cyq002] [Medline: 20150530]
10. Rizzo A, Newman B, Parsons T. Development and clinical results from the virtual iraq exposure therapy application for PTSD. In: Virtual Rehabilitation International Conference. In: Virtual Rehabilitation International Conference, 2009: IEEE; 2009 Presented at: Virtual Rehabilitation International Conference; June 29 2009-July 2 2009; Haifa p. 8-15. [doi: 10.1109/ICVR.2009.5174198]
11. Fleming T, Cheek C, Merry S, Thabrew H, Bridgman H, Stasiak K, et al. Serious games for the treatment or prevention of depression: a systematic review. evista de Psicopatología y Psicología Clínica - Spanish Journal of Clinical Psychology 2015 Jan 15;19(3):227-242. [doi: 10.5944/rppc.vol.19.num.3.2014.13904]
12. Dunn DS, Elliott TR. The Place and Promise of Theory in Rehabilitation Psychology. Rehabil Psychol 2008 Aug;53(3):254-267 [FREE Full text] [doi: 10.1037/a0012962] [Medline: 19649146]
13. Noar SM, Benac CN, Harris MS. Does tailoring matter? Meta-analytic review of tailored print health behavior change interventions. Psychol Bull 2007 Jul;133(4):673-693. [doi: 10.1037/0033-2909.133.4.673] [Medline: 17592961]
14. Limerick H, Coyle D, Moore J. The experience of agency in human-computer interactions: A review. Front Hum Neurosci 2014;8:643. [doi: 10.3389/fnhum.2014.00643]
15. Knowles SE, Toms SG. Qualitative Meta-Synthesis of User Experience of Computerised Therapy for Depression and Anxiety. PLoS one 2014 Jan;9(1):e84323. [doi: 10.1371/journal.pone.0084323]
16. Cayton H. The flat-pack patient? Creating health together. Patient Educ Couns 2006 Sep;62(3):288-290. [doi: 10.1016/j.pec.2006.06.016] [Medline: 16876378]
17. Detmer DE. Building the national health information infrastructure for personal health, health care services, public health, and research. BMC Med Inform Decis Mak 2003 Jan;3(1):1 [FREE Full text] [doi: 10.1186/1472-6947-3-1]
18. Hesse BW, Hansen D, Finholt T, Munson S, Kellogg W, Thomas JC. Social Participation in Health 2.0. Computer (Long Beach Calif) 2010 Nov 11;43(11):45-52 [FREE Full text] [doi: 10.1109/MC.2010.326] [Medline: 21379365]
19. Ryan RM, Deci EL. Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. Am Psychol 2000 Jan;55(1):68-78. [Medline: 11392867]
20. Deci E, Ryan R. Self-determination theory: A macrotheory of human motivation, development, and health. Canadian Psychology/Psychologie canadienne 2008;49(3):182-185. [doi: 10.1037/a0012801]
21. Bandura A. Health promotion by social cognitive means. Health Educ Behav 2004 Apr;31(2):143-164. [doi: 10.1177/1090198104263660] [Medline: 15090118]
22. Rogers RW. A protection motivation theory of fear appeals and attitude change. J Psychol 2010 Jul 02;91(1):93-114. [doi: 10.1080/00223980.1975.9915803]
23. Schwarzer R, Lippke S, Luszczynska A. Mechanisms of health behavior change in persons with chronic illness or disability: the Health Action Process Approach (HAPA). Rehabil Psychol 2011 Aug;56(3):161-170. [doi: 10.1037/a0024509] [Medline: 21767036]

XSL•FO
RenderX

24. Prochaska J, Velicer WF. The transtheoretical model of health behavior change. Am J Health Promotion 1997 Sep;12(1):38-48. [doi: 10.4278/0890-1171-12.1.38]
25. Schwarzer R. Self-efficacy in the adoption and maintenance of health behaviors: Theoretical approaches and a new model. In: Self-Efficacy: Thought Control Of Action. New York: Taylor & Francis; 1992.
26. Marne B, Wisdom J, Huynh-Kim-Bang B, Labat JM. A design pattern library for mutual understanding and cooperation in serious game design. In: Intelligent tutoring systems. New York: Springer; 2012:135-140.
27. Arnab S, Lim T, Carvalho M, Bellotti F, de FS, Louchart S, et al. Mapping learning and game mechanics for serious games analysis. Brit J Educ Technol 2014. [doi: 10.1111/bjet.12113]
28. Fleming T, Lucassen M, Stasiak K, Shepherd M, Merry S. The impact and utility of computerised therapy for educationally alienated teenagers: The views of adolescents who participated in an alternative education-based trial. Clin Psychol 2015 Jan 05. [doi: 10.1111/cp.12052]
29. Shepherd M, Fleming T, Lucassen M, Stasiak K, Lambie I. The design and relevance of a computerised therapy program for indigenous Maori adolescents. JMIR Serious Games 2015;15(3):e1. [doi: 10.2196/games.3804]
30. Lucassen MFG, Hatcher S, Stasiak K, Fleming T, Shepherd M, Merry SN. The views of lesbian, gay and bisexual youth regarding computerised self-help for depression: An exploratory study. Advances in Mental Health 2014 Dec 17;12(1):22-33. [doi: 10.5172/jamh.2013.12.1.22]
31. Fleming TM, Dixon RS, Merry SN. 'It's mean!' The views of young people alienated from mainstream education on depression, help seeking and computerised therapy. Advances in Mental Health 2014 Dec 17;10(2):195-203. [doi: 10.5172/jamh.2011.10.2.195]
32. Cheek C, Bridgman H, Fleming T, Cummings E, Ellis L, Lucassen MFG, et al. Views of young people in rural Australia on SPARX, a fantasy world developed for New Zealand youth with depression. JMIR Serious Games 2014 Feb;2(1):e3 [FREE Full text] [doi: 10.2196/games.3183] [Medline: 25659116]
33. Dede C. Immersive interfaces for engagement and learning. Science 2009 Jan 2;323(5910):66-69. [doi: 10.1126/science.1167311] [Medline: 19119219]
34. Chen KC, Jang SJ. Motivation in online learning: Testing a model of self-determination theory. Comput Hum Behav 2010 Jul;26(4):741-752. [doi: 10.1016/j.chb.2010.01.011]
35. Standage M, Duda JL, Ntoumanis N. A test of self-determination theory in school physical education. Br J Educ Psychol 2005 Sep;75(Pt 3):411-433. [doi: 10.1348/000709904X22359] [Medline: 16238874]
36. Xie K, Debacker T, Ferguson C. Extending the traditional classroom through online discussion: The role of student motivation. J Educ Comput Res 2005 Jul 1;34(1):67-89. [doi: 10.2190/7bak-egah-3mh1-k7c6]
37. Sorebo O, Haehre R. Investigating students' perceived discipline relevance subsequent to playing educational computer games: A personal interest and self-determination theory approach. Scandinavian J Educ Res 2012;56:345-362. [doi: 10.1080/00313831.2011.594609]
38. Ryan R, Deci E. A self-determination theory approach to psychotherapy: The motivational basis for effective change. Canadian Psychology/Psychologie canadienne 2008;49(3):186-193. [doi: 10.1037/a0012753]
39. Schoech D, Boyas J, Black B, Elias-Lambert N. Gamification for Behavior Change: Lessons from Developing a Social, Multiuser, Web-Tablet Based Prevention Game for Youths. J Technol Hum Services 2013 Jul;31(3):197-217. [doi: 10.1080/15228835.2013.812512]
40. Vallerand R, Pelletier LG, Blais M, Briere N, Senecal C, Vallieres E. The Academic Motivation Scale: A Measure of Intrinsic, Extrinsic, and Amotivation in Education. Educ Psychol Meas 1992 Dec 01;52(4):1003-1017. [doi: 10.1177/0013164492052004025]
41. Hunicke R, LeBlanc M, Zubek R. MDA: A formal approach to game design and game research. In: Proceedings of the nineteenth national conference on Artificial Intelligence. 2004 Presented at: Association for the Advancement of Artificial Intelligence Conference; 2004; San Jose URL: http://www.aaai.org/Papers/Workshops/2004/WS-04-04/WS04-04-001.pdf
42. King D, Delfabbro P, Griffiths M. Video game structural characteristics: a new psychological taxonomy. Int J Ment Health Addiction 2009 Apr 7;8(1):90-106. [doi: 10.1007/s11469-009-9206-4]
43. Quick J, Atkinson R, Lin L. Confirming the taxonomy of video game enjoyment. In: Games, Learning, Society 8.0. Madison, WI: ETC Press; 2013.
44. Quick J, Atkinson R, Lin L. Empirical taxonomies of gameplay enjoyment: Personality and video game preference. Int J Game-Based Learning 2012:11-31. [doi: 10.4018/ijgbl.2012070102]
45. WCAG.: World Wide Web Consortium; 2008. Web content accessibility guidelines 2.0 URL: http://www.w3.org/TR/2008/REC-WCAG20-20081211/ [accessed 2015-02-11] [WebCite Cache ID 6WFiBmZnV]
46. Howgego IM, Yellowlees P, Owen C, Meldrum L, Dark F. The therapeutic alliance: the key to effective patient outcome? A descriptive review of the evidence in community mental health case management. Aust N Z J Psychiatry 2003 Apr;37(2):169-183. [Medline: 12656956]
47. Bordin E. The generalizability of the psychoanalytic concept of the working alliance. Psychotherapy: Theory, Research and Practice; 1979:252-260.
48. Lave J, Wenger E. Situated learning: legitimate peripheral participation. Cambridge [England]: Cambridge University Press; 1991.

XSL•FO
RenderX

49. Dede C. Theoretical perspectives influencing the use of information technology in teachinglearning. In: International handbook of information technology in primary and secondary education. New York: Springer Science+Business Media; 2008:43-62.
50. Hickie IB, McGorry PD. Increased access to evidence-based primary mental health care: will the implementation match the rhetoric? Med J Aust 2007 Jul 16;187(2):100-103. [Medline: 17635093]
51. MacLeod KB, Brownlie E. Mental Health and Transitions from Adolescence to Emerging Adulthood: Developmental and Diversity Considerations. Canadian Journal of Community Mental Health 2014 Jul;33(1):77-86. [doi: 10.7870/cjcmh-2014-007]
52. Weisz JR, Sandler IN, Durlak JA, Anton BS. Promoting and protecting youth mental health through evidence-based prevention and treatment. Am Psychol 2005 Sep;60(6):628-648. [doi: 10.1037/0003-066X.60.6.628] [Medline: 16173895]
53. Ryan S, Slavatore K, Green J, Jongewaard D. Online evolution: is there life beyond elves and dwarves? Computer Gaming World 258 2006;1.
54. Spek V, Cuijpers P, Nyklícek I, Riper H, Keyzer J, Pop V. Internet-based cognitive behavioural therapy for subthreshold depression in people over 50 years old: a randomized controlled clinical trial. Psychol Med 2007 Mar;37(3):319-328. [doi: 10.1017/S0033291707000542] [Medline: 17466110]
55. Lederman R, Wadley G, Gleeson J, Bendall S, Álvarez-Jiménez M. Moderated online social therapy. ACM Trans. Comput.-Hum. Interact 2014 Feb 01;21(1):1-26. [doi: 10.1145/2513179]
56. Brand JE, Lorentz P, Mathew T. Digital Australia. 2014. Interactive games and entertainment association URL: http://igea.wpengine.com/wp-content/uploads/2013/11/Digital-Australia-2014-DA14.pdf [accessed 2015-01-27] [WebCite Cache ID 6VsahDXsb]

Abbreviations

CBT: cognitive behavioral therapy
cCBT: computerized cognitive behavioral therapy
HAPA: health action process approach
SDT: self-determination theory
SPARX: smart, positive, active, realistic, X-factor thoughts

Edited by G Eysenbach; submitted 10.12.14; peer-reviewed by S Knowles, C Hawke; comments to author 30.12.14; revised version received 11.02.15; accepted 16.03.15; published 21.04.15.

Please cite as:
Cheek C, Fleming T, Lucassen MFG, Bridgman H, Stasiak K, Shepherd M, Orpin P
Integrating Health Behavior Theory and Design Elements in Serious Games
JMIR Mental Health 2015;2(2):e11
URL: http://mental.jmir.org/2015/2/e11/
doi:10.2196/mental.4133
PMID:26543916

©Colleen Cheek, Theresa Fleming, Mathijs FG Lucassen, Heather Bridgman, Karolina Stasiak, Matthew Shepherd, Peter Orpin. Originally published in JMIR Mental Health (http://mental.jmir.org), 21.04.2015. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Identifying Chinese Microblog Users With High Suicide Probability Using Internet-Based Profile and Linguistic Features: Classification Model

Li Guan[1,2], MS; Bibo Hao[2], MS; Qijin Cheng[3], PhD; Paul SF Yip[3], PhD; Tingshao Zhu[1,4], PhD

[1]Key Lab of Behavioral Science of Chinese Academy of Sciences, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

[2]University of Chinese Academy of Sciences, Beijing, China

[3]HKJC Center for Suicide Research and Prevention, The University of Hong Kong, Hong Kong SAR, China (Hong Kong)

[4]Key Lab of Intelligent Information Processing of Chinese Academy of Sciences, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China

Corresponding Author:
Tingshao Zhu, PhD
Key Lab of Behavioral Science of Chinese Academy of Sciences
Institute of Psychology
Chinese Academy of Sciences
Room 821, Building He-xie, 16th Lincui Road
Chaoyang District
Beijing, 100101
China
Phone: 86 15010965509
Fax: 86 010 64851661
Email: tszhu@psych.ac.cn

Abstract

Background: Traditional offline assessment of suicide probability is time consuming and difficult in convincing at-risk individuals to participate. Identifying individuals with high suicide probability through online social media has an advantage in its efficiency and potential to reach out to hidden individuals, yet little research has been focused on this specific field.

Objective: The objective of this study was to apply two classification models, Simple Logistic Regression (SLR) and Random Forest (RF), to examine the feasibility and effectiveness of identifying high suicide possibility microblog users in China through profile and linguistic features extracted from Internet-based data.

Methods: There were nine hundred and nine Chinese microblog users that completed an Internet survey, and those scoring one SD above the mean of the total Suicide Probability Scale (SPS) score, as well as one SD above the mean in each of the four subscale scores in the participant sample were labeled as high-risk individuals, respectively. Profile and linguistic features were fed into two machine learning algorithms (SLR and RF) to train the model that aims to identify high-risk individuals in general suicide probability and in its four dimensions. Models were trained and then tested by 5-fold cross validation; in which both training set and test set were generated under the stratified random sampling rule from the whole sample. There were three classic performance metrics (Precision, Recall, F1 measure) and a specifically defined metric "Screening Efficiency" that were adopted to evaluate model effectiveness.

Results: Classification performance was generally matched between SLR and RF. Given the best performance of the classification models, we were able to retrieve over 70% of the labeled high-risk individuals in overall suicide probability as well as in the four dimensions. Screening Efficiency of most models varied from 1/4 to 1/2. Precision of the models was generally below 30%.

Conclusions: Individuals in China with high suicide probability are recognizable by profile and text-based information from microblogs. Although there is still much space to improve the performance of classification models in the future, this study may shed light on preliminary screening of risky individuals via machine learning algorithms, which can work side-by-side with expert scrutiny to increase efficiency in large-scale-surveillance of suicide probability from online social media.

(JMIR Mental Health 2015;2(2):e17) doi:10.2196/mental.4227

XSL•FO
RenderX

KEYWORDS

suicide probability; microblog; Chinese; classification model

Introduction

Clinical Features of High Suicide Probability

Identifying individuals with suicide probability at an early stage is vital for suicide intervention and prevention. Over the past few decades, people have dedicated themselves to identifying the characteristics of individuals with high suicide probability. Clinicians found high suicide risk in individuals with physical or psychological disease, for example, cancer, Acquired Immune Deficiency Syndrome, and depression [1-3]. There exists a strong connection between high suicide probability and certain personality traits [4,5]; individuals in Asia under specific age groups were reported to be with potential high risk, such as the elderly (especially those in rural areas) and teenagers [6-9]. As for the emotional level, research has shown that hostility, suicide ideation, negative self-evaluation, and depression are the key indicators of suicide. Although many risk factors have been reported to be correlated with suicide probability, it is still difficult to identify suicidal individuals, since suicidal behavior consists of a constellation of complex factors and everyone is unique [10-12]. Moreover, preventive intervention for high suicide probability individuals is often lagged behind, as efforts to track suicidal individuals in populations are hampered by difficulties in data collection and identification of suicide probability [13].

Research of Internet Suicide Probability Analysis

As the Internet has become a fast growing platform for social interaction in recent years, there are a large number of social network platforms containing suicide related information, which provide a rich source for monitoring suicide probability [14,15]. Researchers have been trying to figure out suicide features and trends from the Internet [13,16-20], and some have managed to locate certain high suicide risk groups by social network analysis [21]. Nevertheless, to the best of our knowledge, little research has been conducted for identifying high suicide probability of individuals using a constellation of Internet features.

Research Objective

In this study, we examine the feasibility and effectiveness of identifying high suicide probability microblog users automatically based on Internet accessible data. As the dominant microblog service provider in China, Sina Weibo now has 167 million active users, and more than 100 million posts are published daily [22], which provide rich behavior and linguistic information of individuals for any further analysis. As almost all of Weibo users are 35 or younger, this brings us an excellent opportunity to investigate the suicide risk of Weibo youth. We adopt the Suicide Probability Scale in Mandarin to label the suicide probability level of the Weibo users that participated in our Internet survey, and to determine our target group, for example, participants with high risk. We employed two machine learning algorithms, Simple Logistic Regression (SLR) and Random Forest (RF), to train classifiers to predict individual suicide probability via their profile and linguistic features extracted from Sina Weibo, and evaluated the performance of these classifiers on the labeled target group. We expect that the classifier with the best performance can properly identify high-risk individuals through their Weibo data with acceptable accuracy.

Methods

Participants and Procedures

Participants were invited to take part in this Internet survey via three approaches on Sina Weibo: (1) recruiting information was published on our laboratory's official Sina Weibo account with over 5000 followers. Some of the followers took part in the survey voluntarily; (2) a verified celebrity of Sina Weibo, who is a prestigious psychologist in mainland China and has more than 970,000 followers, retweeted our recruiting information and attracted more participants; and (3) another nonofficial Weibo account had been created to send invitation messages randomly on user's home page. All participants interested in this survey were asked to log on to the Internet survey system by their Sina Weibo account. After they finished reading and signing an informed consent form specifying the objective of the survey and their rights, they were invited to fulfill a survey on demographic information and mental health status, including the Suicide Probability Scale (SPS) in Mandarin. They received a compensation of 30 Renminbi if they completed the whole survey. Contact information of a national suicide prevention hotline was shown on the survey Web page, and the participants were encouraged to seek help if they felt stressful or suicidal. Ethical considerations of the study have been reviewed and granted by the Review Board of the Institute of Psychology, Chinese Academy of Sciences.

Participant Exclusion Criteria

A participant screening was conducted to assure the quality of this whole process. First, to comply with ethic code, only participants above 18 years of age would be involved. Next, to decrease the possibility that one fulfilled the survey more than once with different microblog accounts, participants' Internet Protocol (IP) addresses were examined. Survey submissions from the same IP would be eliminated, thus only the first submission would be used. Last, but not least, it was considered that one should have an adequate amount of microblog posts for feature extraction to avoid the "floor effect", and we only kept participants with more than 100 posts in total.

From May 22th to July 13th, 2014, 1196 Weibo users took part in the survey, 1040 completed the whole survey and 909 of them passed the screening. The final sample pool consisted of 909 Sina Weibo users (561 female, 348 male, mean of age 24.3, SD 5.0).

Measures

Labeling High-Risk Participants

The SPS was developed by Cull and Gill to assess suicide risk of adults and adolescents above the age of 14. Previous studies have verified that SPS could be utilized as an effective screening

XSL•FO
RenderX

tool in the community for individual suicide prevention and intervention [23,24]. Liang et al have translated the standardized scale into Mandarin and verified its reliability and validity [25]. SPS consists of 36 self-report questions using a 4-point Likert scale ranging from "none" to "all of the time". Participants would get a total score of overall suicide probability, as well as scores in four subscales: (1) hostility, (2) suicide ideation, (3) negative self-evaluation, and (4) desperation.

SPS is substantially related to an externally developed index of suicide risk; individuals identified with high suicide probability require further expert scrutiny, or conditional evaluation with family members and friends. The Ontario Hospital Association and Canadian Patient Safety Institute suggested a total raw score of 78 as the cutoff point for high suicide risk [26]. Since there has been no standard norm of SPS score for microblog users in China yet, participants who scored one SD above either the mean of total SPS score or the mean in each subscale score in our Weibo user sample were labeled as high-risk individuals respectively (details in Table 1).

Table 1. SPS score distribution and score-based categorization.

Name of scale	Average score x (SD)	Cutoff for high score class > cutoff point n (%)	Cutoff for low score class < cutoff point n (%)
SPS	69.4 (11.8)	>81 144/909 (15.8)	<58 125/909 (13.8)
Hostility subscale	13.0 (2.5)	>15 137/909 (15.1)	<11 142/909 (15.6)
Suicide ideation subscale	11.5 (3.2)	>14 156/909 (17.2)	<9 94/909 (10.3)
Negative self-evaluation subscale	20.5 (4.4)	>24 173/909 (19.0)	<17 166/909 (18.3)
Desperation subscale	24.6 (4.7)	>29 135/909 (14.9)	<20 110/909 (12.1)

Extracting Features From Microblogs

Calling application programming interfaces, provided by Sina Weibo Data Center, allowed all of the publically available digital records of users to be downloaded, from which profile and linguistic features were extracted to train models.

Profile features consist of three types of categories: (1) participant profile or general behavior; (2) user settings; and (3) participant's microblog behavior.

Category (1) includes: gender; length of username; total number of favorites/followers/follows/friends (mutual follow); length of self-description; length of domain name; count of numbers in domain name; number of openly published microblogs; number of originally published microblogs; number of originally published microblogs with photos; number of originally published posts with URL; numbers of originally published posts with "@"; number of microblogs published between 22:00 and 6:00; number of times that participant used first person plural/singular words; number of total/positive/negative emoticons; and number of days that participant stayed active. To determine positive and negative emoticons, five psychology professionals were recruited to evaluate all 1983 Sina Weibo emoticons. Based on their agreement, 48 positive emoticons and 118 negative emoticons were ultimately identified.

Category (2) includes: whether the user enables private message sending; whether the user allows all users to leave comments; whether the user enables geotagging of their account; and whether the user includes "I" in self-description.

Category (3) includes: the average/maximum/minimum/median number of words in participant's single microblog; the average number of comments on participant's single microblog; the average number of times that participant's single microblog was retweeted; the average number of "likes" for participant's single microblog; microblog originality (original posts/total posts in public domain); microblog transitivity (posts containing hyperlinks/total posts in public domain); microblog interaction (posts @ other users/total posts in public domain); group reference (the average number of first person plural words per post); self-reference (the average number of first person singular words per post); nocturnal activeness (posts published during 22:00 to 6:00/total posts in public domain); adoption of positive emoticons (the average number of positive emoticons per post); adoption of negative emoticons (the average number of negative emoticons per post); and social activeness (number of friends/number of followers). Ratio data were adopted in many of the Category (3) features to eliminate the impact of time discontinuity, since participants varied in the Weibo active period.

We adopted those features according to three criteria: (1) very few features are raised in previous research. For example, there has been a lot of work focusing on the connection between suicide intention, depressed thinking, and insomnia [27,28], based on this, we adopted the feature of "nocturnal activeness"; (2) some features are defined intuitively, as we think there might exist some kind of relation between the feature and suicide risk (eg, the average number of negative emoticons used per post); and (3) for all the rest, they seem to be common, but important,

XSL•FO
RenderX

and we should pay attention to them. Although they have never been mentioned, it is possible that they turn out to be useful for identifying suicide risk.

Using Simplified Chinese Micro-blog Word Count Dictionary (SCMBWC), a Chinese version of Language Inquiry and Word Count [29], which is an effective lexicon for Weibo text analysis [30], linguistic features were extracted. There are 88 features in SCMBWC, covering basic categories in Chinese linguistics such as language process, psychological process, person concern, and oral language. TextMind, a Chinese text analysis system [31], was used in this study to carry out the task of linguistic feature extraction [30].

Modeling

Methodology for Modeling

We built our models on a training set and then evaluated them on a hold-out test set. To do so, we first divided all the participants into three classes. As mentioned above, participants scoring one SD above the mean (mean+1SD) were labeled as high-risk individuals. Accordingly, participants scoring below mean-1SD were labeled as low-risk ones, and those scoring in between were labeled as medium-risk ones. Intuitively, there may exist significant difference in behavioral and linguistic features between high-risk individuals and low-risk ones, thus, models built upon these two groups might capture the appropriate patterns to differentiate high-risk individuals from low-risk ones. To ensure model applicability for the general Weibo user crowd, the proportion of each class in a test set follows the same distribution of the whole participant sample, in which case the performance of models can be genuinely reflected.

Therefore, the training sets are from two extreme groups only, but test sets consist of participants in all three groups, since we want to test the performance of the model in a real world scenario. Here, we run training and testing by 5-fold cross validation. Each training set consisted of 80% of the high-risk and low-risk individuals (suicide probability, 216/269; hostility, 224/279; suicide ideation, 201/250; negative self-evaluation, 272/339; and desperation, 196/245), and each test set consisted of 20% of high-risk, medium-risk, and low-risk individuals (181/909). Both training set and test set were randomly generated 5 times from the whole participant pool to balance the variance of stratified random sampling.

Modeling Algorithms and Performance Metrics

There were two machine learning algorithms that were employed for training classification models, SLR and RF. SLR is a type of probabilistic classification model which is a special case of linear model with binary dependent variable. RF is an ensemble method, training multiple decision trees and the final result is the mode of all decision trees' outputs. The two algorithms have both been used in previous research to triage health problems [32-36]. To evaluate the models, three classic performance metrics were used: (1) Precision (number of true positives/total number of instances predicted to be positive), (2) Recall (number of true positives/total number of positive instances), and (3) F1 measure, which considers the 1:1 tradeoff between precision and recall to give a balanced view [37].

In addition, we also defined "Screening Efficiency" to measure the capacity of workload saved comparing with traditional clinical suicide scrutiny. Screening Efficiency was calculated as, (total number of instances - total number of instances predicted to be positive)/total number of instances. For example, if there were in total 100 individuals, and 40 of them were prescreened by our model as highly risky, then only 40 of them would have to move forward for expert evaluation, thus the workload we might save should be (100-40)/100*100%=60%. Training and testing of models were all conducted via WEKA, a widely adopted machine learning workbench for data mining [38].

Results

User Statistics

The majority of users (873/909, 96.0%) were adults below the age of 35, which is consistent with the current age distribution in Sina Weibo. Table 1 summarizes the score distribution and categorization in the whole participant sample pool for total suicide probability and four subscale dimensions. The sample size of each training set (containing 80% of high-score and low-score users) was summarized as follows: 216/269 for SPS total score, 224/279 for hostility score, 201/250 for suicide ideation score, 272/339 for negative self-evaluation score, and 196/245 for desperation score. The sample size of all testing sets was 181/909 (20% of total users under stratified sampling).

Evaluation

Tables 2-6 show performance of the models on overall suicide probability, as well as four subscale dimensions. SLR and RF were generally matched in performance of classifying potentially risky individuals. For overall suicide probability, the optimal model output was able to achieve a Recall value of 0.82, and Screening Efficiency varied between 0.32-0.46. For hostility dimension, the optimal model output was able to achieve a Recall value of 0.70, and Screening Efficiency varied between 0.42-0.65. For suicide ideation dimension, the optimal model output was able to achieve a Recall value of 0.84, and Screening Efficiency varied between 0.15-0.33. For negative self-evaluation dimension, the optimal model output was able to achieve a Recall value of 0.74, and Screening Efficiency varied between 0.38-0.55. For desperation dimension, apart from two outputs from SLR that tended to identify all individuals as high score, the optimal model output was able to achieve a Recall value of 0.89, and Screening Efficiency varied between 0.21-0.48. Precision values in model outputs varied between 0.1-0.25, and F1 measures varied between 0.17-0.37.

XSL•FO
RenderX

Table 2. Model performance for classifying overall suicide probability.

Classifier	Trial number	Performance metrics			
		Precision	Recall	F1 measure	Screening efficiency
SLR	1	0.13	0.50	0.20	0.38
	2	0.14	0.54	0.23	0.42
	3	0.23	0.79	0.35	0.46
	4	0.13	0.50	0.21	0.41
	5	0.19	0.79	0.31	0.36
RF	1	0.13	0.57	0.21	0.32
	2	0.18	0.75	0.29	0.34
	3	0.20	0.82	0.32	0.36
	4	0.16	0.64	0.26	0.38
	5	0.15	0.64	0.24	0.33

Table 3. Model performance for classifying hostility.

Classifier	Trial number	Performance metrics			
		Precision	Recall	F1 measure	Screening efficiency
SLR	1	0.12	0.30	0.17	0.62
	2	0.16	0.37	0.22	0.65
	3	0.18	0.52	0.26	0.56
	4	0.16	0.44	0.24	0.60
	5	0.21	0.70	0.33	0.50
RF	1	0.14	0.56	0.22	0.40
	2	0.17	0.67	0.27	0.42
	3	0.14	0.48	0.21	0.47
	4	0.12	0.44	0.18	0.42
	5	0.14	0.52	0.22	0.44

Table 4. Model performance for classifying suicide ideation.

Classifier	Trial number	Performance metrics			
		Precision	Recall	F1 measure	Screening efficiency
SLR	1	0.19	0.81	0.31	0.29
	2	0.22	0.84	0.34	0.33
	3	0.19	0.74	0.30	0.33
	4	0.16	0.65	0.26	0.31
	5	0.20	0.81	0.32	0.30
RF	1	0.17	0.84	0.28	0.15
	2	0.17	0.81	0.29	0.20
	3	0.18	0.84	0.29	0.18
	4	0.17	0.77	0.28	0.21
	5	0.17	0.77	0.27	0.20

XSL•FO
RenderX

Table 5. Model performance for classifying negative self-evaluation.

Classifier	Trial number	Performance metrics			
		Precision	Recall	F1 measure	Screening efficiency
SLR	1	0.25	0.68	0.37	0.49
	2	0.24	0.59	0.34	0.53
	3	0.20	0.47	0.29	0.55
	4	0.21	0.62	0.32	0.45
	5	0.24	0.74	0.36	0.41
RF	1	0.22	0.71	0.33	0.39
	2	0.23	0.65	0.34	0.47
	3	0.22	0.65	0.33	0.46
	4	0.22	0.74	0.34	0.38
	5	0.20	0.62	0.30	0.41

Table 6. Model performance for classifying desperation.

Classifier	Trial number	Performance metrics			
		Precision	Recall	F1 measure	Screening efficiency
SLR	1	0.15	1.00	0.26	0
	2	0.17	0.89	0.29	0.22
	3	0.15	1.00	0.26	0
	4	0.14	0.48	0.21	0.48
	5	0.15	0.63	0.24	0.36
RF	1	0.14	0.67	0.24	0.31
	2	0.13	0.67	0.22	0.26
	3	0.13	0.56	0.21	0.37
	4	0.10	0.44	0.17	0.37
	5	0.15	0.78	0.25	0.21

Discussion

Principal Results and Comparison With Prior Work

The key finding of our study is that a high level of suicide probability along the dimension of hostility, suicide ideation, negative self-evaluation, and desperation can be identified with acceptable performance via the profile and text data of microblog users. It is shown that classification performance was generally matched between SLR and RF. Precision varies from 10% to 25%, Recall varies from 30% to 89%, F1 measures vary from 17% to 37%, and the Screening Efficiency varies from 21% to 65%. The performance of the classifiers seems to depend on the randomization of data between the training and testing sets. For example, the Recall on hostility using SLR varies by 40% (0.30-0.70), but only by 7% for suicide ideation using RF (0.77-0.84). It may suggest that the degree of generalizability is different for the four risk factors measured in subscales; for example, future studies may be designed to verify whether suicide ideation has the greatest potential in identifying individual suicide risk among all the emotional factors.

For any risky individual, suicide prevention and intervention is a continuous process, involving a constantly alternating process of suicide risk evaluation and intervention therapy [39]. The traditional process is both time and effort consuming, and because many suicidal individuals in China don't actively seek help [39], they are often beyond the reach of professional service. Researchers in the suicide prevention and intervention fields have realized the great potential of Web-based intervention; Internet programs have been developed to help people diagnosed as suicidal [40-42]. Our study aims at providing empirical evidence that a suicide risk evaluation process can be conducted through examining online social media content. A computerized algorithm evaluation can work side-by-side with traditional questionnaire methodology to provide reference information for identifying potentially risky individuals and guide them to further intervention.

As the evaluation result shows, among the three classic performance metrics, Recall is generally higher than the other two. This suggests that the models attempt to retrieve as many risky suicidal individuals as possible, even at the cost of partly increasing false alarm. Considering the severity of the suicide

XSL•FO
RenderX

act, we do not want to miss any risky individual. Therefore, Recall is our primary concern in this study. However, low Precision and F1 measure indicate that the current model alone can only serve as a preliminary screening tool for suicide probability. Some of the latest research findings also suggest that even though prediction of psychological problems by machine learning algorithms have advanced in accuracy, they still cannot take the place of expert scrutiny [43-47]. To apply our current findings, we can work together with suicide prevention organizations, the computerized program prescreens Weibo users' suicide risk and then automatically refer high-risk individuals to such organizations. They will further manually examine and provide intervention services according to their professional assessment.

It is thus of our particular interest to explore to what extent preliminary screening of high-risk individuals via machine learning algorithms can reduce the workload in traditional scale assessment for suicide risk. It is shown from our newly defined metric "Screening Efficiency" that, assuming the proposed models serve at their best performance, currently we are just able to save less than half of the traditional workload in general. Although not directly complementary to Recall, a sign of tradeoff has been revealed in many of the experiment trials between the amount of saved workload for further scrutiny, and the proportion of correctly retrieved high-risk individuals. Combining the model evaluation results, we believe there is still much space for advancement in improving the predictive power of models in successive research. Nevertheless, it has been a good start to concentrate on the progressive attempt of feature extraction, modeling design, and classifier selection.

Limitations

In order to facilitate the usability of our Internet survey system, we allowed participants to complete the survey discontinuously. In other words, if a participant was interrupted and forced to pause the survey partly completed, the progress could be saved for the next access. We did find a few participants with long fulfilling time, and were unable to tell whether they were interrupted, or other reasons that might potentially bias the value of self-report assessment. This concern calls for the optimization of Internet assessment methodology. Some researchers have already been working on developing short, good quality tools to test suicidal behavior on the Internet [48], but more efforts need to be spent to reduce response burden and improve accuracy for Internet self-report evaluation.

It is natural to wonder whether there are some features with the strongest predictive power among all the proposed features. According to the model outputs of our study, the powerful indicators are not consistent among different models; the predictive features in models with the same algorithm would even appear different among different trials. In addition, the predictive features are often uninterpretable. Although one of the advantages of machine learning is to discover hidden relations that do not fit in with the current knowledge system, we admit that currently we have better knowledge concerning the overall predictive power of modeling than the specific predictive power of a single feature. It is of our interest to consolidate feature systems and to strengthen output interpretation.

In this pilot study, we categorized users into three classes, and particularly labeled those who scored mean+1SD as high-risk individuals to indicate that they are more likely in need of careful clinical evaluation of suicide risk. Because there has been no norm group with regard to suicide probability scores among China's Sina Weibo users, we are aware of the possibility of potential bias with regard to this user sample and the based cutoff points for high suicide probability. For future studies that intend to advance in the suicide Internet research in China, they may investigate the localization of this measuring tool into a specific Internet group.

Conclusions

Social media is widely used at the present time. Our study indicates that high suicide probability can be evaluated via the publicized profile and text information of microblog users. Although currently our model is unable to reach sufficient accuracy to provide diagnosis, this innovative approach does shed light on the value of monitoring large-scale populations, and enables detecting potentially suicidal individuals for suicide prevention professionals' further follow-up. Future studies need to focus on increasing the accuracy of classification, and testing the performance on a larger scope of social media users.

Acknowledgments

The authors gratefully acknowledge the generous support from the National High-Tech R&D Program of China (2013AA01A606), the National Basic Research Program of China (2014CB744600), the Key Research Program of Chinese Academy of Sciences (CAS) (KJZD-EWL04), and the CAS Strategic Priority Research Program (XDA06030800). The study was also partly supported by the Research Grants Council Strategic Public Policy Grant (HKU 7003-SPPR-12).

Conflicts of Interest

None declared.

References

1. Innos K, Rahu K, Rahu M, Baburin A. Suicides among cancer patients in Estonia: A population-based study. European Journal of Cancer 2003 Oct;39(15):2223-2228. [doi: 10.1016/S0959-8049(03)00598-7]

XSL•FO
RenderX

2. Kim YK, Lee SW, Kim SH, Shim SH, Han SW, Choi SH, et al. Differences in cytokines between non-suicidal patients and suicidal patients in major depression. Prog Neuropsychopharmacol Biol Psychiatry 2008 Feb 15;32(2):356-361. [doi: 10.1016/j.pnpbp.2007.08.041] [Medline: 17919797]
3. Leserman J. HIV disease progression: Depression, stress, and possible mechanisms. Biological Psychiatry 2003 Aug;54(3):295-306. [doi: 10.1016/S0006-3223(03)00323-8]
4. Wang CW, Chan CL, Yip, PS. Suicide rates in China from 2002 to 2011: An update. Social Psychiatry Psychiatric Epidemiology 2014;49(6):929-941. [doi: 10.1007/s00127-013-0789-5]
5. Conner KR, Meldrum S, Wieczorek WF, Duberstein PR, Welte JW. The association of irritability and impulsivity with suicidal ideation among 15- to 20-year-old males. Suicide Life Threat Behav 2004;34(4):363-373. [doi: 10.1521/suli.34.4.363.53745] [Medline: 15585458]
6. World Health Organization. Preventing suicide: A global imperative. Geneva: WHO Publications; 2014.
7. Awata S, Seki T, Koizumi Y, Sato S, Hozawa A, Omori K, et al. Factors associated with suicidal ideation in an elderly urban Japanese population: A community‐based, cross‐sectional study. Psychiatry and clinical neurosciences 2005;59(3):327-336. [doi: 10.1111/j.1440-1819.2005.01378.x]
8. Bjørngaard JH, Bjerkeset O, Vatten L, Janszky I, Gunnell D, Romundstad P. Maternal age at child birth, birth order, and suicide at a young age: A sibling comparison. American journal of epidemiology 2014;kwt014. [doi: 10.1093/aje/kwt014]
9. Harwood DMJ, Hawton K, Hope T, Harriss L, Jacoby R. Life problems and physical illness as risk factors for suicide in older people: A descriptive and case-control study. Psychological Medicine 2006;36(09):1265-1274. [doi: 10.1017/S0033291706007872]
10. Mościcki EK. Identification of suicide risk factors using epidemiologic studies. Psychiatric Clinics of North America 1997 Sep;20(3):499-517. [doi: 10.1016/S0193-953X(05)70327-0]
11. Phillips MR, Yang G, Zhang Y, Wang L, Ji H, Zhou M. Risk factors for suicide in China: A national case-control psychological autopsy study. The Lancet 2002 Nov;360(9347):1728-1736. [doi: 10.1016/S0140-6736(02)11681-3]
12. Borges G, Angst J, Nock MK, Ruscio AM, Kessler RC. Risk factors for the incidence and persistence of suicide-related outcomes: A 10-year follow-up study using the National Comorbidity Surveys. Journal of Affective Disorders 2008;105(1):25-33. [doi: 10.1016/j.jad.2007.01.036]
13. McCarthy MJ. Internet monitoring of suicide risk in the population. J Affect Disord 2010 May;122(3):277-279 [FREE Full text] [doi: 10.1016/j.jad.2009.08.015] [Medline: 19748681]
14. Westerlund M, Hadlaczky G, Wasserman D. The representation of suicide on the Internet: Implications for clinicians. Journal of medical Internet research 2012;14(5). [doi: 10.2196/jmir.1979]
15. Kemp CG, Collings SC. Hyperlinked suicide. Crisis: The Journal of Crisis Intervention and Suicide Prevention 2011;32(3):143-151. [doi: 10.1027/0227-5910/a000068]
16. Chen P, Chai J, Zhang L, Wang D. International conference on advanced information engineering and education science (ICAIEES 2013). 2013. Development and application of a Chinese webpage suicide information mining system (SIMS) URL: http://www.atlantis-press.com/php/download_paper.php?id=10818 [accessed 2014-10-30] [WebCite Cache ID 6ThjoJo3S]
17. Jashinsky J, Burton SH, Hanson CL, West J, Giraud-Carrier C, Barnes MD, et al. Tracking suicide risk factors through Twitter in the US. Crisis 2014;35(1):51-59. [doi: 10.1027/0227-5910/a000234] [Medline: 24121153]
18. Li TM, Chau M, Yip PS, Wong PW. Temporal and computerized psycholinguistic analysis of the blog of a Chinese adolescent suicide. Crisis: The Journal of Crisis Intervention and Suicide Prevention 2014;35(3):168-175. [doi: 10.1027/0227-5910/a000248]
19. Fu KW, Cheng Q, Wong PW, Yip PS. Responses to a self-presented suicide attempt in social media: A social network analysis. Crisis: The Journal of Crisis Intervention and Suicide Prevention 2013;34(6):406-412. [doi: 10.1027/0227-5910/a000221]
20. Cheng Q, Chang S, Yip PS. Opportunities and challenges of online data collection for suicide prevention. The Lancet 2012 May;379(9830):e53-e54. [doi: 10.1016/S0140-6736(12)60856-3]
21. Silenzio V, Duberstein PR, Tang W, Lu N, Tu X, Homan CM. Connecting the invisible dots: Reaching lesbian, gay, and bisexual adolescents and young adults at risk for suicide through online social networks. Social Science & Medicine 2009;69(3):469-474. [doi: 10.1016/j.socscimed.2009.05.029]
22. 2014 Sina Weibo User Report. Sina Weibo Data Center, 2014. URL: http://www.199it.com/archives/324955.html [accessed 2015-04-20] [WebCite Cache ID 6XvKIVl00]
23. Naud H, Daigle MS. Predictive validity of the suicide probability scale in a male inmate population. J Psychopathol Behav Assess 2009 Sep 2;32(3):333-342. [doi: 10.1007/s10862-009-9159-8]
24. Gençöz T, Or P. Associated factors of suicide among university students: Importance of family environment. Contemp Fam Ther 2006 May 10;28(2):261-268. [doi: 10.1007/s10591-006-9003-1]
25. Liang Y, Yang L. China Journal of Psychology. 2010. Study on reliability and validity of the Suicide Probability Scale URL: http://www.cqvip.com/qk/98348a/201002/33125858.html [accessed 2015-03-25] [WebCite Cache ID 6XHni3FBA]
26. Perlman CM, Neufeld E, Martin L, Goy M, Hirdes JP. Ontario Hospital Association and Canadian Patient Safety Institute. Toronto: ON; 2011. Suicide risk assessment inventory: A resource guide for Canadian health care organizations URL:

XSL•FO
RenderX

http://www.oha.com/KnowledgeCentre/Documents/Final%20-%20Suicide%20Risk%20Assessment%20Guidebook.pdf [accessed 2015-04-28] [WebCite Cache ID 6Y7UsbvJM]
27. Nadorff MR, Fiske A, Sperry JA, Petts R, Gregg JJ. Insomnia symptoms, nightmares, and suicidal ideation in older adults. The Journals of Gerontology Series B: Psychological Sciences and Social Sciences 2013;68(2):145-152. [doi: 10.1093/geronb/gbs061]
28. Woosley JA, Lichstein KL, Taylor DJ, Riedel BW, Bush AJ. Hopelessness mediates the relation between insomnia and suicidal ideation. J Clin Sleep Med 2014 Nov 15;10(11):1223-1230. [doi: 10.5664/jcsm.4208] [Medline: 25325598]
29. Pennebaker JW, Francis ME, Booth RJ. Mahway: Lawrence Erlbaum Associates. Linguistic inquiry and word count: LIWC 2001 URL: http://homepage.psy.utexas.edu/HomePage/Faculty/Pennebaker/Reprints/LIWC2007_OperatorManual.pdf [accessed 2015-03-25] [WebCite Cache ID 6XHnLiOE0]
30. Gao R, Hao B, Li H, Gao Y, Zhu T. Developing simplified Chinese psychological linguistic analysis dictionary for microblog. Brain and Health Informatics 2013;8211:359-368. [doi: 10.1007/978-3-319-02753-1_36]
31. Textmind system. 2013. URL: http://ccpl.psych.ac.cn/textmind/ [accessed 2015-04-26] [WebCite Cache ID 6Y6PqNLE8]
32. Díaz-Uriarte R, Alvarez de Andrés Sara. Gene selection and classification of microarray data using random forest. BMC Bioinformatics 2006;7:3 [FREE Full text] [doi: 10.1186/1471-2105-7-3] [Medline: 16398926]
33. Kacar K, Rocca MA, Copetti M, Sala S, Mesaroš Š, Opinćal TS. Overcoming the clinical–MR imaging paradox of multiple sclerosis: MR imaging data assessed with a random forest approach. American Journal of Neuroradiology 2011;32(11):2098-2102. [doi: 10.3174/ajnr.A2864]
34. Gray KR, Aljabar P, Heckemann RA, Hammers A, Rueckert D, Alzheimer's Disease Neuroimaging Initiative. Random forest-based similarity measures for multi-modal classification of Alzheimer's disease. Neuroimage 2013 Jan 15;65:167-175 [FREE Full text] [doi: 10.1016/j.neuroimage.2012.09.065] [Medline: 23041336]
35. Theodossiou I. The effects of low-pay and unemployment on psychological well-being: A logistic regression approach. Journal of Health Economics 1998 Jan;17(1):85-104. [doi: 10.1016/S0167-6296(97)00018-0]
36. Smith V, Decuman S, Sulli A, Bonroy C, Piettte Y, Deschepper E, et al. Do worsening scleroderma capillaroscopic patterns predict future severe organ involvement? A pilot study. Ann Rheum Dis 2012 Oct;71(10):1636-1639. [doi: 10.1136/annrheumdis-2011-200780] [Medline: 22402146]
37. Goutte C, Gaussier E. A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. Advances in Information Retrieval 2005;3408:345-359. [doi: 10.1007/978-3-540-31865-1_25]
38. Frank E, Hall M, Holmes G, Kirkby R, Pfahringer B, Witten IH, et al. Weka-a machine learning workbench for data mining. Data Mining and Knowledge Discovery Handbook. Springer US 2010. [doi: 10.1007/978-0-387-09823-4_66]
39. Li G. Suicide and self-harm. 1st ed. Beijing: People's Medical Publishing House; 2007:196-197.
40. Furber G, Jones GM, Healey D, Bidargaddi N. A comparison between phone-based psychotherapy with and without text messaging support in between sessions for crisis patients. J Med Internet Res 2014;16(10):e219 [FREE Full text] [doi: 10.2196/jmir.3096] [Medline: 25295667]
41. Stjernswärd S, Hansson L. A web-based supportive intervention for families living with depression: Content analysis and formative evaluation. JMIR research protocols 2014;3(1). [doi: 10.2196/resprot.3051]
42. Whiteside U, Lungu A, Richards J, Simon GE, Clingan S, Siler J, et al. Designing messaging to engage patients in an online suicide prevention intervention: Survey results from patients with current suicidal ideation. Journal of medical Internet research 2014;16(2). [doi: 10.2196/jmir.3173]
43. Wald R, Khoshgoftaar TM, Napolitano A, Sumner C. Using Twitter content to predict psychopathy. In Machine Learning and Applications (ICMLA) 11th International Conference 2012 Dec;2:394-401. [doi: 10.1109/ICMLA.2012.228]
44. Poulin C, Shiner B, Thompson P, Vepstas L, Young-Xu Y, Goertzel B. PLoS ONE. 2014 Mar 20. Predicting the risk of suicide by analyzing the text of clinical notes URL: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0085733 [accessed 2015-04-27] [WebCite Cache ID 6Y6UQv2TP]
45. De Choudhury M, Gamon M, Counts S, Horvitz E. In ICWSM. 2013. Predicting depression via social media URL: http://research.microsoft.com/pubs/192721/icwsm_13.pdf [accessed 2015-04-07] [WebCite Cache ID 6XbueguaQ]
46. De Choudhury M, Counts S, Horvitz E. Predicting postpartum changes in emotion and behavior via social media. 2013 Apr Presented at: the SIGCHI Conference on Human Factors in Computing Systems. ACM; 2013; Paris p. 3267-3276. [doi: 10.1145/2470654.2466447]
47. De Choudhury M, Counts S, Horvitz EJ, Hoff A. Characterizing and predicting postpartum depression from shared facebook data. 2014 Presented at: the 17th ACM conference on Computer supported cooperative work & social computing ACM; February 2014; Baltimore p. 626-638 URL: http://www.msr-waypoint.net/en-us/um/people/horvitz/FB-cscw2014.pdf [doi: 10.1145/2470654.2466447]
48. De Beurs Derek Paul, de Vries Anton Lm, de Groot Marieke H, de KJ, Kerkhof AJ. Applying computer adaptive testing to optimize online assessment of suicidal behavior: A simulation study. J Med Internet Res 2014;16(9):e207 [FREE Full text] [doi: 10.2196/jmir.3511] [Medline: 25213259]

XSL•FO
RenderX

Abbreviations

CAS: Chinese Academy of Sciences
IP: Internet Protocol
RF: Random Forest
SCMBWC: Simplified Chinese Micro-blog Word Count dictionary
SLR: Simple Logistic Regression
SPS: Suicide Probability Scale

Edited by G Eysenbach; submitted 12.01.15; peer-reviewed by B O'Dea, M Larsen; comments to author 18.03.15; revised version received 30.03.15; accepted 03.04.15; published 12.05.15.

Please cite as:
Guan L, Hao B, Cheng Q, Yip PSF, Zhu T
Identifying Chinese Microblog Users With High Suicide Probability Using Internet-Based Profile and Linguistic Features: Classification Model
JMIR Mental Health 2015;2(2):e17
URL: http://mental.jmir.org/2015/2/e17/
doi:10.2196/mental.4227
PMID:26543921

©Li Guan, Bibo Hao, Qijin Cheng, Paul SF Yip, Tingshao Zhu. Originally published in JMIR Mental Health (http://mental.jmir.org), 12.05.2015. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Word Recall: Cognitive Performance Within Internet Surveys

Shannon K Runge[1], MA; Benjamin M Craig[2], PhD; Heather S Jim[2], PhD

[1]University of South Florida and Moffitt Cancer Center, Tampa, FL, United States

[2]Moffitt Cancer Center and University of South Florida, Tampa, FL, United States

Corresponding Author:
Shannon K Runge, MA
University of South Florida and Moffitt Cancer Center
MRC-CANCONT
12902 Magnolia Dr.
Tampa, FL,
United States
Phone: 1 813 745 1245
Fax: 1 813 745 6525
Email: shannon.runge@moffitt.org

Abstract

Background: The use of online surveys for data collection has increased exponentially, yet it is often unclear whether interview-based cognitive assessments (such as face-to-face or telephonic word recall tasks) can be adapted for use in application-based research settings.

Objective: The objective of the current study was to compare and characterize the results of online word recall tasks to those of the Health and Retirement Study (HRS) and determine the feasibility and reliability of incorporating word recall tasks into application-based cognitive assessments.

Methods: The results of the online immediate and delayed word recall assessment, included within the Women's Health and Valuation (WHV) study, were compared to the results of the immediate and delayed recall tasks of Waves 5-11 (2000-2012) of the HRS.

Results: Performance on the WHV immediate and delayed tasks demonstrated strong concordance with performance on the HRS tasks (ρc=.79, 95% CI 0.67-0.91), despite significant differences between study populations (P<.001) and study design. Sociodemographic characteristics and self-reported memory demonstrated similar relationships with performance on both the HRS and WHV tasks.

Conclusions: The key finding of this study is that the HRS word recall tasks performed similarly when used as an online cognitive assessment in the WHV. Online administration of cognitive tests, which has the potential to significantly reduce participant and administrative burden, should be considered in future research studies and health assessments.

(JMIR Mental Health 2015;2(2):e20) doi:10.2196/mental.3969

KEYWORDS

cognition; online surveys; episodic memory; Health and Retirement Study; Women's Health Valuation Study

Introduction

The use of Internet-enabled devices, such as computers, smartphones, and tablets, to conduct cognitive research has increased dramatically over the past decade [1-3]. These devices allow researchers to use application-based cognitive assessments that have distinct advantages over more traditional assessment methods (ie, face-to-face interviews), including rapid data collection, reduced participant and administrative burden, and access to diverse or hard-to-reach populations [4,5]. When used in either a community or clinic setting, such online applications may detect cognitive and behavioral information that is missed with face-to-face assessments [6], including millisecond changes in cognitive processes [2]. Furthermore, in light of recent recommendations that cognitive screenings be included as a part of routine personalized health care [7], online cognitive assessments may play an important role in detecting subtle changes in cognitive function for both healthy and clinical populations at times when prevention and intervention strategies may have an optimal impact [8].

Application-based administration of cognitive tests has the potential to significantly advance research examining changes

XSL•FO
RenderX

in cognition due to aging or illness. Repeated, short online cognitive batteries can provide a fine-grained assessment of cognitive capabilities in everyday life. For example, studies could examine situations or times of day in which cognitive lapses are most likely to occur (ie, during stress) [9,10], which can be used to devise targeted behavioral interventions to improve cognition. Similarly, more frequent cognitive assessments may help to better understand patterns of cognitive change over time in research cohorts or clinical settings.

Frequent use of cognitive assessments may be particularly important in clinical and primary care settings, where early indicators of mild cognitive impairment can be misdiagnosed as typical age-related declines in as many as 91% of cases [11]. This rate of misdiagnosis may be attributable to the frequent use of the Mini-Mental Status Examination, which lacks sensitivity to detect subclinical levels of cognitive decline compared to other assessments [12,13]. Rates of misdiagnosis are further exacerbated by individual subjective memory complaints [14]. Measures that evaluate more specific cognitive domains like episodic memory may be more specific for the detection of early changes in cognitive performance.

Episodic memory is one of the first domains in which people experience subclinical changes in cognitive performance [15,16]. Broadly described as a person's ability to recall temporally related events or dates [17], episodic memory is particularly sensitive to the effects of aging [18-20]. This is likely a reflection of age-related neurobiological changes that occur in areas of the brain associated with episodic memory (eg, prefrontal cortex, medial temporal lobes, and hippocampus [19,21,22]), such as the decreased availability of the neurotransmitter dopamine [23], changes in functional connectivity between brain regions [24,25], and volumetric reductions of the hippocampus and prefrontal cortex [21].

Recent evidence indicates that subtle changes in episodic memory can be detected in individuals with normal or slightly impaired cognitive abilities [26]. Examining episodic memory in clinical or research settings may be particularly valuable since lower baseline scores and greater rates of changes in episodic memory are likely to precede the onset of clinical symptoms of cognitive decline [16,27], especially for individuals with a genetic risk for Alzheimer's disease [28]. Recall tests are frequently used to estimate episodic memory as a part of larger interview-based [26,29] and online [3] neuropsychological batteries. Despite the clear advantages and potential benefits of application-based cognitive assessments, researchers often fail to demonstrate equivalence between their application-based assessment and its interview-based counterpart [3]. Ideally, equivalence between assessments (ie, construct validity) would be evaluated using a gold standard measure [30]. In the absence of such a standard, it is preferable to use an internally consistent and valid measure that has demonstrated response stability across samples [31,32].

In response to this gap, the current study opted to replicate the episodic memory tasks (immediate and delayed recall) of the Health and Retirement Study (HRS) in an online survey. These tasks were selected for a number of reasons. First, performance on the cognitive measures of the HRS has shown to be stable from wave to wave, after controlling for cohort effects and test-retest bias [33]. Second, none of these measures has been adapted for use in application-based assessments and tested for equivalence. Third, the format and presentation of the episodic tasks of the HRS were most easily replicated in an online format and would not require the use of complex computer technology that may be difficult or unavailable for older populations (eg, microphones). Finally, due to the authors' interest in age, this study was further motivated by evidence that episodic memory is more susceptible to increasing age compared to semantic memory (ie, abilities related to vocabulary and general knowledge [34]), which has been shown to remain stable well into later decades of life [18,20]. Given the age range of the online sample in the current study (40-69 years), as well as the previous methodological considerations, the replication of the episodic memory tasks was prioritized higher than the other HRS measures.

This study examines the performance of an online word recall task that was originally developed as part of the HRS for cognitively healthy adults. Specifically, the results of an online immediate and delayed word recall task in a nationally representative sample of women aged 40 to 69 years were compared to the results of female respondents from waves 5-11 (2000-2012) of the HRS. Using these primary and secondary data, two questions were examined: (1) Do the online word recall tasks demonstrate sufficient equivalence to the HRS word recall tasks? (2) Does word recall performance vary as a function of respondent characteristics and task modality? Ultimately, the results of this study will aid in the evaluation of the potential of cognitive assessments in online surveys and health assessments.

Methods

Study Samples

The Health and Retirement Study

Since its launch in 1992, the goal of the Health and Retirement Study (HRS) has been to provide a detailed, national representation of US adults aged 50 years and older. Jointly managed through the National Institute on Aging (U01 AG009740), the Institute for Social Research, and the University of Michigan (IRB Protocols HUM00056464, HUM00061128, HUM00002562, HUM00079949, HUM00080925, and HUM00074501), the HRS is widely cited as an excellent source of data for use in examining cognitive trends and abilities of the aging US population [35]. Data is collected via telephone and face-to-face interviews in 2-year cycles, with new cohorts added every 6 years. The HRS uses a dual modality approach, where initial interviews are conducted face-to-face and the majority of successive interviews are conducted over the telephone (unless participants are older than 80 years of age). Hispanic and black adults are oversampled. Spouses of HRS participants are also included, regardless of age.

The Women's Health Valuation Study

Conducted at Moffitt Cancer Center in Tampa, Florida, the Women's Health Valuation (WHV) study is an Internet-based health valuation study that included health measures and a

XSL•FO
RenderX

discrete choice experiment (DCE) where respondents reported their preferences between possible health outcomes. The approach and methods, including its sampling design and survey instrument, were adapted from the PROMIS-29 valuation study (1R01CA160104) [36] and approved by the University of South Florida Institutional Review Board (USF IRB Protocol 8236).

The WHV online survey instrument had four components: screener, health, DCE, and follow-up. Each component had a series of questions distributed across a continuous series of pages, and responses were recorded by clicking or typing answers and then hitting the *Next* button. Each page included a *Back* button so the respondent could return to previous pages and change previous answers; however, to discourage participants from returning to previous pages of the survey, the *Back* button was disabled. To exit the survey, respondents could close their browser at any time. If the browser was closed prior to completing the survey, the data were not recorded. Responses to all questions were mandatory in order to proceed to the next page.

Participants were recruited from a pre-existing national panel of US adults. To promote concordance with the 2010 US Census, participants were sampled according to 6 demographic quotas: age in years (40-54 and 55-69) and race/ethnicity (Hispanic; black, non-Hispanic; white; and other, non-Hispanic). Further details about the methods of this study are available online [37]. Overall, 4474 women completed the survey between April 3, 2013 and April 21, 2013.

Cognitive Measures of the Health and Retirement Study

Episodic Memory

The cognitive battery of the HRS has been evaluated for internal consistency and validity [38]. Latent factor path modeling has identified three cognitive domains: episodic memory (immediate and delayed recall), mental status (serial 7s, backward counting from 20, naming), and vocabulary (ie, semantic memory) [35]. Measures of episodic memory include an immediate and delayed recall task. Mental status is measured by a serial 7s subtraction test, counting backwards from 20, and naming (the last name of the current president and vice president; two objects [scissors and cactus] based on a brief verbal description; and the current month, day, year, and day of week). Semantic memory is assessed using a baseline measure of vocabulary (5 words) [39].

As a measure of episodic memory, the immediately and delayed recall tasks are drawn from four categorized lists of 10 English nouns that did not overlap in content. Respondents are randomly assigned to one of the four lists at the initial interview. Longitudinally, each respondent is randomly assigned to receive an alternative word list, such that each respondent is assigned to a different set of words for the three successive waves of data collection. With this counterbalanced approach, each respondent was assigned to each word list only once over 4 waves of data collection, and approximately 8 years will pass before a respondent is reassigned to the same set of words as their initial interview.

During the immediate recall task, an interviewer reads a list of 10 words at a rate of approximately 2 seconds per word to each respondent, who verbally recalled as many words as possible. Approximately 5 minutes after the immediate word recall test, during which respondents answered questions about their emotional state and completed two mental status tasks (eg, counting backwards, serial 7s), respondents were asked to recall the words from the immediate recall task. For each task, the number of correctly recalled words is scored, with higher scores indicating better performance.

Self-Reported Memory

In addition to episodic memory, HRS respondents are also asked to self-report their memory at the present time (excellent, very good, good, fair, or poor) and compare their current memory to their memory 2 years ago (better, same, or worse).

For the purpose of comparison, this study examines all word recall responses from waves 5-11 (2000-2012) of the HRS. Since the WHV was restricted to female respondents, we excluded male respondents from the HRS to decrease the risk of gender bias. Participants of the HRS who reported using a proxy respondent; refused to respond to word recall tasks; or had missing data on demographic, memory, or word recall variables (less than 2.0% of the sample) were also excluded. Aside from these exclusion criteria, 12,545 women completed between 1 and 7 word recall tasks with a median (interquartile range) of 3 tasks (2-5 tasks). These tasks were restructured to represent a cross-sectional dataset with a total of 43,417 word recall tasks.

Cognitive Measures of the Women's Health Valuation Study

Episodic Memory

The episodic memory of the WHV replicated the word recall task conducted as part of the HRS. All respondents were asked to recall 10 English nouns immediately after they were presented on-screen (immediate recall) and after a delay (delayed recall). Each respondent received one of four randomly assigned sets of words, which were taken verbatim from the HRS and presented in the same order. Prior to the immediate recall task, respondents were presented with a screen that informed them that they would be shown a set of 10 words and would be asked to recall as many words as they could. These instructions were largely based on those given to HRS respondents but modified for online presentation. Words appeared on the computer screen one at a time for approximately 3 seconds. Respondents were asked to recall the words directly after the presentation of all 10 words (immediate recall) and then approximately 20 minutes later at the end of the DCE component (delayed recall). For each recall, respondents typed as many words as they could remember, in any order, in empty text boxes within the survey. As with the HRS, the primary measure of episodic memory was the sum of correctly recalled words for each task, regardless of order.

Self-Reported Memory

The self-reported memory questions of the WHV were replicated from the self-reported memory questions of the HRS. As part

XSL•FO
RenderX

of the health component, the self-reported memory questions asked participants to rate their memory at the present time (excellent, very good, good, fair, or poor) and compare their current memory to their memory 2 years ago (better, same, or worse).

Compared to the word recall task in the HRS, the online task in WHV differed in the several ways. The word lists were displayed visually on a computer device/browser as opposed to being spoken by an interviewer (basic literacy skills were required, with less reliance on verbal communication), respondents recalled words by typing them versus speaking them (basic typing skills were required, with less reliance on verbal communication), and the words can sound the same with different spelling (eg, see vs sea and rock vs roc), which may make the WHV task more specific. In addition, the delay between the immediate and delayed recalls task was shorter (5 minutes vs 20 minutes) and the WHV version was purely cross-sectional, whereas HRS respondents may have completed the tasks up to seven times. Nevertheless, the study took all available steps possible to replicate the original HRS tasks.

Statistical Analyses

Demographic and descriptive statistics (Table 1) obtained on both groups were analyzed using independent sample *t* tests, Pearson chi-square, and one-way analyses of variance, where appropriate. In order to estimate the precision and accuracy of the two word recall tasks, Lin's concordance correlation coefficient (ρ_c) [30] was used to collectively compare the average frequency with which the WHV and HRS participants recalled each word. Unlike Pearson's correlation coefficient, which estimates only the linear covariation between variables, Lin's concordance quantifies the degree of agreement between two measures of the same variable by providing a measure of covariation and correspondence [30]. Finally, multivariate linear regression models adjusted for cluster errors (ie, multiple tasks per respondent) were used to estimate the associations between characteristics of each study sample and number of correctly recalled words for the immediate and delayed recall tasks. All analyses were conducted using Stata 13 software (StataCorp).

Results

Overview

The WHV online survey had 4474 respondents, each of whom completed 1 word recall task. The HRS survey had 12,545 respondents who completed between 1 and 7 recall tasks. As shown in Table 1, WHV respondents differed significantly from HRS respondents along each characteristic. Overall, WHV respondents were more likely to be white or Hispanic, younger, and better educated and report excellent or very good memory compared to HRS respondents, possibly due to sampling from an online panel.

Figure 1 is a scatterplot of the likelihood of immediate recall for each word by modality, which ranges from 0.49 to 0.85 for WHV respondents and 0.33 to 0.91 for HRS respondents. Out of the 40 words, 35 words had greater recall for the WHV versus HRS task with a mean difference of 11.82% (95% CI −0.31 to 0.08). At first glance, Lin's concordance correlation coefficient (ρ_c=.57, 95% CI 0.42-0.722) indicated mild correspondence. Once the likelihoods were normalized (ie, subtracting the sample mean and dividing by the standard deviation), Lin's concordance correlation coefficient increased to .789 (95% CI 0.67-0.91), indicating strong correspondence. Similarly, the delayed recall task showed Lin's concordance correlation coefficient with and without normalization that suggested strong concordance (ρ_c=.82, 95% CI 0.72-0.91 and ρ_c=.86, 95% CI 0.76-0.94, respectively; not shown).

For the immediate and delayed recall tasks, this study assessed differences in association between the number of correctly recalled words by study sample and word list assignment (Table 2), as well as sociodemographic differences between samples (Table 3). Results from the regression analyses were interpreted using a base scenario that represents the median sociodemographic characteristics of the sample (ie, the average number of words that are correctly recalled by a white female aged 50-54 years who is married, has a high school diploma, and self-reports her current memory as good). For immediate or delayed recall, WHV respondents recalled significantly more words than HRS respondents, except for List 3 in delayed recall. For both WHV and HRS respondents, the number of correctly recalled words varied significantly depending on which list was assigned; however, these differences were small (<0.28 words).

XSL•FO
RenderX

Table 1. Respondent characteristics by modality.

	WHV	HRS	*P* value
Number of respondents	4474	12,545	
Number of tasks per respondent, median (IQR)	1	3 (2-5)	
Total number of tasks	4474	43,417	
Age in years, median (IQR)	53 (48-61)	60 (55-64)	<.001
40-44, n (%)	629 (14.17)	735 (1.68)	
45-49, n (%)	754 (16.98)	2158 (4.94)	
50-54, n (%)	1051 (23.67)	7701 (17.63)	
55-59, n (%)	661 (14.89)	10, 951 (25.07)	
60-64, n (%)	641 (14.44)	11,260 (25.78)	
65-69, n (%)	704 (15.86)	10,868 (24.88)	
Race			<.001
White, n (%)	3556 (80.09)	33,992 (75.14)	
Black, n (%)	632 (14.23)	8832 (19.52)	
Other, n (%)	252 (5.68)	2412 (5.33)	
Hispanic ethnicity			<.001
No, n (%)	3743 (84.30)	39,604 (87.27)	
Yes, n (%)	697 (15.70)	5774 (12.72)	
Educational attainment			<.001
No degree, n (%)	168 (3.78%)	8334 (18.40)	
High school diploma/GED, n (%)	1955 (44.03)	24,941 (55.05)	
Associates degree/some college, n (%)	1257 (28.31)	2747 (6.06)	
Bachelor's degree, n (%)	669 (15.07)	5664 (12.50)	
Master's degree, n (%)	320 (7.21)	3198 (7.06)	
Law/MD/PhD, n (%)	71 (1.60)	419 (0.92)	
Marital status			<.001
Married, n (%)	2338 (53.78)	28,333 (62.35)	
Partnered, n (%)	231 (5.20)	2069 (4.55)	
Separated/divorced, n (%)	1048 (23.60)	8654 (19.04)	
Widowed, n (%)	262 (5.90)	4922 (10.83)	
Never married, n (%)	511 (11.51)	1467 (3.23)	
Self-reported current memory			<.001
Excellent, n (%)	468 (10.54)	2518 (5.77)	
Very good, n (%)	1743 (39.04)	11,353 (26.00)	
Good, n (%)	1704 (38.38)	18,991 (43.48)	
Fair, n (%)	486 (10.93)	9137 (20.92)	
Poor, n (%)	48 (1.08)	1654 (3.79)	

XSL•FO
RenderX

Figure 1. Likelihood of immediate recall by word.

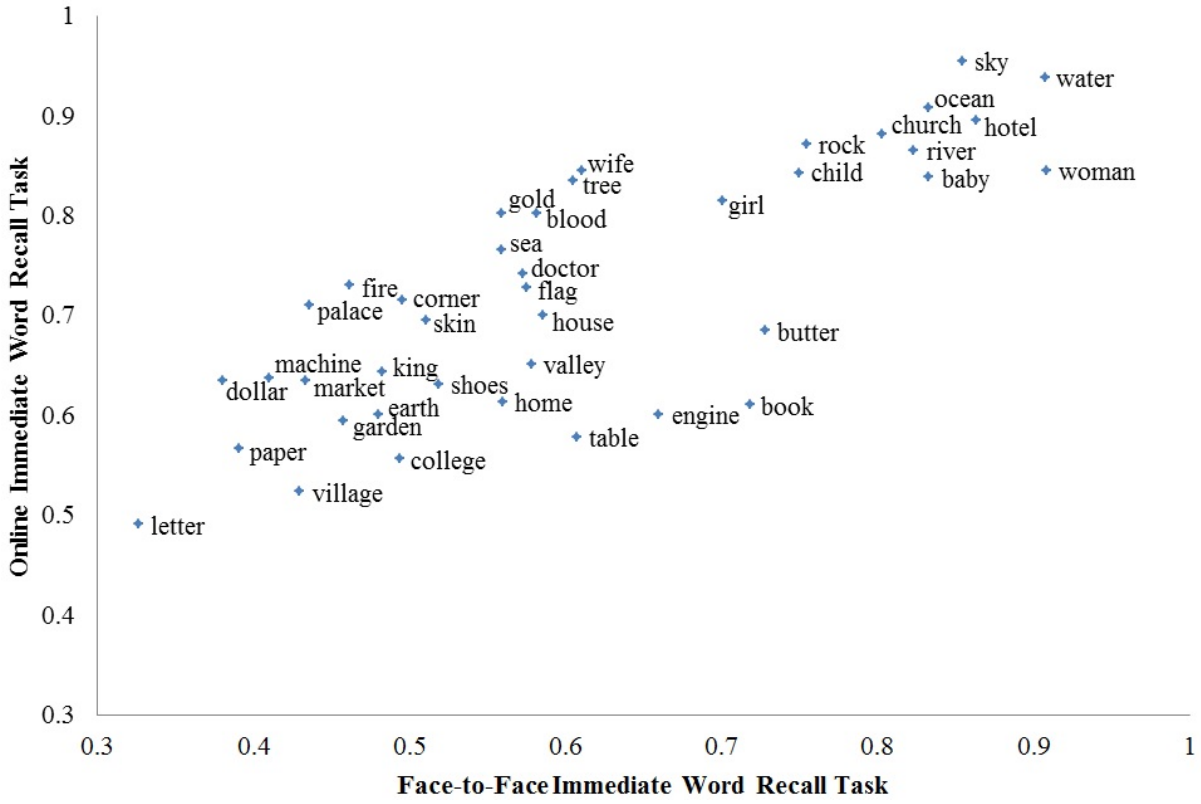


Table 2. Average number of correctly recalled words by list and modality.

	Immediate recall			Delayed recall		
	WHV	HRS	*P* value[a]	WHV	HRS	*P* value[a]
Overall	7.24	6.06	1	5.41	5.21	<.001
List 1[a]	7.40	6.14	.<001	5.57	5.24	<.001
List 2[a]	7.16	5.93	.<001	5.29	5.08	.013
List 3[a]	7.12	6.12	<.002	5.27	5.28	.922
List 4[a]	7.30	6.07	<.001	5.50	5.23	.002

[a]Significant differences were detected between lists for immediate (P_{WHV} < .001 and P_{HRS} < .001) and delayed (P_{WHV} =.002 and P_{HRS} < .001) word recall tasks.

XSL•FO
RenderX

Table 3. Associated respondent characteristics and number of correctly recalled words by survey modality: WHV versus HRS.

	Immediate recall			Delayed recall		
	WHV	HRS	*P* value[a]	WHV	HRS	*P* value[a]
Constant[b]	7.19[d]	6.34[d]	<.001	5.62[d]	5.48[d]	<.001
Age in years						
40-44	−.04	.25[d]	.01	−.13	.31[d]	<.001
45-49	−.01	.01	.84	−.05	.01	.65
50-54	—	—	—	—	—	—
55-59	.14	−.01	.10	−.04	.04	.57
60-64	.01	−.07[d]	.48	−.10	.02	.39
65-69	.02	−.24[d]	<.001	.15	−.16[d]	.03
Race						
White	—	—	—	—	—	—
Black	−.55[d]	−.50[d]	.57	−.71[d]	−.74[d]	.86
Other	−.16	−.34[d]	.26	−.32	−.41[d]	.70
Hispanic ethnicity						
No	—	—	—	—	—	—
Yes	−.19[c]	−.36[d]	.08	−.10	−.35[d]	.09
Educational attainment						
No degree	−.36[c]	−.61[d]	.14	−.25	−.64[d]	.07
High school diploma/GED	—	—	—	—	—	—
Associate's degree/ some college	.12	.19[d]	.35	−.13	.18[d]	<.001
Bachelor's degree	.35[d]	.47[d]	.19	.26[c]	.50[d]	.07
Master's degree	.60[d]	.61[d]	.92	.40[c]	.68[d]	.12
Law/MD/PhD	.67[d]	1.01[d]	.13	.52	1.09[d]	.16
Marital						
Married	—	—	—	—	—	—
Partnered	−.14	−.10[d]	.79	.17	−.11[c]	.17
Separated/divorced	−.02	−.07[d]	.45	−.17	−.10[d]	.52
Widowed	.04	−.04	.48	.05	−.03	.67
Never married	.05	−.17[d]	.04	−.05	−.14[c]	.56
Self-reported current memory						
Excellent	.18	.13[d]	.64	.12	.12[d]	.99
Very good	.22[d]	.21[d]	.97	.07	.24[d]	.10
Good	—	—	—	—	—	—
Fair	−.52[d]	−.31[d]	.03	−.60[d]	−.35[d]	<.001
Poor	−1.53[d]	−.74[d]	.01	−1.93[d]	−.85[d]	<.001

[a]Represents *P* value for H^0: No difference between online and face-to face.

[b]Base scenario represents the average number of words that are correctly recalled by a white female, aged 50-54 years, who is married, has a high school

XSL•FO
RenderX

diploma, and self-reports her current memory as good.

[c]*P* value <.05.

[d]*P* value <.01.

Immediate Word Recall

Immediate word recall was significantly associated with respondent characteristics in WHV and HRS tasks, and there were significant modality differences between the online and HRS studies. Overall, WHV respondents immediately recalled about one more word (0.85) than HRS respondents did, after adjusting for respondent characteristics. In terms of demographics, age was significantly associated with immediate recall for the HRS task but not the WHV task. Specifically, younger respondents recalled more words than older respondents in the HRS tasks but not in the WHV tasks. Non-white and/or Hispanic respondents were significantly associated with reduced immediate recall for either modality; however, their associations were not significantly different by modality.

Levels of educational attainment were significantly associated with immediate recall for both the HRS and WHV tasks. Detrimental effects were seen for the lowest education level; respondents with less than a high school diploma recalled fewer words. The benefits of obtaining education beyond high school were incrementally significant, with the exception of WHV respondents who earned an associate's degree. Marital status was significantly associated with immediate recall in the HRS tasks but not the WHV tasks. Specifically, respondents who reported being partnered, separated, divorced, or never married recalled fewer words than their married counterparts. However, the only associations that differed significantly between modalities were those for individuals who were never married.

Self-reported current memory was significantly associated with immediate word recall in both modalities. As expected, those who reported their memory as excellent or very good were more likely to recall more words than those with a fair or poor memory. However, it is unclear whether those who reported excellent memory had better recall than those who reported very good memory. The association between a poor memory and immediate word recall was statistically significant with a noteworthy effect (1.53 words less than good memory). The association with fair or poor was greater for the WHV task than the HRS task, possibly because of interviewer biases (eg, slowing the task for persons who reported poor memory).

Delayed Word Recall

As with immediate word recall, the associations between respondent characteristics and delayed word recall were significant, and their associations differed by modality. Adjusting for respondent characteristics, WHV respondents recalled approximately 0.14 more words after a delay than HRS respondents. Like the immediate recall results, the association between age and delayed recall was significant for the HRS task but not the WHV task. For both modalities, respondents who were Non-white and/or Hispanic performed significantly worse on the delayed recall tasks, but the associations did not differ significantly.

Levels of educational attainment were significantly associated for both modalities and differed slightly from what was seen for the immediate recall task. Significant detrimental effects were no longer seen for WHV respondents with less than a high school diploma but persisted for HRS respondents. Higher levels of education beyond an associate's degree remained significantly associated with greater delayed recall, with the exception of WHV respondents who earned an associate's or advanced degree. The association between advanced education levels and recall was very strong for HRS respondents, who recalled approximately 0.50 more words compared to similarly educated WHV respondents. Marital status was significantly associated with delayed recall for the HRS modality but not the online modality. HRS respondents who reported being partnered, separated or divorced, or never married recalled significantly fewer words compared to married respondents. The associations between modalities were not significantly different.

Self-reported current memory was significantly associated with delayed word recall in both modalities. Similar to the immediate recall task, respondents who reported their memory as excellent or very good were more likely to recall more words than those with a fair or poor memory. The association between poor memory and delayed recall intensified for WHV respondents, who recalled nearly 2 words less compared to the base scenario and more than 1 word less compared to HRS respondents with a similar memory rating.

In order to explore the possibility that word recall scores for WHV respondents were influenced by literacy level and typing skills (ie, misspelled words would not be counted as correct), the previous analyses were rerun after correcting words that were misspelled by one letter. This arbitrary adjustment was based on the number of WHV responses that appeared to be related to misspellings (eg, doller for dollar) or mistyping (eg, ovean for ocean), and is akin to the best-judgment practice granted to HRS interviewers when determining whether a HRS response should be counted as correct (eg, woman for women or shoe for shoes). When the analyses were rerun using the spell-corrected word counts, no significant differences were seen for any of the results. Therefore, the results reported here were conducted using the uncorrected word recall responses for WHV respondents.

Discussion

Principal Findings

This study compared and characterized the results of the WHV word recall task to those of a gold standard HRS word recall task in order to determine reliability for future surveys. The results of this study provide support for the inclusion of online cognitive assessments in health surveys. This is the first study attempting to replicate the HRS word recall tasks in an application-based assessment. The results indicate that the immediate and delayed word recall tasks were equivalent to the HRS tasks, as evidenced by high levels of concordance

XSL•FO
RenderX

(precision) and association with self-reported memory (convergent validity). Even after controlling for age, education, and self-reported memory, WHV respondents recalled nearly one more word than HRS respondents for the immediate recall tasks. This difference decreased but remained significant for the delayed recall and may be attributed to study design differences or other unobservable sample selection biases. In summary, both HRS and WHV tasks appear to perform well despite key differences between the studies.

While our normalized results demonstrated a high level of concordance between the WHV and HRS tasks and thus support the primary goal of this study, we did note significant differences between samples that may be related to a number of potential confounders, such as differences in study design. For example, the HRS recall lists were presented verbally, whereas the words of the WHV lists were presented visually. Upon initial review, one may think that differences in how the brain processes auditory versus visual information may contribute to modality differences. However, research has shown that auditory and visual recall tasks activate overlapping regions of the brain, and while the left hemisphere of the brain is activated slightly more during visual tasks, there is no evidence that recall performance is impacted by modality [40].

An additional difference in study design is the length of time and type of activities that were completed by respondents between the immediate and delayed recall tasks. HRS respondents answered questions regarding their emotional state over the past week (eg, levels of motivation, happiness, and loneliness) and completed two mental math tasks (ie, counting backwards and subtracting 7s) for 5 minutes. WHV respondents completed a series of DCE tasks during the 20-minute delay, which may arguably require greater levels of cognitive engagement. These dissimilarities in the amount of delay and the complexity of the tasks completed during the delay may have contributed to the observed modality differences. The regression analysis may control for some of the sample selection issues, but panel and delay attributes may also explain differences by modality.

In addition to modality differences, there is a potential concern for practice effects to bias the results of repeated word recall tasks, particularly since such effects mask true declines in cognitive performance [41]. Practice effects have been associated with the cognitive data of the HRS [33,35]; however, the interpretation of these results is muddied by the complex methodology of the earliest waves of data collection. For example, Rodgers et al examined practice effects in the word recall tasks of the 1993 and 1995 waves of the Asset and Health Dynamics Among the Oldest Old Study (AHEAD) to word recall performance of the 1998 and 2000 waves of the HRS (the AHEAD and HRS were merged in 1998 due to methodological and content similarities) [33]. Although significant practice effects were identified from wave 1 (1993) to wave 2 (1995) and from wave 2 to wave 3 (1998), none were identified from wave 3 to wave 4 (2000) [33]. The authors note these results are difficult to interpret given the considerable methodological changes that were made from wave to wave, most notable of which is the implementation of the counterbalanced word recall list assignment in wave 2 of AHEAD (1995). Additionally, there is the possibility that the original word list used in 1993 was simply more difficult compared to word lists used in subsequent waves [33].

In a more recent analysis, McArdle et al found evidence of practice effects in cognitive data from earlier waves of the HRS (1992-2004) [35]; however, this result may also be affected by substantive changes in study design. Specifically, the word recall tests of 1992 and 1994 included only one word list with 20 nouns; the counterbalanced approach of randomly assigned four lists of 10 words was first implemented with the HRS in 1996. As with the results of the previous study, the presence of practice effects could be attributed to respondents receiving the same list of words in 1992 and 1994. Additionally, greater levels of recall in subsequent waves could be attributed to the fact that respondents may find it easier to recall 10 words as opposed to 20.

These methodological changes clearly restrict the interpretability of potential practice effects noted within the HRS. The results of the current study are less subjective to such biases since the analyses are restricted to the 2000-2012 waves of the HRS (ie, the counterbalanced assignment of word recall lists is uniform across waves). Despite this counterbalanced approach, it is not possible to completely rule out the potential influence of practice effects. Future studies should attempt to measure the presence and impact of practice effects in the HRS using only the waves with identical methodological approaches.

We also found several interesting associations between episodic memory performance and sociodemographic characteristics. The effect of marital status on word recall was significant only for HRS respondents; individuals who were partnered, separated or divorced, or never married performed worse compared to those who were married. The presence of significant results in the HRS sample but not the WHV sample may be related to the fact that married/partnered HRS respondents are often interviewed one after the other. Previous research has indicated that spouses who are interviewed second may be at a disadvantage in free recall tasks [42], possibly due to the fact the first interviewed spouse may be healthier. Another possible explanation of these results is that those who are partnered have been shown to perform better on episodic memory tasks in general compared to non-partnered individuals [35].

Education was another sociodemographic characteristic that was significantly associated with word recall performance, with higher levels of education significantly predicting higher episodic memory performance. Higher levels of education are thought to influence cognitive function by increasing individual levels of brain and cognitive reserve [43]. Brain reserve refers to the inherent efficiency and capability of the brain to support and execute cognitive functions [43]. Conversely, cognitive reserve represents the brain's ability to maintain this efficiency despite the accumulation of structural and neural damage that occurs as a result of natural aging, disease, or injury [43]. Increased levels of cognitive reserve may be particularly beneficial during later stages of life [44-46]. Previous researchers have argued against controlling for the impact of education, stating that growing levels of education represent cohort trends that contribute to overall increases in cognitive

XSL•FO
RenderX

performance [33]. However, it is possible that other factors associated with higher education (eg, increased socioeconomic status, better nutrition, greater availability of resources) may have attributed to this positive relationship.

While several computer-based cognitive batteries have been developed [47,48] to date, these have lacked correspondence to HRS tasks used in large cohort studies. The goal of the current study was to develop an application-based cognitive measure for episodic memory that could be easily used in future research studies and health assessments. The potential benefits of such online tasks can be inferred from evidence showing that including short cognitive tests as a part of a routine evaluation in the clinical or community setting aids in the early detection of cognitive decline. Individuals who self-report problems with memory may be more aware of adverse changes in cognitive performance [49]. Additionally, older adults who report problems with memory but perform normally have been shown to have structural brain changes similar to those seen in mild cognitive impairment [50].

Future Research

Future research should assess additional cognitive tasks included in the HRS. This type of research might expand the results of the current study to investigate the effects of setting (eg, waiting room, hospital room, home use of online tasks) or to support the use of routine online cognitive assessments to track cognitive change in healthy older adults or clinical populations. Furthermore, clear standards for measurement using online tasks similar to the electronic patient-reported outcome literature should be created [51]. Development of such standards is likely complicated by the fact that device and software technology continues to evolve and age-related rates of cognitive change vary across a range of domains and birth cohorts with varying computer aptitudes [52,53].

Limitations

A key limitation of the study is the use of an existing panel in the community setting. While some may argue that sampling bias is introduced by using research panels who demonstrate high levels of technological capabilities (ie, use of computers, smartphones, tablets), it has also been noted that such panels allow researchers to collect large amounts of data from diverse populations [2]. A further limitation is the lack of access to medical records that verify quality of self-reported health. Older individuals tend to rate their health more highly than younger individuals despite increases in chronic medical problems [54-56], and this overestimation of health may inadvertently bias results. The biases associated with self-reported health and behavior measures are well documented; however, expanding the current research into clinical settings would alleviate this issue. Also, the community setting adds a lack of environmental control (eg, interruptions) that may increase variability. A future project may compare interview-based and application-based tasks in a clinical population (eg, Alzheimer patients) during set times. Additionally, the current study focuses on episodic memory; in order to obtain a more robust estimation of cognitive abilities, future efforts should identify the correspondence between interview-based and online versions of other cognitive assessments of such as measures of semantic memory and vocabulary.

Inability to monitor respondent behavior is a limitation of online and telephone surveys [1]. For example, respondents of online or telephone word recall tasks could have written down the words on paper as they were presented. Examination of eye-tracking or client-side paradata [57] (ie, information about respondent behavior recorded by respondents' computers, such as the number of times and locations of mouse clicks) has the potential to be extremely valuable in the analysis of online survey data. Nevertheless, further technological advancements are needed before such evidence can be incorporated into cognitive measures.

In summary, this study found a high level of convergent validity between the WHV and HRS word recall tasks, after controlling for age, education, and self-reported memory. Use of application-based cognitive assessments should continue to expand in community research and clinical settings, but greater efforts need to be made in regards to validating such online measures. Additionally, researchers should be wary of a number of potential biases, including modality differences, retest effects, and gender differences in cognitive performance.

Acknowledgments

The authors thank Carol Templeton, Michelle Owens, and Nawreen Jahan at Moffitt Cancer Center for their contributions to the research and creation of this paper.

Conflicts of Interest

None declared.

References

1. Lee H, Baniqued PL, Cosman J, Mullen S, McAuley E, Severson J, et al. Examining cognitive function across the lifespan using a mobile application. Comput Hum Behav 2012 Sep;28(5):1934-1946. [doi: 10.1016/j.chb.2012.05.013]
2. Dufau S, Duñabeitia JA, Moret-Tatay C, McGonigal A, Peeters D, Alario FX, et al. Smart phone, smart science: how the use of smartphones can revolutionize research in cognitive science. PLoS One 2011;6(9):e24974 [FREE Full text] [doi: 10.1371/journal.pone.0024974] [Medline: 21980370]

XSL•FO
RenderX

3. Wild K, Howieson D, Webbe F, Seelye A, Kaye J. Status of computerized cognitive testing in aging: a systematic review. Alzheimers Dement 2008 Nov;4(6):428-437 [FREE Full text] [doi: 10.1016/j.jalz.2008.07.003] [Medline: 19012868]
4. Bohannon J. Human subject research: social science for pennies. Science 2011 Oct 21;334(6054):307. [doi: 10.1126/science.334.6054.307] [Medline: 22021834]
5. Schonlau M, van Soest A, Kapteyn A, Couper M. Selection bias in web surveys and the use of propensity scores. Sociol Methods Res 2009 Feb 01;37(3):291-318. [doi: 10.1177/0049124108327128]
6. Parsey CM, Schmitter-Edgecombe M. Applications of technology in neuropsychological assessment. Clin Neuropsychol 2013 Nov;27(8):1328-1361 [FREE Full text] [doi: 10.1080/13854046.2013.834971] [Medline: 24041037]
7. Borson S, Frank L, Bayley PJ, Boustani M, Dean M, Lin PJ, et al. Improving dementia care: the role of screening and detection of cognitive impairment. Alzheimers Dement 2013 Mar;9(2):151-159 [FREE Full text] [doi: 10.1016/j.jalz.2012.08.008] [Medline: 23375564]
8. Grodstein F. How early can cognitive decline be detected? Brit Med J 2012 Jan 05;344(jan04 4):d7652-d7652. [doi: 10.1136/bmj.d7652]
9. Sliwinski MJ, Smyth JM, Hofer SM, Stawski RS. Intraindividual coupling of daily stress and cognition. Psychol Aging 2006 Sep;21(3):545-557 [FREE Full text] [doi: 10.1037/0882-7974.21.3.545] [Medline: 16953716]
10. Stawski RS, Mogle J, Sliwinski MJ. Intraindividual coupling of daily stressors and cognitive interference in old age. J Gerontol B Psychol Sci Soc Sci 2011 Jul;66 Suppl 1:i121-i129 [FREE Full text] [doi: 10.1093/geronb/gbr012] [Medline: 21743045]
11. Boustani M, Callahan CM, Unverzagt FW, Austrom MG, Perkins AJ, Fultz BA, et al. Implementing a screening and diagnosis program for dementia in primary care. J Gen Intern Med 2005 Jul;20(7):572-577 [FREE Full text] [doi: 10.1111/j.1525-1497.2005.0126.x] [Medline: 16050849]
12. Alagiakrishnan K, Zhao N, Mereu L, Senior P, Senthilselvan A. Montreal Cognitive Assessment is superior to Standardized Mini-Mental Status Exam in detecting mild cognitive impairment in the middle-aged and elderly patients with type 2 diabetes mellitus. Biomed Res Int 2013;2013:186106 [FREE Full text] [doi: 10.1155/2013/186106] [Medline: 23936778]
13. Velayudhan L, Ryu SH, Raczek M, Philpot M, Lindesay J, Critchfield M, et al. Review of brief cognitive tests for patients with suspected dementia. Int Psychogeriatr 2014 Aug;26(8):1247-1262 [FREE Full text] [doi: 10.1017/S1041610214000416] [Medline: 24685119]
14. Edmonds EC, Delano-Wood L, Galasko DR, Salmon DP, Bondi MW. Subjective cognitive complaints contribute to misdiagnosis of mild cognitive impairment. J Int Neuropsychol Soc 2014 Sep;20(8):836-847 [FREE Full text] [doi: 10.1017/S135561771400068X] [Medline: 25156329]
15. Albert M, Blacker D, Moss MB, Tanzi R, McArdle JJ. Longitudinal change in cognitive performance among individuals with mild cognitive impairment. Neuropsychology 2007 Mar;21(2):158-169. [doi: 10.1037/0894-4105.21.2.158] [Medline: 17402816]
16. Mickes L, Wixted JT, Fennema-Notestine C, Galasko D, Bondi MW, Thal LJ, et al. Progressive impairment on neuropsychological tasks in a longitudinal study of preclinical Alzheimer's disease. Neuropsychology 2007 Nov;21(6):696-705. [doi: 10.1037/0894-4105.21.6.696] [Medline: 17983283]
17. Tulving E. Episodic and semantic memory. In: Tulving E, Donaldson E, editors. Organization of Memory. New York: Academic Press; 1972:381-403.
18. Small BJ, Dixon RA, McArdle JJ, Grimm KJ. Do changes in lifestyle engagement moderate cognitive decline in normal aging? Evidence from the Victoria Longitudinal Study. Neuropsychology 2012 Mar;26(2):144-155 [FREE Full text] [doi: 10.1037/a0026579] [Medline: 22149165]
19. Tulving E. Episodic memory: from mind to brain. Annu Rev Psychol 2002;53:1-25. [doi: 10.1146/annurev.psych.53.100901.135114] [Medline: 11752477]
20. Salmon DP, Ferris SH, Thomas RG, Sano M, Cummings JL, Sperling RA, et al. Age and apolipoprotein E genotype influence rate of cognitive decline in nondemented elderly. Neuropsychology 2013 Jul;27(4):391-401 [FREE Full text] [doi: 10.1037/a0032707] [Medline: 23876113]
21. Head D, Rodrigue KM, Kennedy KM, Raz N. Neuroanatomical and cognitive mediators of age-related differences in episodic memory. Neuropsychology 2008 Jul;22(4):491-507 [FREE Full text] [doi: 10.1037/0894-4105.22.4.491] [Medline: 18590361]
22. Mayes AR, Roberts N. Theories of episodic memory. Philos Trans R Soc Lond B Biol Sci 2001 Sep 29;356(1413):1395-1408 [FREE Full text] [doi: 10.1098/rstb.2001.0941] [Medline: 11571031]
23. Li SC, Rieckmann A. Neuromodulation and aging: implications of aging neuronal gain control on cognition. Curr Opin Neurobiol 2014 Dec;29:148-158. [doi: 10.1016/j.conb.2014.07.009] [Medline: 25064177]
24. Ford JH, Kensinger EA. The relation between structural and functional connectivity depends on age and on task goals. Front Hum Neurosci 2014;8:307 [FREE Full text] [doi: 10.3389/fnhum.2014.00307] [Medline: 24904351]
25. Addis DR, Leclerc CM, Muscatell KA, Kensinger EA. There are age-related changes in neural connectivity during the encoding of positive, but not negative, information. Cortex 2010 Apr;46(4):425-433 [FREE Full text] [doi: 10.1016/j.cortex.2009.04.011] [Medline: 19555933]

XSL•FO
RenderX

26. Dixon RA, de Frias CM. Cognitively elite, cognitively normal, and cognitively impaired aging: neurocognitive status and stability moderate memory performance. J Clin Exp Neuropsychol 2014;36(4):418-430 [FREE Full text] [doi: 10.1080/13803395.2014.903901] [Medline: 24742143]
27. Albert M, Soldan A, Gottesman R, McKhann G, Sacktor N, Farrington L, et al. Cognitive changes preceding clinical symptom onset of mild cognitive impairment and relationship to ApoE genotype. Curr Alzheimer Res 2014;11(8):773-784 [FREE Full text] [Medline: 25212916]
28. Caselli RJ, Dueck AC, Osborne D, Sabbagh MN, Connor DJ, Ahern GL, et al. Longitudinal modeling of age-related memory decline and the APOE epsilon 4 effect. N Engl J Med 2009 Jul 16;361(3):255-263 [FREE Full text] [doi: 10.1056/NEJMoa0809437] [Medline: 19605830]
29. Klekociuk SZ, Summers JJ, Vickers JC, Summers MJ. Reducing false positive diagnoses in mild cognitive impairment: the importance of comprehensive neuropsychological assessment. Eur J Neurol 2014 Oct;21(10):1330-1336. [doi: 10.1111/ene.12488] [Medline: 24943259]
30. Lin LI. A concordance correlation coefficient to evaluate reproducibility. Biometrics 1989 Mar;45(1):255-268. [Medline: 2720055]
31. Kelly PA, O'Malley KJ, Kallen MA, Ford ME. Integrating validity theory with use of measurement instruments in clinical settings. Health Serv Res 2005 Oct;40(5 Pt 2):1605-1619 [FREE Full text] [doi: 10.1111/j.1475-6773.2005.00445.x] [Medline: 16178998]
32. Sechrest L. Validity of measures is no simple matter. Health Serv Res 2005;40:1584-1604. [doi: 10.1111/j.1475-6773.2005.00043.x]
33. Rodgers WL, Ofstedal MB, Herzog AR. Trends in scores on tests of cognitive ability in the elderly U.S. population, 1993-2000. J Gerontol B Psychol Sci Soc Sci 2003 Nov;58(6):S338-S346. [Medline: 14614121]
34. Greenberg DL, Verfaellie M. Interdependence of episodic and semantic memory: evidence from neuropsychology. J Int Neuropsychol Soc 2010 Sep;16(5):748-753 [FREE Full text] [doi: 10.1017/S1355617710000676] [Medline: 20561378]
35. McArdle JJ, Fisher GG, Kadlec KM. Latent variable analyses of age trends of cognition in the Health and Retirement Study, 1992-2004. Psychol Aging 2007 Sep;22(3):525-545. [doi: 10.1037/0882-7974.22.3.525] [Medline: 17874952]
36. Craig B, Schell M, Brown P, Reeve B, Cella D, Hays R, et al. HRQoL Values for Cancer Survivors: Enhancing PROMIS Measures for CER. Tampa FL: H Lee Moffitt Cancer Center; 2011.
37. Craig B, Owens M. Methods Report of the Women's Health Valuation Study. Methods report on the Women's Health Valuation (WHV): Year 1 Moffitt Cancer Center; 2013. URL: http://labpages.moffitt.org/craigb/Publications/WHVMethods_140106.pdf: [accessed 2015-05-07] [WebCite Cache ID 6YLq8ZkuX]
38. Ofstedal MB, Fisher GG, Herzog AR. Documentation of cognitive functioning measures in the Health and Retirement Study. Ann Arbor, MI: Survey Research Center (University of Michigan); 2005.
39. Wechsler D. Wechsler Adult Intelligence Scale-Revised. New York: Psychological Corporation; 1981.
40. Crottaz-Herbette S, Anagnoson RT, Menon V. Modality effects in verbal working memory: differential prefrontal and parietal responses to auditory and visual stimuli. Neuroimage 2004 Jan;21(1):340-351. [Medline: 14741672]
41. Rabbitt P, Diggle P, Holland F, McInnes L. Practice and drop-out effects during a 17-year longitudinal study of cognitive aging. J Gerontol B Psychol Sci Soc Sci 2004 Mar;59(2):P84-P97. [Medline: 15014091]
42. Herzog AR, Rodgers WL. Cognitive performance measures in survey research on older adults. In: Schwarz N, Park DC, Knauper B, Sudman S, editors. Cognition, Aging, and Self-Reports. Philadelphia PA: Psychology Press; 1999.
43. Stern Y. Cognitive reserve. Neuropsychologia 2009 Aug;47(10):2015-2028 [FREE Full text] [doi: 10.1016/j.neuropsychologia.2009.03.004] [Medline: 19467352]
44. Deary IJ, Corley J, Gow AJ, Harris SE, Houlihan LM, Marioni RE, et al. Age-associated cognitive decline. Br Med Bull 2009;92:135-152. [doi: 10.1093/bmb/ldp033] [Medline: 19776035]
45. Perlmutter M, Nyquist L. Relationships between self-reported physical and mental health and intelligence performance across adulthood. J Gerontol 1990 Jul;45(4):P145-PP55. [Medline: 2365970]
46. Fritsch T, McClendon MJ, Smyth KA, Lerner AJ, Friedland RP, Larsen JD. Cognitive functioning in healthy aging: the role of reserve and lifestyle factors early in life. Gerontologist 2007 Jun;47(3):307-322. [Medline: 17565095]
47. Wesnes K. Assessing cognitive function in clinical trials: latest developments and future directions. Drug Discov Today 2002 Jan 1;7(1):29-35. [Medline: 11790601]
48. Makdissi M, Collie A, Maruff P, Darby DG, Bush A, McCrory P, et al. Computerised cognitive assessment of concussed Australian Rules footballers. Br J Sports Med 2001;35:354-360. [doi: 10.1136/bjsm.35.5.354]
49. Kalbe E, Salmon E, Perani D, Holthoff V, Sorbi S, Elsner A, et al. Anosognosia in very mild Alzheimer's disease but not in mild cognitive impairment. Dement Geriatr Cogn Disord 2005;19(5-6):349-356. [doi: 10.1159/000084704] [Medline: 15802909]
50. Saykin AJ, Wishart HA, Rabin LA, Santulli RB, Flashman LA, West JD, et al. Older adults with cognitive complaints show brain atrophy similar to that of amnestic MCI. Neurology 2006 Sep 12;67(5):834-842 [FREE Full text] [doi: 10.1212/01.wnl.0000234032.77541.a2] [Medline: 16966547]
51. Coons SJ, Gwaltney CJ, Hays RD, Lundy JJ, Sloan JA, Revicki DA, et al. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ISPOR ePRO

XSL•FO
RenderX

Good Research Practices Task Force report. Value Health 2009 Jun;12(4):419-429. [doi: 10.1111/j.1524-4733.2008.00470.x] [Medline: 19900250]
52. Small BJ, Dixon RA, McArdle JJ. Tracking cognition-health changes from 55 to 95 years of age. J Gerontol B Psychol Sci Soc Sci 2011 Jul;66 Suppl 1:i153-i161 [FREE Full text] [doi: 10.1093/geronb/gbq093] [Medline: 21196437]
53. Dixon RA, Wahlin A, Maitland SB, Hultsch DF, Hertzog C, Bäckman L. Episodic memory change in late adulthood: generalizability across samples and performance indices. Mem Cognit 2004 Jul;32(5):768-778. [Medline: 15552354]
54. Vuorisalmi M, Lintonen T, Jylhä M. Comparative vs global self-rated health: associations with age and functional ability. Aging Clin Exp Res 2006 Jun;18(3):211-217. [Medline: 16804367]
55. Jylhä M, Guralnik JM, Balfour J, Fried LP. Walking difficulty, walking speed, and age as predictors of self-rated health: the women's health and aging study. J Gerontol A Biol Sci Med Sci 2001 Oct;56(10):M609-M617. [Medline: 11584033]
56. Jylhä M. What is self-rated health and why does it predict mortality? Towards a unified conceptual model. Soc Sci Med 2009 Aug;69(3):307-316. [doi: 10.1016/j.socscimed.2009.05.013] [Medline: 19520474]
57. Heerwegh D. Explaining response latencies and changing answers using client-side paradata from a web survey. Soc Sci Comput Rev 2003 Aug 01;21(3):360-373. [doi: 10.1177/0894439303253985]

Abbreviations

DCE: discrete choice experiment
HRS: Health and Retirement Study
WHV: Women's Health Valuation Study

Edited by G Eysenbach; submitted 24.10.14; peer-reviewed by P Wicks, C Brett; comments to author 10.01.15; revised version received 12.02.15; accepted 18.03.15; published 02.06.15.

Please cite as:
Runge SK, Craig BM, Jim HS
Word Recall: Cognitive Performance Within Internet Surveys
JMIR Mental Health 2015;2(2):e20
URL: http://mental.jmir.org/2015/2/e20/
doi:10.2196/mental.3969
PMID:26543924

©Shannon K Runge, Benjamin M Craig, Heather S Jim. Originally published in JMIR Mental Health (http://mental.jmir.org), 02.06.2015. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Review

Online Peer-to-Peer Support for Young People With Mental Health Problems: A Systematic Review

Kathina Ali[1], BSc (Psych); Louise Farrer[1], PhD (Clinical Psych); Amelia Gulliver[1], PhD; Kathleen M Griffiths[1], PhD

National Institute for Mental Health Research, The Australian National University, Canberra, Australia

Corresponding Author:
Kathina Ali, BSc (Psych)
National Institute for Mental Health Research
The Australian National University
Building 63, Eggleston Road
Canberra,
Australia
Phone: 61 261259155
Fax: 61 261250733
Email: kathina.ali@anu.edu.au

Abstract

Background: Adolescence and early adulthood are critical periods for the development of mental disorders. Online peer-to-peer communication is popular among young people and may improve mental health by providing social support. Previous systematic reviews have targeted Internet support groups for adults with mental health problems, including depression. However, there have been no systematic reviews examining the effectiveness of online peer-to-peer support in improving the mental health of adolescents and young adults.

Objective: The aim of this review was to systematically identify available evidence for the effectiveness of online peer-to peer support for young people with mental health problems.

Methods: The PubMed, PsycInfo, and Cochrane databases were searched using keywords and Medical Subject Headings (MeSH) terms. Retrieved abstracts (n=3934) were double screened and coded. Studies were included if they (1) investigated an online peer-to-peer interaction, (2) the interaction discussed topics related to mental health, (3) the age range of the sample was between 12 to 25 years, and (4) the study evaluated the effectiveness of the peer-to-peer interaction.

Results: Six studies satisfied the inclusion criteria for the current review. The studies targeted a range of mental health problems including depression and anxiety (n=2), general psychological problems (n=1), eating disorders (n=1), and substance use (tobacco) (n=2). The majority of studies investigated Internet support groups (n=4), and the remaining studies focused on virtual reality chat sessions (n=2). In almost all studies (n=5), the peer support intervention was moderated by health professionals, researchers or consumers. Studies employed a range of study designs including randomized controlled trials (n=3), pre-post studies (n=2) and one randomized trial. Overall, two of the randomized controlled trials were associated with a significant positive outcome in comparison to the control group at post-intervention. In the remaining four studies, peer-to-peer support was not found to be effective.

Conclusions: This systematic review identified an overall lack of high-quality studies examining online peer-to-peer support for young people. Given that peer support is frequently used as an adjunct to Internet interventions for a variety of mental health conditions, there is an urgent need to determine the effectiveness of peer support alone as an active intervention.

(JMIR Mental Health 2015;2(2):e19) doi:10.2196/mental.4418

KEYWORDS

mental health; Internet; young people; peer-to-peer support; Internet support groups; technology; systematic reviews

Introduction

The mental health and well-being of young people is a major public health concern [1]. Findings suggest that one in four young people aged 16 to 25 years have experienced at least one mental health problem during their lifetime [2]. Although young people have one of the highest prevalence rates for mental health problems, their needs are often unmet and access to mental health services is limited [1]. Previous research has identified

XSL•FO
RenderX

various barriers to treatment including concerns about confidentiality, lack of knowledge of resources, cost, and inaccessibility of services [3]. More broadly, young people are reluctant to seek professional face-to-face help for their problems [4].

It has been reported that a growing number of young people use the Internet to seek help and information regarding their mental health [5], and 93% of young people report being online on a regular basis [6]. Furthermore, online interventions might be of particular interest for this age group due to the high levels of anonymity, easy access independent of time and location, cost-effectiveness for large populations, and the potential of such interventions to be perceived as less stigmatizing [7-10]. These advantages could potentially overcome some of the barriers young people face when seeking help for mental health problems.

Peer-to-peer support is a promising aspect of online mental health interventions. Prior research estimated that millions of people access online support groups daily [11]. Peer-to-peer support enables young people to connect with others, share experiences, seek and provide information, advice, and emotional support, and is often delivered as part of complex multi-component online interventions [8]. Research has shown that a majority of young people use the Internet to connect with others [12] suggesting that online peer-to-peer support could be a powerful tool to help reduce stigma and increase help-seeking for mental health problems.

A wide variety of online peer support platforms exist, including asynchronous (Internet support groups (ISGs)/discussion groups/bulletin boards/forums) and synchronous (chatrooms, virtual reality environments) formats. Available evidence suggests that online peer-to-peer support interventions might be beneficial for users [11]. Research has shown that users of online forums improve their coping strategies, both in social interaction and with regard to their health condition [13]. Peer support may also increase supportive communication [14] and emotional well-being [15]. Previous studies showed that greater participant involvement in an online forum was associated with lower levels of emotional distress among adolescents [16]. A meta-analysis of peer support interventions for depression found evidence that peer support leads to improvements in depressive symptoms relative to usual care [17]. A systematic review of Internet support groups for a wide variety of health conditions also found positive effects for depressive symptoms [18].

Although reviews have evaluated the effectiveness of face-to-face and online peer support networks in adults with a variety of health conditions, none have directly examined peer-to-peer support for mental health problems in young people. A recent systematic review examined online and social network interventions for depression in young people [19]. However, this review did not investigate other mental health conditions in young people. Other systematic reviews have investigated the effectiveness of health-related online peer support networks [11], and the impact of Internet support groups on depressive symptoms [18]. These reviews, however, did not focus on young people and mental health problems in general.

Therefore, it is unclear whether online peer-to-peer support is beneficial for young people and their mental health. The aim of the current review is to systematically evaluate the evidence regarding the effectiveness of online peer-to-peer support for young people with mental health problems.

Methods

Databases

PubMed, PsycInfo, and Cochrane databases were searched using keywords, phrases, and Medical Subject Headings (MeSH) terms in June 2014 (see Multimedia Appendix 1).

Search Methodology

The search strategy covered the following concepts: (1) technology, online communities, and methods of peer-to-peer interaction; (2) young people; and (3) mental health. Search terms regarding concept (1) were based on those used by Eysenbach et al [11], and Griffiths et al [18]. Search terms regarding concept (2) (young people) were developed by the researchers. Search terms regarding concept (3) (mental health) were based on the International Classification of Diseases (ICD-10) list of mental disorders and the National Health and Medical Research Council (NHMRC) keywords for mental health research [20]. The current systematic review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement [21]. A PRISMA checklist is available in Multimedia Appendix 2.

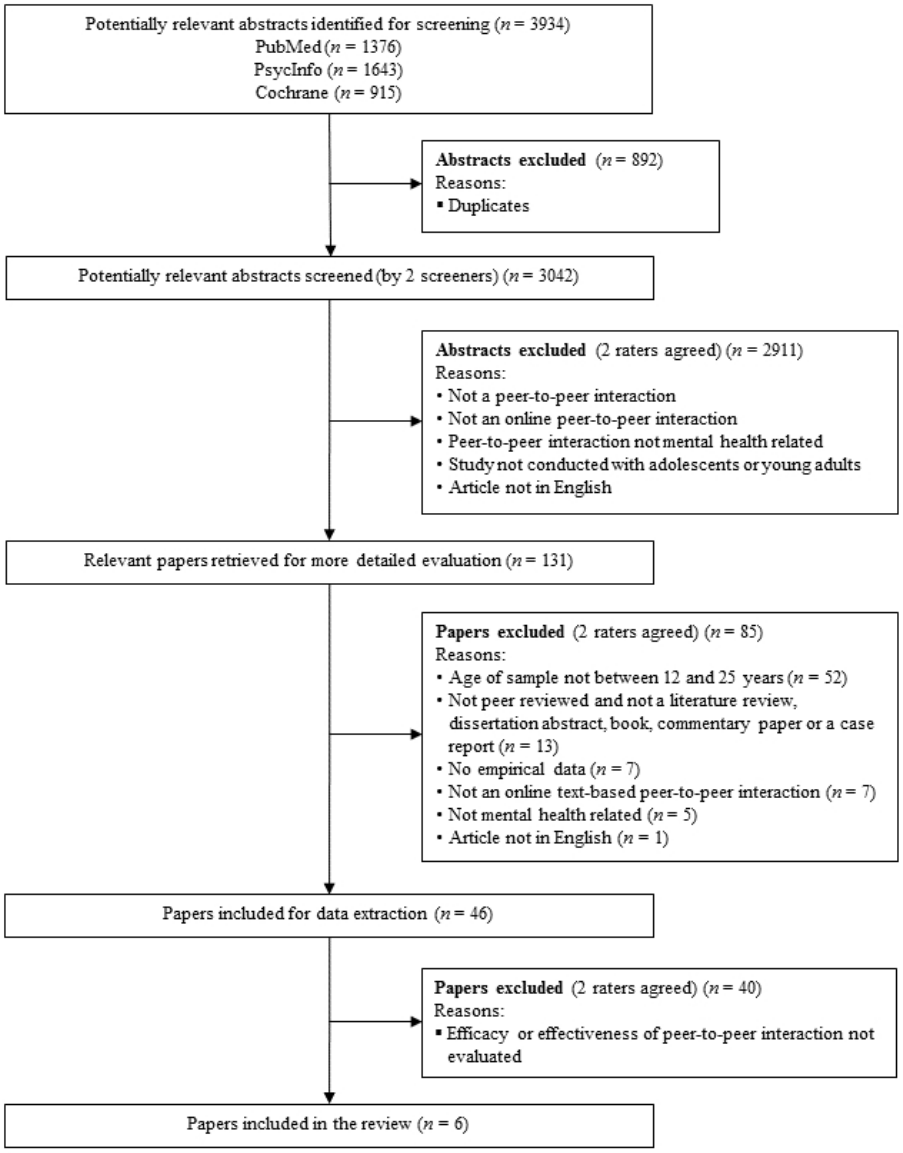
Study Identification

Overview

Figure 1 presents the flowchart for the selection of included studies. A process involving three screening stages was applied to select relevant studies for the present review. In total, the database searches yielded 3934 abstracts, of which 892 duplicates were removed.

XSL•FO
RenderX

Figure 1. Study identification flowchart.



Stage 1 Screening

For the first stage, in order to eliminate clearly irrelevant abstracts, two independent raters (KA and BC or RR) screened the remaining 3042 abstracts for relevant studies according to the following inclusion criteria (Textbox 1). This first screening stage yielded a total of 131 relevant papers.

Stage 2 Screening

In the second stage, the inclusion criteria were refined (Textbox 2) and the remaining 131 papers were screened according to these criteria by two independent raters (KA and LF or AG).

Textbox 1. Inclusion criteria for the fist stage.

1. The study discussed or investigated a peer-to-peer interaction.
2. The study discussed or investigated at least one of the following: online/electronic support groups, online/electronic social or peer support, online computer-based communication or interaction, collaborative virtual environments or interventions.
3. The peer-to-peer interaction focused on mental health or psychology related conditions (eg, mental illness, smoking, etc).
4. The study sample was composed of adolescents (12-17) or young adults (18-25).
5. The article was written in English.

XSL•FO
RenderX

Textbox 2. Inclusion criteria for the second stage.

1. Age: the mean age or the age range of the sample was between 12 to 25 years. If sample age was not specified in the paper, studies that related to students were included.
2. Peer-to-peer interaction: the peer-to-peer interaction was online and text-based. Studies that involved interactions in virtual reality environments that were not also text-based did not satisfy this criterion.
3. Peer-to-peer focus: the peer-to-peer interaction discussed or investigated mental health or psychology related conditions.
4. The study reported empirical data.
5. The study was peer-reviewed and not a literature review, dissertation abstract, book, commentary paper or a case report.
6. The article was in English.

Stage 3 Screening

A total of 46 papers were retained for the third screening stage and were included if they met the following criterion, rated by KA and LF: The study evaluated the effectiveness of peer-to-peer interaction, either as a stand-alone intervention or as a component of an intervention. Component intervention studies (eg, online cognitive behavioral therapy plus a support group) where the effectiveness of the peer support component of the intervention could not be isolated were excluded.

In addition, previous reviews, key journals, and reference lists of key papers were hand searched. However, no eligible studies were identified using this method. Finally, six papers were included for coding by two coders (KA and LF). At all stages of abstract screening, studies that raters mutually agreed on were retained and any disagreement was resolved by discussion.

Coding of Included Papers

The six papers evaluating the effectiveness of peer-to-peer interaction were coded independently by two raters with a preformulated coding sheet. Included studies were coded for (1) participant characteristics, (2) intervention design, (3) peer-to-peer support interaction, and (4) study design characteristics. Data coding of *participant characteristics* comprised the following: participant type, symptom level of recruited participants, sample size, age, and sex. Data coding regarding the *intervention design* and *peer-to-peer support interaction* included the following: intervention description, the format of the peer-to-peer interaction (Internet support groups/discussion groups/bulletin boards/forums, chatrooms, virtual reality), peer support type (public, research, other), and whether and by whom (consumers/health professionals/unknown) it was moderated. Finally, coding of *study design characteristics* included study design (randomized controlled trial [RCT]/randomized trial/pre-post), whether or not intention-to-treat (ITT) analysis was employed, dropout (n, %), the primary outcome measure for the study, measurement time points, whether or not the intervention yielded a statistically significant positive outcome and where possible Hedge's *g* effect sizes for differences between the intervention and the control group post-intervention.

In order to evaluate study quality, the risk of bias criteria proposed by the Cochrane Effective Practice and Organisation of Care Group (EPOC) was used [22]. These criteria are designed to assess potential sources of bias for studies involving a control group. The EPOC criteria assess the following nine study characteristics: random allocation sequence and allocation concealment, differences in baseline outcome measurements and characteristics, treatment of missing outcome data, researcher knowledge of allocated interventions, contamination between conditions, selective outcome reporting, and any other risk of bias.

Data Analysis

Given the small number and the heterogeneous nature of the studies, a quantitative meta-analysis was not undertaken. Study results pertaining to the primary outcome measure(s) of each study were reported. For the randomized controlled trials and the randomized trial, between group results were reported, including the group by time interaction. For pre-post studies, within group results were reported. Where studies reported mean scores and standard deviations for the primary outcome measures, between group Hedge's *g* corrections for small sample size [23] and confidence intervals were calculated for the randomized controlled trials and the randomized trial. One of the randomized controlled trials and both of the pre-post studies contained insufficient data to calculate effect sizes.

Results

Study Characteristics

Detailed characteristics of the included studies (n=6) are provided in Multimedia Appendix 3. Half of the studies were RCTs [24-26], two studies used a pre-post design [27,28] and one study was a randomized trial without a control group [29]. All three RCTs employed a no-intervention control group. Sample sizes ranged from 26 to 283 (median=95) across all studies. The studies were categorized according to the mental health topic of the online peer support network. The conditions targeted were depression and anxiety [26,27], general psychological problems [29], eating disorders [24], and substance use (tobacco) [25,28]. The symptom level at baseline included low to moderate levels of psychological distress, depressive symptoms, psychological problems, and regular smoking.

Origin

Half of the studies were conducted in the United States (n=3), and the remaining studies were from Australia, England, and Ireland.

XSL•FO
RenderX

Interventions

The majority of studies employed Internet support groups, bulletin boards, or forums (n=4), and the remaining studies focused on virtual reality chat (n=2). Most of the studies investigated peer-to-peer support platforms that were developed for research purposes and not available to the public. Five studies reported that support groups were moderated by either health professionals or consumers. The intervention length ranged from 3 to 10 weeks with a mean of 6.8 (SD 2.3) and the length of the longest follow-up ranged from 1-12 months post-intervention.

Participants

Most of the samples consisted of university students (n=4), and the remaining studies included rural teens or adolescent smokers. The range of mean age of participants in the sample fell between 15 to 21 years. Most of the studies targeted young adults (n=4) aged 18 to 25. In half of the studies (n=3), either all or a majority of the participants were female; one study contained equal numbers of males and females, one study contained predominantly males, and one study did not report participant gender. In most studies, participants were recruited at universities or schools (n=5).

Outcome Measures

The two studies targeting depression and anxiety used the Depression Anxiety Stress Scale (DASS) or the Center for Epidemiologic Studies Depression Scale (CES-D) as their primary outcome measure. The single study targeting psychological stress used the Clinical Outcomes in Routine Evaluation-Outcome Measure (CORE-OM). One study targeting eating disorders used the Eating Disorder Inventory (EDI), and the two remaining studies targeting substance use problems (tobacco) used past-week abstinence rates.

Study Quality

Of the six included studies, three studies reported completer analyses; two used an ITT design, and one study reported completer analyses for all outcomes and ITT analyses for some outcomes. Among the studies reporting ITT analyses, one study reported data from a full sample (no dropouts), one study estimated missing data using a generalized estimating equations approach, and one study used baseline scores to estimate missing data. Among all samples dropout ranged from to 0% to 86%.

Table 1 displays the ratings for the EPOC quality criteria for studies involving a control group [24-26,29]. None of the studies applied or indicated adequate randomization methods. Half of the studies indicated significant differences in the outcome measures across study groups at baseline. More than half of the studies reported that baseline characteristics of the study and control providers were similar. More than half of the studies reported that they used measures to adequately address incomplete data. None of the studies reported that the knowledge of the allocated interventions was adequately prevented during the study. In contrast, almost all studies reported that the study was adequately protected against contamination, that the study was free from selective outcome reporting and that the study was free from other risks of bias.

Table 1. Details of studies meeting quality rating criteria.

Study	Ellis 2011 [26]	Freeman 2008 [29]	Low 2006 [24]	Woodruff 2007 [25]
Was the allocation sequence adequately generated?				
Was the allocation adequately concealed?				
Were baseline outcome measurements similar?	✔	✔		
Were baseline characteristics similar?	✔	✔	✔	
Were incomplete outcome data adequately addressed?	✔		✔	✔
Was knowledge of the allocated interventions adequately prevented during the study?				
Was the study adequately protected against contamination?	✔	✔	✔	✔
Was the study free from selective outcome reporting?	✔	✔	✔	✔
Was the study free from other risks of bias?	✔	✔		✔

Intervention Efficacy

Depression and Anxiety Symptoms

Both the RCT and the pre-post study targeting depression and anxiety symptoms examined moderated forums [26,27]. The RCT demonstrated that an online peer support forum was effective compared with the control condition at post-intervention in reducing anxiety (g=-.91), but not depression (g=-.63) [26]. In the pre-post study, a reduction in depressive symptoms between pre and post-intervention was found among users of the forum [27]. However, this reduction was not statistically significant.

Psychological Problems

The single study targeting psychological problems was a randomized trial and compared an electronic bulletin board plus online information with online information alone [29]. The moderation status of the bulletin board was unknown. A significant reduction in depressive symptoms was observed in both intervention groups from pre- to post-intervention. There was no evidence for an additional effect of the electronic support group post-intervention in reducing depressive symptoms (g=-.22).

XSL•FO
RenderX

Eating Disorders

The single RCT targeting eating disorders compared a no intervention control condition with three conditions, comprising the Student Bodies program plus either (1) moderated peer support, (2) unmoderated peer support, or (3) no peer support [24]. No significant differences were found between the three groups, post-intervention and at follow-up, indicating that no additional effect of peer-to-peer support to the Student Bodies program was found. Between group effect sizes for Bulimia, Body Dissatisfaction and Drive for Thinness subscales ranged from -1.05 to 0.98.

Smoking

The RCT and the pre-post study targeting tobacco use, both examined virtual world chat rooms [25,28]. While in both studies participants were smokers at baseline, the RCT reported significantly higher abstinence rates in the intervention group compared to the control group post intervention [25]. The pre-post study showed an increase in smoking abstinence (11%) in the past week from pre- to post-intervention, although this change was not significant [28].

Discussion

Principal Findings

This systematic review identified six studies (3 RCTs, 2 pre-post, 1 randomized trial) examining the effectiveness of online peer-to-peer support for young people with mental health problems. The studies targeted a range of mental health issues, including depression and anxiety, general psychological problems, eating disorders and substance use (tobacco). Overall, two of the 4 RCTs/randomized trials yielded a positive effect for the peer-support group relative to the comparison group at post-intervention: the RCT targeting anxiety and the RCT targeting tobacco [25,26]. There was no evidence that peer-to-peer support was effective for eating disorder or depressive symptoms [24,26]. However, the study targeting depressive symptoms might have been underpowered given the magnitude of the effect size. Of the two trials that yielded positive effects, one used a moderated discussion group [26] and the other utilized virtual reality chat [25]. In general, studies were of low quality, scoring 4.6 out of 9 on average.

Thus, the current review found some evidence for the efficacy of peer-to-peer support alone or as an adjunct to other treatment programs for mental health problems in young people. These findings are similar to a previous systematic review targeting online and social networking interventions in young people for depression [19], which reported positive outcomes for some of the social networking studies. The results are also consistent with findings from two studies of ISGs that have targeted adults with depression [30,31]. It is encouraging in the present review that two of the interventions tested in randomized controlled trials were shown to be effective in comparison to the control groups. Although possibly limited in their generalizability due to limited sample sizes, high numbers of female participants, large variations in dropout rates and specific mental health problems (anxiety and tobacco) these findings are of interest, particularly given the extensive use of peer support in the field of mental health [16].

The majority of discussion groups were moderated by health professionals, researchers or consumers. However, there was limited information on moderators, their level of skills, and their engagement with the discussion group. In addition, although the type of moderation might impact the outcome for participants, included studies did not provide detailed information on the use of theory driven moderation or discussion of risk management. Given that moderation is a critical aspect of peer-to-peer support, future studies should include details about the type of moderation and moderators. This may shed light on which level of moderation works best for participants.

The paucity of high-quality randomized controlled trials examining the effects of peer support in young people makes it difficult to draw firm conclusions about its effectiveness. These findings are similar to previous systematic reviews of online peer-to-peer support in adults. For example, one review found that all identified studies that evaluated the additional effect of peer support were pre-post studies with low-quality research designs [11]. Another review that focused on Internet support groups and depression in adults found similar results, emphasizing the need for high-quality research in this field [18]. The lack of high-quality studies is especially concerning, given young people's extensive use of the Internet to search for mental health information and connect with others [12].

From this review it is clear that research on peer-to-peer support among young people has been dominated by studies that employed asynchronous communication. In this review only two studies investigated synchronous communication; both involved a virtual reality chat component targeting tobacco. There is a need to investigate the effectiveness of synchronous chat sessions for other mental health problems in young people.

It is also important to note that although peer support is frequently used as an adjunct to online interventions, very few studies have isolated and investigated the additional effect of online peer support. The present review addressed this specific gap in the literature. This resulted in the exclusion of 40 studies. It is disappointing that so few studies sought to identify the specific contribution of peer-to-peer support. Opinions on this approach are mixed and the suitability and practicability of peer-to-peer support as an intervention in and of itself has been the subject of debate [11,32]. Qualitative studies of online peer support contribute to understanding user characteristics, perceived benefits, potential risks, and the self-help process [32]. Yet, the question remains under which conditions and for whom these support groups are effective and how social support can be improved [11]. Future studies should use both qualitative and quantitative methods to investigate this question.

The field of online peer-to-peer support is still in its infancy and many questions remain unanswered. Despite this fact, given that many online interventions for young people include a peer support component, there is an urgent need to distinguish which parts of these interventions are effective.

XSL•FO
RenderX

Limitations

There are several limitations to this systematic review. Studies were identified based on searches in three databases. It is possible that this search strategy failed to identify some eligible studies. To address this in part, previous reviews, key journals, and key papers were hand searched. Despite extensive search, no additional eligible studies were identified using this method. A further level of bias might be due to the criterion that only English language papers were included in the present review. Finally, the review may be subject to publication bias if authors failed to publish some studies with null findings.

Implications for Future Research

High-quality research on online peer-to-peer support for young people is currently lacking. Many studies examining Internet interventions for young people use peer support as an adjunct. However, there is limited evidence for the effectiveness of this addition to care. It is vital that future research explores the specific contribution of peer support to these interventions. Based on the findings published to date, it is not possible to conclude for whom and under which conditions peer support interactions work. In addition, the absence of studies involving individuals experiencing severe problems or in recovery shows it is not known if peer-to-peer support is appropriate for this group.

In summary, given that a majority of young people are using the Internet routinely, further research is needed to explore the role that peer-to-peer support might play in assisting young people with mental health problems.

Acknowledgments

This project was resourced by the Young and Well CRC. The Young and Well CRC is established under the Australian Government's Cooperative Research Centres Program. The authors would like to acknowledge Bradley Carron-Arthur and Rebecca Randall for screening the abstracts in the initial stage of the study.

Conflicts of Interest

KG is a co-developer of MoodGYM, which was evaluated as part of one [26] of the included trials.

Multimedia Appendix 1

Search terms/search history.

[PDF File (Adobe PDF File), 48KB - mental_v2i2e19_app1.pdf]

Multimedia Appendix 2

PRISMA checklist.

[PDF File (Adobe PDF File), 149KB - mental_v2i2e19_app2.pdf]

Multimedia Appendix 3

Full data from the review.

[PDF File (Adobe PDF File), 80KB - mental_v2i2e19_app3.pdf]

References

1. Patel V, Flisher AJ, Hetrick S, McGorry P. Mental health of young people: a global public-health challenge. Lancet 2007 Apr 14;369(9569):1302-1313. [doi: 10.1016/S0140-6736(07)60368-7] [Medline: 17434406]
2. Australian Institute of Health and Welfare. Young Australians: their health and wellbeing. 2011. 2011 URL: http://www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=10737419259 [accessed 2015-03-10] [WebCite Cache ID 6WuvqEl8k]
3. Booth ML, Bernard D, Quine S, Kang MS, Usherwood T, Alperstein G, et al. Access to health care among Australian adolescents young people's perspectives and their sociodemographic distribution. J Adolesc Health 2004 Jan;34(1):97-103. [Medline: 14706412]
4. Rickwood D, Deane FP, Wilson CJ, Ciarrochi J. Young people's help-seeking for mental health problems. Australian e-Journal for the Advancement of Mental Health 2005;4:1-34 [FREE Full text]
5. Wetterlin FM, Mar MY, Neilson EK, Werker GR, Krausz M. eMental health experiences and expectations: a survey of youths' Web-based resource preferences in Canada. J Med Internet Res 2014;16(12):e293 [FREE Full text] [doi: 10.2196/jmir.3526] [Medline: 25519847]
6. Lenhart A, Purcell K, Smith A, Zickuhr K. Social media & mobile internet use among teens and young adults. 2010 URL: http://www.pewinternet.org/files/old-media/Files/Reports/2010/PIP_Social_Media_and_Young_Adults_Report_Final_with_toplines.pdf [accessed 2015-03-10] [WebCite Cache ID 6WuwAlBfp]

XSL•FO
RenderX

7. Bauer S, Moessner M. Harnessing the power of technology for the treatment and prevention of eating disorders. Int J Eat Disord 2013 Jul;46(5):508-515. [doi: 10.1002/eat.22109] [Medline: 23658102]
8. Campbell AJ, Robards F. Using technologies safely and effectively to promote young people's wellbeing: a better practice guide for services. 2013 URL: http://www.youngandwellcrc.org.au/wp-content/uploads/2014/03/Guide_Using_Technologies_Safely_and_Effectively_Young_and_Well_CRC_Mar2013.pdf [accessed 2015-03-10] [WebCite Cache ID 6WuwKH8td]
9. Griffiths K, Christensen H. Internet-based mental health programs: A powerful tool in the rural medical kit. Aust J Rural Health 2007;15:81-87. [Medline: 17441815]
10. King R, Bickman L, Shochet I, McDermott B, Bor B. Use of the internet for provision of better counselling and psychotherapy services to young people, their families and carers. Psychotherapy in Australia 2010;17(1):66-74.
11. Eysenbach G, Powell J, Englesakis M, Rizo C, Stern A. Health related virtual communities and electronic support groups: systematic review of the effects of online peer to peer interactions. BMJ 2004 May 15;328(7449):1166 [FREE Full text] [doi: 10.1136/bmj.328.7449.1166] [Medline: 15142921]
12. Burns JM, Davenport TA, Durkin LA, Luscombe GM, Hickie IB. The internet as a setting for mental health service utilisation by young people. Med J Aust 2010 Jun 7;192(11 Suppl):S22-S26. [Medline: 20528703]
13. Tanis M. Health-related on-line forums: what's the big attraction? J Health Commun 2008;13(7):698-714. [doi: 10.1080/10810730802415316] [Medline: 18958781]
14. Wright KB, Bell SB, Wright KB, Bell SB. Health-related Support Groups on the Internet: Linking Empirical Findings to Social Support and Computer-mediated Communication Theory. J Health Psychol 2003 Jan;8(1):39-54. [doi: 10.1177/1359105303008001429] [Medline: 22113899]
15. Batenburg A, Das E. Emotional approach coping and the effects of online peer-led support group participation among patients with breast cancer: a longitudinal study. J Med Internet Res 2014;16(11):e256 [FREE Full text] [doi: 10.2196/jmir.3517] [Medline: 25474819]
16. Barak A, Dolev-Cohen M. Does activity level in online support groups for distressed adolescents determine emotional relief. Counselling and Psychotherapy Research 2006 Sep;6(3):186-190. [doi: 10.1080/14733140600848203]
17. Pfeiffer PN, Heisler M, Piette JD, Rogers Mary A M, Valenstein M. Efficacy of peer support interventions for depression: a meta-analysis. Gen Hosp Psychiatry 2011;33(1):29-36 [FREE Full text] [doi: 10.1016/j.genhosppsych.2010.10.002] [Medline: 21353125]
18. Griffiths KM, Calear AL, Banfield M. Systematic review on Internet Support Groups (ISGs) and depression (1): Do ISGs reduce depressive symptoms? J Med Internet Res 2009;11(3):e40 [FREE Full text] [doi: 10.2196/jmir.1270] [Medline: 19793719]
19. Rice SM, Goodall J, Hetrick SE, Parker AG, Gilbertson T, Amminger GP, et al. Online and social networking interventions for the treatment of depression in young people: a systematic review. J Med Internet Res 2014;16(9):e206 [FREE Full text] [doi: 10.2196/jmir.3304] [Medline: 25226790]
20. National Health and Research Council (NHMRC). Research Keywords/Phrases URL: https://www.nhmrc.gov.au/_files_nhmrc/file/grants/policy/rgms_nhmrc_research_keywords_110923.pdf [accessed 2015-04-12] [WebCite Cache ID 6Xj285z1p]
21. Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. Ann Intern Med 2009 Aug 18;151(4):264-9, W64. [Medline: 19622511]
22. Effective Practice Organisation of Care (EPOC). Suggested risk of bias criteria for EPOC reviews. 2015 URL: http://epoc.cochrane.org/sites/epoc.cochrane.org/files/uploads/14%20Suggested%20risk%20of%20bias%20criteria%20for%20EPOC%20reviews%2009%2002%2015_0.pdf [accessed 2015-03-10] [WebCite Cache ID 6WuwffsPH]
23. Borenstein M. Effect sizes for continuous data. In: Cooper H, Hedges LV, Valentine JC, editors. The handbook of research synthesis and meta-analysis. New York: Russell Sage Foundation; 2009:221-236.
24. Low KG, Charanasomboon S, Lesser J, Reinhalter K, Martin R, Jones H, et al. Effectiveness of a computer-based interactive eating disorders prevention program at long-term follow-up. Eat Disord 2006;14(1):17-30. [doi: 10.1080/10640260500403816] [Medline: 16757446]
25. Woodruff SI, Conway TL, Edwards CC, Elliott SP, Crittenden J. Evaluation of an Internet virtual world chat room for adolescent smoking cessation. Addict Behav 2007 Sep;32(9):1769-1786. [doi: 10.1016/j.addbeh.2006.12.008] [Medline: 17250972]
26. Ellis LA, Campbell AJ, Sethi S, O'Dea BM. Comparative randomized trial of an online cognitive-behavioral therapy program and an online support group for depression and anxiety. J Cyber Ther Rehabil 2011;4(4):461-467.
27. Horgan A, McCarthy G, Sweeney J. An evaluation of an online peer support forum for university students with depressive symptoms. Arch Psychiatr Nurs 2013 Apr;27(2):84-89. [doi: 10.1016/j.apnu.2012.12.005] [Medline: 23540518]
28. Woodruff SI, Edwards CC, Conway TL, Elliott SP. Pilot test of an Internet virtual world chat room for rural teen smokers. J Adolesc Health 2001 Oct;29(4):239-243. [Medline: 11587907]
29. Freeman E, Barker C, Pistrang N. Outcome of an online mutual support group for college students with psychological problems. Cyberpsychol Behav 2008;11:591-593. [Medline: 18817485]

XSL•FO
RenderX

30. Houston TK, Cooper LA, Ford DE. Internet support groups for depression: a 1-year prospective cohort study. Am J Psychiatry 2002 Dec;159(12):2062-2068. [Medline: 12450957]
31. Griffiths KM, Mackinnon AJ, Crisp DA, Christensen H, Bennett K, Farrer L. The effectiveness of an online support group for members of the community with depression: a randomised controlled trial. PLoS One 2012;7(12):e53244 [FREE Full text] [doi: 10.1371/journal.pone.0053244] [Medline: 23285271]
32. Barak A, Grohol JM, Pector E. BMJ. Methodology, validity, and applicability: a critique on Eysenbach URL: http://www.bmj.com/rapid-response/2011/10/30/methodology-validity-and-applicability-critique-eysenbach-et-al [accessed 2015-05-10] [WebCite Cache ID 6YQR26cBg]

Abbreviations

CES-D: Center for Epidemiologic Studies Depression Scale
CORE-OM: Clinical Outcomes in Routine Evaluation-Outcome Measure
DASS: Depression Anxiety Stress Scale
EDI: Eating Disorder Inventory
EPOC: Effective Practice and Organisation of Care Group
ICD-10: International Classification of Diseases
ISGs: Internet support groups
ITT: Intention-to-Treat
MeSH: Medical Subject Headings
NHMRC: National Health and Medical Research Council
PRSIMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT: randomized controlled trial

Edited by G Eysenbach; submitted 11.03.15; peer-reviewed by S Rice, N Thomas; comments to author 01.04.15; revised version received 16.04.15; accepted 17.04.15; published 19.05.15.

Please cite as:
Ali K, Farrer L, Gulliver A, Griffiths KM
Online Peer-to-Peer Support for Young People With Mental Health Problems: A Systematic Review
JMIR Mental Health 2015;2(2):e19
URL: http://mental.jmir.org/2015/2/e19/
doi:10.2196/mental.4418
PMID:26543923

©Kathina Ali, Louise Farrer, Amelia Gulliver, Kathleen M Griffiths. Originally published in JMIR Mental Health (http://mental.jmir.org), 19.05.2015. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Short Paper

Preferences of Young Adults With First-Episode Psychosis for Receiving Specialized Mental Health Services Using Technology: A Survey Study

Shalini Lal[1], BScOT, MSc, PhD; Jennifer Dell'Elce[2], BA; Natasha Tucci[3], BA; Rebecca Fuhrer[4], PhD; Robyn Tamblyn[4], BScN, MSc, PhD; Ashok Malla[5], MD, FRCPC

[1]School of Rehabilitation, University of Montreal, Montreal, QC, Canada

[2]School of Physical and Occupational Therapy, McGill University, Montreal, QC, Canada

[3]Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, United Kingdom

[4]McGill University, Department of Epidemiology, Biostatistics, and Occupational Health, Montreal, QC, Canada

[5]Department of Psychiatry, McGill University, Montreal, QC, Canada

Corresponding Author:
Shalini Lal, BScOT, MSc, PhD
School of Rehabilitation
University of Montreal
CP 6128, succursale Centre-ville
Montreal, QC, H3C 3J7
Canada
Phone: 1 514 761 6131 ext 2426
Fax: 1 514 888 4064
Email: shalini.lal@umontreal.ca

Abstract

Background: Despite the potential and interest of using technology for delivering specialized psychiatric services to young adults, surprisingly limited attention has been paid to systematically assess their perspectives in this regard. For example, limited knowledge exists on the extent to which young people receiving specialized services for a first-episode psychosis (FEP) are receptive to using new technologies as part of mental health care, and to which types of technology-enabled mental health interventions they are amenable to.

Objective: The purpose of this study is to assess the interest of young adults with FEP in using technology to receive mental health information, services, and supports.

Methods: This study uses a cross-sectional, descriptive survey design. A convenience sample of 67 participants between the ages of 18 and 35 were recruited from two specialized early intervention programs for psychosis. Interviewer-administered surveys were conducted between December 2013 and October 2014. Descriptive statistics are reported.

Results: Among the 67 respondents who completed the survey, the majority (85%, 57/67) agreed or strongly agreed with YouTube as a platform for mental health-related services and supports. The top five technology-enabled services that participants were amenable to were (1) information on medication (96%, 64/67); (2) information on education, career, and employment (93%, 62/67); (3) decision-making tools pertaining to treatment and recovery (93%, 62/67); (4) reminders for appointments via text messaging (93%, 62/67); and (5) information about mental health, psychosis, and recovery in general (91%, 61/67). The top self-reported barriers to seeking mental health information online were lack of knowledge on how to perform an Internet search (31%, 21/67) and the way information is presented online (27%, 18/67). Two thirds (67%; 45/67) reported being comfortable in online settings, and almost half (48%; 32/67) reported a preference for mixed formats when viewing mental health information online (eg, text, video, visual graphics).

Conclusions: Young people diagnosed with FEP express interest in using the Internet, social media, and mobile technologies for receiving mental health-related services. Increasing the awareness of young people in relation to various forms of technology-enabled mental health care warrants further attention. A consideration for future research is to obtain more in-depth knowledge on young people's perspectives, which can help improve the design, development, and implementation of integrated technological health innovations within the delivery of specialized mental health care.

XSL•FO
RenderX

(JMIR Mental Health 2015;2(2):e18) doi:10.2196/mental.4400

KEYWORDS

eHealth; health services accessibility; medical informatics; mHealth; patient engagement; telemedicine

Introduction

Internet and mobile technologies are increasingly being considered as a promising avenue to improve access and quality of mental health services [1-5]. This is particularly true for young people, given the omnipresence of online technologies in their daily lives [6-8]. Technology can be used to support and complement many areas of mental health service delivery including providing information, conducting screenings and assessments, monitoring symptoms and behaviors, delivering psychosocial interventions, and providing peer support [1]. Technology-enabled services offer a less intensive and arguably a more engaging format that is commensurate with the developmental culture of young people growing up in the 21st century [9]. However, there is a significant gap between the role that technologies such as the Internet, social media, and mobile devices play in young people's lives and the role that these technologies have in the delivery of specialized psychiatric services to young people. For example, specialized early intervention (SEI) services for young people diagnosed with a first-episode psychosis (FEP) are predominantly based on models of care that are delivered in person.

As specialized psychiatric services begin to consider various forms of technology to augment, complement, or extend the reach of care for young people, it is important to obtain their interests and perspectives in this regard. For example, information is needed on young people's preferences for receiving technology-enabled mental health care, and which types of services they are amenable to (eg, online peer support, text reminders for medication and appointments, online counseling). This knowledge can guide the development of Internet-based and mobile innovations that are engaging, useful, and patient centered. As such, the aim of this paper is to describe the methods and results of a survey we conducted to assess the preferences of young adults with FEP on using technology for receiving mental health services, information, and supports.

Methods

Recruitment

Using a cross-sectional, descriptive survey design, a total of 67 participants were recruited from two SEI programs for psychosis: Prevention and Early Intervention Program for Psychoses-Montréal and Prevention and Early Intervention Program for Psychoses-McGill University Health Centre. Before being admitted to these services, the majority of patients received less than 1 month of medication for their FEP. Ethical approval was received from the Institutional Review Board of McGill University's Faculty of Medicine, and all participants provided informed consent.

Participants were eligible if they met the following criteria: between 18 and 35 years of age, diagnosed with an affective or nonaffective psychotic disorder, within 5 years of treatment, clinically stable, and able to communicate in English or French. Participants were excluded if they had no contact with clinical staff for over 3 months, were hospitalized at the time of recruitment, were unable to concentrate or attend to a conversation, or were intoxicated at the time of recruitment. Recruitment occurred between December 2013 and October 2014.

Measures

Data collection involved an interviewer-administered survey (20-30 minutes in duration), developed iteratively over a 3-month period with feedback from young adult patients and family representatives, service providers, and the research team. The survey consisted of quantitative questions to assess the following data: (1) demographics; (2) access and use of technology and social media; and (3) preferences regarding the use of technology for various types of mental health services, information, and supports. In this paper, items 1 and 3 were included. Examples of questions from the survey are provided in Multimedia Appendix 1. Responses were numerically coded in an SPSS database (version 20; SPSS Inc, Chicago, IL, USA) and descriptive statistics (ie, mean, standard deviation, frequencies) are reported in the following sections.

Results

Participants

Of the 76 individuals approached, 67 provided consent and completed the survey. The mean age of the participants was 25.6 (standard deviation = 5.1), and 76% were men (51/67). Table 1 presents the demographic details of the participants.

XSL•FO
RenderX

Table 1. Demographic characteristics (n=67).

Characteristics		n	%
Sex (male/female)			
	Male	51	76.1
Language of survey (English/French)			
	English	33	49.3
Length of treatment			
	>12 months	39	58.2
	6-12 months	17	25.4
	<6 months	11	16.4
Race [a]			
	White	43	64.2
	Asian	10	14.9
	Black	5	7.5
	Latin American	5	7.5
	Arab	1	1.5
	Mixed	3	4.5
Education			
	High school complete	23	34.3
	High school incomplete	16	23.9
	College complete	15	22.4
	Undergraduate studies complete	10	14.9
	Graduate studies complete	3	4.5
Employment status			
	Not employed/looking for work	42[b]	62.7
	Full time	17	25.4
	Part-time	8	11.9
Education status			
	Not in school/not student	49[c]	73.1
	Part-time	10	14.9
	Full time	8	11.9
Living situation			
	With family/guardian	42	62.7
	Alone	10	14.9
	With roommates	7	10.4
	With spouse	6	9.0
	Supported housing	2	3.0
Annual income (CAD$)			
	<14,999	49	73.1
	15,000-29,999	7	10.4
	30,000-49,999	9	13.4
	>50,000	2	3.0
Children (Yes/No)			

XSL•FO
RenderX

Characteristics		n	%
	No	60	89.6

[a]Race categories were adapted from Statistics Canada race breakdown.

[b]Eight of the 42 are full-time students and four others are on medical leave.

[c]Two of the 49 are on medical leave from school.

Social Media and Mental Health Services

In descending order, participants agreed or strongly agreed with the idea of using the following social media sites as platforms to receive mental health information and supports: YouTube (85%, 57/67), Facebook (58%, 39/67), Skype (40%, 27/67), and Twitter (39%, 26/67).

Technology and Various Types of Mental Health-Related Services

As illustrated in Table 2, in descending order, the participants were amenable to using technology for the following top 10 types of services:(1) information on medication and side effects; (2) information and support related to education, career, and employment; (3) decision-making tools regarding treatment and recovery; (4) reminders for appointments through text messaging; (5) information about mental health, psychosis, and recovery in general; (6) information about physical health; (7) contact with mental health care providers; (8) scheduling appointments; (9) information about program events; and (10) education on coping skills. More than half (66%, 44/67) agreed or strongly agreed with using technology to facilitate social contact with other young people receiving services for FEP, and a little more than half of the participants (52%, 35/67) agreed or strongly agreed with receiving counseling or therapy online.

Table 2. Types of technology-enabled services (n=67).

Service	Strongly agree/agree n (%)	Strongly disagree/disagree n (%)	Undecided n (%)
Reminders for appointments			
By text message	62 (92.5)	3 (4.5)	2 (3.0)
By email	53 (79.1)	9 (13.4)	5 (7.5)
By applications	49 (73.1)	9 (13.4)	9 (13.4)
Reminders for medication			
By text message	45 (67.1)	16 (23.9)	6 (9.0)
By email	38 (56.7)	22 (32.8)	7 (10.4)
By applications	42 (62.7)	15 (22.4)	10 (14.9)
Information on medications and side effects	64 (95.5)	1 (1.5)	2 (3.0)
Information/support related to education, career, and employment	62 (92.5)	3 (4.5)	2 (3.0)
Tools to enable decision making regarding treatment and recovery	62 (92.5)	3 (4.5)	2 (3.0)
Information on mental health, psychosis, and recovery	61 (91.0)	3 (4.5)	3 (4.5)
Information on physical health	61 (91.0)	4 (6.0)	2 (3.0)
Contact with treatment team	57 (85.1)	6 (9.0)	4 (6.0)
Appointment scheduling	56 (83.6)	7 (10.4)	4 (6.0)
Information on program events	56 (83.6)	7 (10.4)	4 (6.0)
Education on coping skills	56 (83.6)	7 (10.4)	4 (6.0)
Online social contact between clients in the program and clients from similar programs	44 (65.7)	13 (19.4)	10 (14.9)
Counseling/therapy services	35 (52.2)	18 (26.9)	14 (20.9)

Barriers to Accessing Mental Health Support and Information Online

Table 3 describes the barriers that participants (n=67) reported in accessing mental health information and support online. In descending order, the top five barriers were (1) lack of knowledge on how to perform an Internet search (31%, 21/67), (2) the way in which information is presented online (27%, 18/67), (3) no interest or need (22%, 15/67), (4) lack of time (19%, 13/67), and (5) cost of Internet access (19%, 13/67). A little less than a third (30%; 20/67) reported “no barriers” to accessing mental health information and support online.

XSL•FO
RenderX

Table 3. Barriers to accessing mental health information and support online (n=67).

Barriers	n	%
Lack of knowledge on how to search Internet	21	31.3
No barriers	20	29.9
Way the information is presented online	18	26.9
Not interested/no need	15	22.4
Lack of time	13	19.4
Cost of Internet access	13	19.4
Lack of skills using technological devices	12	17.9
Fear or discomfort with technological devices	11	16.4
Lack of access to a device	9	13.4
Lack of access to Internet	5	7.5
Due to disability	4	6.0
Other[a]	2	3.0

[a]Other includes too much information, and knowing what information is good and bad.

Preferred Formats for Receiving Mental Health Information Online

Table 4 describes preferences for receiving mental health information online. Almost half of the participants (48%, 32/67) preferred mixed formats including a combination of text, video, graphics, and audio. Almost one third (31%, 21/67) preferred text as the method for receiving information, and 15% (10/67) preferred video (see Figure 1 for screenshot).

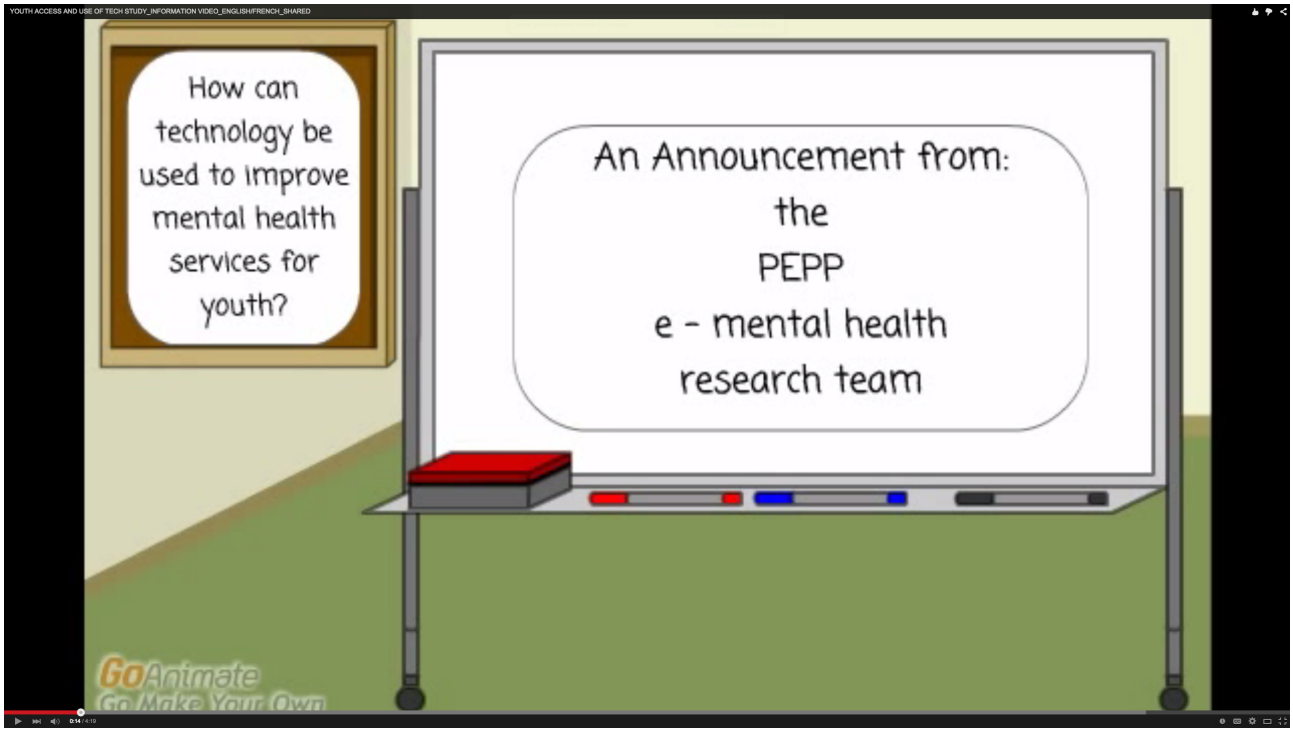
Table 4. Preferred formats for mental health information online (n=67).

Format	n	%
Mixed formats	32	47.8
Text	21	31.3
Video	10	14.9
Graphics	2	3.0
Audio	1	1.5
Other[a]	1	1.5

[a]The response indicated “no preference.”

XSL•FO
RenderX

Figure 1. Screenshot from information video about the study.



Comfort Levels in Different Online and Offline Social Environments

When asked about the level of comfort in different types of social environments, the following portion of participants reported feeling comfortable or very comfortable in online settings, over the phone, and in-person group settings: 67% (45/67), 66% (44/67), and 67% (45/67), respectively, whereas 81% of the sample (54/67) expressed being comfortable or very comfortable with one-to-one, in-person interactions.

Discussion

Principal Findings

To our knowledge, this study is the first to systematically assess preferences of young people with FEP for using various technologies to receive mental health care. We found that the majority of participants strongly agreed with using social media such as YouTube and Facebook as a platform for mental health-related services and supports. The majority of the participants also agreed with the idea of using technology for receiving information on topics such as medication, education, career, employment, mental health, psychosis, and recovery; as decision-making tools; and as reminders for appointments. Lack of knowledge on how to perform an Internet search and the way information is presented online were reported as barriers to seeking mental health information. Almost half of the participants reported a preference for mixed formats when viewing mental health information online (eg, text, video, visual graphics). It is also noteworthy that one third of participants reported lack of comfort in online settings.

YouTube and Facebook are the two most popular social media sites with young people in North America [10,11]. Social media offer young people a forum for asking questions, receiving feedback, sharing personal stories, and receiving information about clinical services and related events. However, it is important to note that there are many questions in terms of governance, ethics, professionalism, privacy, and confidentiality pertaining to the use of social media in the health care sector that remain unanswered [12,13]. Future research that includes the subjective perspectives of young people on these factors is warranted.

Given the extent to which young people are increasingly socialized to technology at a very young age, it is surprising that approximately one third of the sample reported lack of comfort in online environments. This may be related to privacy concerns, symptoms, or lack of skill level and confidence with technology; however, further research is needed to understand the factors contributing to this lack of comfort. Such knowledge can provide insight into the need for raising the literacy, confidence, and skills of young people in accessing mental health information, services, and supports online.

About one third disagreed, strongly disagreed, or were undecided regarding receiving reminders for taking medication through texting or email, which may be related to negative attitudes and perceptions about medication. Moreover, approximately half of the sample agreed with using technology to connect with peers or receive counseling. These findings can be better understood through qualitative research methods.

Limitations

The findings of this study are based on young adults between 20 and 30 years of age receiving services for an FEP in Quebec, Canada, and may not be generalizable to youth elsewhere. The majority of the sample population was men and although FEP incidence rates are approximately two times as high in male participants compared with female participants [14], our sample is still gender biased. It is important for future research to

XSL•FO
RenderX

consider gender-related preferences in relation to receiving technology-enabled mental health care. The findings are also limited by the fact that the study is based on self-reported data. Given the small sample size, the results should be interpreted with caution. Future research with a larger sample that can assess how sociodemographic factors (eg, gender, economic status, ethnicity, and education) influence preferences in relation to using technology for various types of services and supports is warranted.

Conclusions

Our study provides preliminary indications that young adults with FEP are interested in using the Internet, social media, and mobile technologies for receiving specialized services. Young people with FEP are in an ideal position to contribute to how mental health information, services, and supports can best be translated into online formats. These results will help inform the development of a website tailored to the needs of young people receiving services for an FEP at the recruitment sites. Further research, through qualitative methods, is warranted on patient perspectives regarding how social media and other types of technologies can be used to enhance service delivery and the issues of confidentiality, security, ethics, and professionalism in this regard.

Acknowledgments

We would like to acknowledge (1) all respondents for their participation in the study; (2) Marilyn Henderson for research assistance with survey development, recruitment, and data collection; (3) all individuals who provided feedback on the survey during its development including clinicians, students, patients, and family representatives. During completion of this work, the first author was partially supported by postdoctoral fellowships from the Canadian Institutes of Health Research and Knowledge Translation Canada, and a postdoctoral stipend from the Canadian Institutes of Health Research Interdisciplinary Capacity Enhancement Grant-in-Health Disparities (HOA-80072). The last author is supported by the Canada Research Chair Program.

Authors' Contributions

SL conceived and designed the study, led the development, implementation, analysis, and reporting of the survey results, and drafted the paper. JD and NT assisted in data collection, management, analysis, reporting of results, and review of the manuscript. RF, RT, and AM contributed substantially to the interpretation of the data and critical review and revision of the manuscript.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Examples of survey questions.

[PDF File (Adobe PDF File), 254KB - mental_v2i2e18_app1.pdf]

References

1. Lal S, Adair CE. E-mental health: A rapid review of the literature. Psychiatr Serv 2014 Jan 1;65(1):24-32. [doi: 10.1176/appi.ps.201300009] [Medline: 24081188]
2. Álvarez-Jiménez M, Gleeson JF, Bendall S, Lederman R, Wadley G, Killackey E, et al. Internet-based interventions for psychosis: A sneak-peek into the future. Psychiatr Clin North Am 2012 Sep;35(3):735-747. [doi: 10.1016/j.psc.2012.06.011] [Medline: 22929876]
3. Ben-Zeev D. Mobile technologies in the study, assessment, and treatment of schizophrenia. Schizophr Bull 2012 May;38(3):384-385 [FREE Full text] [doi: 10.1093/schbul/sbr179] [Medline: 22234487]
4. Burns J, Birrell E. Enhancing early engagement with mental health services by young people. Psychol Res Behav Manag 2014;7:303-312 [FREE Full text] [doi: 10.2147/PRBM.S49151] [Medline: 25473320]
5. Wetterlin FM, Mar MY, Neilson EK, Werker GR, Krausz M. eMental health experiences and expectations: A survey of youths' Web-based resource preferences in Canada. J Med Internet Res 2014;16(12):e293 [FREE Full text] [doi: 10.2196/jmir.3526] [Medline: 25519847]
6. Pew Research Center. Social Media Use by Age Group Over Time. Washington, DC: Pew Research Center; 2011. URL: http://www.pewinternet.org/data-trend/social-media/social-media-use-by-age-group/ [accessed 2015-03-05] [WebCite Cache ID 6Wo7C1Mpl]
7. Purcell K. 10 things to know about how teens use technology. Washington, DC: Pew Research Center; 2013. URL: http://www.pewinternet.org/2013/07/10/10-things-to-know-about-how-teens-use-technology/ [accessed 2015-01-28] [WebCite Cache ID 6VvTvNaBx]
8. Mental Health Commission of Canada. A foundation for online knowledge mobilization in child and youth mental health. Alberta: Mental Health Commission of Canada; 2012. URL: http://www.mentalhealthcommission.ca/sites/default/files/

XSL•FO
RenderX

filefield_private_download/document/cy_online_knowledge_mobilization_final_report_eng.pdf [accessed 2015-01-28] [WebCite Cache ID 6VvV0bT5V]
9. Lal S, Malla A. Service engagement in first episode psychosis: Current issues and future directions. Can J Psychiatry 2015 (forthcoming).
10. Steeves V. Young Canadians in a wired world, phase III: Online privacy, online publicity. Ottawa: MediaSmarts; 2014. 2014 URL: http://mediasmarts.ca/sites/mediasmarts/files/pdfs/publication-report/full/YCWWIII_Online_Privacy_Online_Publicity_FullReport.pdf [accessed 2015-01-28] [WebCite Cache ID 6VvVM9Epg]
11. Duggan M, Smith A. Social media update 2013. Washington, DC: Pew Research Center; 2013 Dec 30. URL: http://www.pewinternet.org/2013/12/30/social-media-update-2013/ [accessed 2015-01-28] [WebCite Cache ID 6VvVjJvtH]
12. Bennett K, Bennett AJ, Griffiths KM. Security considerations for e-mental health interventions. J Med Internet Res 2010 Dec;12(5):e61 [FREE Full text] [doi: 10.2196/jmir.1468] [Medline: 21169173]
13. Grajales FJ, Sheps S, Ho K, Novak-Lauscher H, Eysenbach G. Social media: A review and tutorial of applications in medicine and health care. J Med Internet Res 2014;16(2):e13 [FREE Full text] [doi: 10.2196/jmir.2912] [Medline: 24518354]
14. Amminger GP, Harris MG, Conus P, Lambert M, Elkins KS, Yuen HP, et al. Treated incidence of first-episode psychosis in the catchment area of EPPIC between 1997 and 2000. Acta Psychiatr Scand 2006 Nov;114(5):337-345. [doi: 10.1111/j.1600-0447.2006.00790.x] [Medline: 17022793]

Abbreviations

FEP: first-episode psychosis
SEI: specialized early intervention

Edited by G Eysenbach; submitted 07.03.15; peer-reviewed by D Ben-Zeev, L Farrer; comments to author 26.03.15; revised version received 26.04.15; accepted 26.04.15; published 20.05.15.

Please cite as:
Lal S, Dell'Elce J, Tucci N, Fuhrer R, Tamblyn R, Malla A
Preferences of Young Adults With First-Episode Psychosis for Receiving Specialized Mental Health Services Using Technology: A Survey Study
JMIR Mental Health 2015;2(2):e18
URL: http://mental.jmir.org/2015/2/e18/
doi:10.2196/mental.4400
PMID:26543922

©Shalini Lal, Jennifer Dell'Elce, Natasha Tucci, Rebecca Fuhrer, Robyn Tamblyn, Ashok Malla. Originally published in JMIR Mental Health (http://mental.jmir.org), 20.05.2015. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Factor Structure of the Internet Addiction Test in Online Gamers and Poker Players

Yasser Khazaal[1], MD; Sophia Achab[1], MD, PhD; Joel Billieux[2], PhD; Gabriel Thorens[1], MD; Daniele Zullino[1], MD; Magali Dufour[3], PhD; Stéphane Rothen[1], PhD

[1]Geneva University Hospitals, Geneva, Switzerland

[2]Laboratory for Experimental Psychopathology, Psychological Sciences Research Institute, Catholic University of Louvain, Louvain-la-Neuve, Belgium

[3]Sherbrooke University, Montreal, QC, Canada

Corresponding Author:
Yasser Khazaal, MD
Geneva University Hospitals
Grand Pré 70C
Geneva, 1206
Switzerland
Phone: 41 223725550
Fax: 41 223202840
Email: yasser.khazaal@hcuge.ch

Abstract

Background: The Internet Addiction Test (IAT) is the most widely used questionnaire to screen for problematic Internet use. Nevertheless, its factorial structure is still debated, which complicates comparisons among existing studies. Most previous studies were performed with students or community samples despite the probability of there being more problematic Internet use among users of specific applications, such as online gaming or gambling.

Objective: To assess the factorial structure of a modified version of the IAT that addresses specific applications, such as video games and online poker.

Methods: Two adult samples—one sample of Internet gamers (n=920) and one sample of online poker players (n=214)—were recruited and completed an online version of the modified IAT. Both samples were split into two subsamples. Two principal component analyses (PCAs) followed by two confirmatory factor analyses (CFAs) were run separately.

Results: The results of principal component analysis indicated that a one-factor model fit the data well across both samples. In consideration of the weakness of some IAT items, a 17-item modified version of the IAT was proposed.

Conclusions: This study assessed, for the first time, the factorial structure of a modified version of an Internet-administered IAT on a sample of Internet gamers and a sample of online poker players. The scale seems appropriate for the assessment of such online behaviors. Further studies on the modified 17-item IAT version are needed.

(JMIR Mental Health 2015;2(2):e12) doi:10.2196/mental.3805

KEYWORDS

Internet addiction; Internet Addiction Test (IAT); poker players; World of Warcraft; massively multiplayer online role playing; validation; factorial structure

Introduction

As the main medium of modern life, the Internet is used in a wide range of human activities. This expansion has numerous benefits, including its use for social, psychological, and medical purposes, as shown by a wide range of studies on eHealth [1-11]. In parallel, however, serious concerns have been raised related to problematic or excessive uncontrolled Internet use [12-15].

In recent years, several studies proposed incorporating Internet addiction as a new diagnosis into the Diagnostic and Statistical Manual of Mental Disorders (DSM) [16,17]. However, it was recently decided that empirical research is still too scarce to allow its inclusion as a new psychiatric disorder in the fifth edition (DSM-5). The validity of the Internet addiction construct is challenged by both theoretical and empirical concerns. For some researchers, Internet addiction is an umbrella construct that encompasses a variety of different dysfunctional behaviors

XSL•FO
RenderX

related to involvement in different online activities that do not necessarily coexist (eg, video games, cybersex, social networks, online gambling). Internet addiction thus seems to pertain to specific domains of use (ie, gaming, cybersex, social networks, or gambling) rather than to a general pattern of Internet use [12,18]. From this perspective, Internet gaming disorder (IGD) was proposed as a new condition and included in Section III of the DSM-5. The goal of Section III is to foster research on the conditions included therein [19].

With this potential new diagnosis, an important challenge is to develop assessment tools that are able to capture the specificity of this phenomenon, not only in terms of presence or absence of a given diagnostic, but also in terms of gradient severity. Since the initial research on Internet addiction, several psychometric measures have been developed [20-22]. One of the most translated tool is the Internet Addiction Test (IAT) [23]. The IAT is considered to be one of the instruments providing the most relevant clinical information about Internet gaming addiction [22,24]. The IAT has been moderately to highly correlated with a number of other measures of excessive Internet use, such as the Compulsive Internet Use Scale [25-28], the Generalized Problematic Internet Use Scale [29], the Revised Chen Internet Addiction Scale [30], and the Korean Internet Addiction Scale [31].

Despite the large diffusion of the IAT for research purposes, there is wide disagreement related first, to its factor structure [29] and second, to possible problematic items. In regard to the factor structure (see Multimedia Appendix 1), one to six factors were reported across studies published in English [27-45]. Furthermore, when similar numbers of factors were found, differences were observed in the distribution of the items on the factors. In addition, when more than one factor was extracted, a number of studies reported high correlations between factors [30,41,43], or a low contribution of some factors to the percentage of explained variance [27,35,42-45].

The issues related to the factorial structure of the IAT were possibly complicated by specific item-related concerns [28]. This led some authors to modify or discard items [27,28,30,31,34,41,45]. The most repeatedly reported concerns were the following:

1. Item 4—"How often do you form new relationships with fellow online users?"—has problematic loadings in a number of studies [27,28,36]. The recent rise of social networking, which widely disseminated this phenomenon, was suggested as a plausible explanation.

2. Item 6—on consequences on school work—and Item 8—related to job performance—ask about similar fields. However, the answer may differ depending on the participant's understanding and on the specific status of the participant. Unsurprisingly, covariance was repeatedly found between these items [29,30,32,33]. Some authors removed Item 8 [28,34,44,45] or Item 6 [45] after analysis.

3. A similar type of overlap was shown between Item 3—"How often do you prefer the excitement of the Internet to intimacy with your partner?"—and Item 19—"How often do you choose to spend more time online over going out with others?" This lead some authors to discard one of the items [28] or to modify Item 3 [27] because of the influence of lifestyle (ie, having a partner or not) on the answer options.

4. In contrast with the other IAT components, Item 7—"How often do you check your email before something else that you need to do?"—is not related to the Internet in general, but to a specific use (ie, emails). Concerns were reported in a number of studies [27-29,31,35,41,44], in part because of the wide dissemination of email, and in part because of changes in this type of communication—the automatic notification of its reception, thus, no need to "check", and its progressive replacement by social media. The item was deleted by some authors [28,30,41] and modified by others [31].

5. As a result of permanent Internet access (ie, without a specific need to log in), a rewording of Item 14 was proposed, as follows: "How often do you lose sleep due to being online late at night?" [28].

As shown, some of these IAT items involved specific patterns of life, such as being employed or being in a relationship. The "not applicable" answer option was probably included for this reason. It was, however, considered to be problematic by some authors [28] and was not systematically used.

In addition to conflicting results on the structure and certain items of the IAT, the psychometric characteristics were mostly assessed with students or community samples (Multimedia Appendix 1).

To our knowledge, no previous studies have assessed the psychometric characteristics of the instrument specifically for users of a given Internet application such as Internet games or gambling sites, despite the wide use of the scale, with or without modification, in studies related to these specific patterns of use [28,46,47].

In the context of increasing interest in possible Internet addiction-related disorders—with common involvement of gaming and gambling—and the emergence of the DSM-5 concept of the Internet gaming disorder, the use of a modified IAT for assessment of online gaming and online gambling may be worthwhile. Textbox 1 shows a modified version of the IAT for Internet game use.

Moreover, the lack of published studies on the psychometrical properties of the IAT on samples of gamers or gamblers appears to be an important weakness. This is of particular importance considering the increasing resemblance between gambling and gaming [48] and the possible links between related online activities (ie, massively multiplayer online role-playing games, online poker) and patterns of problematic Internet use [10,49-53].

The main goal of this study, therefore, was to investigate the factorial structure of the French version of the IAT modified for Internet gaming—or gambling—when used online, using samples of Internet gamers and Internet gamblers.

Table 1 shows the correspondence of the IAT items in Textbox 1 with the DSM-5 gaming and gambling criteria.

XSL•FO
RenderX

Textbox 1. The Internet Addiction Test modified for Internet game use.

IAT items modified for Internet game use:

1. How often do you find that you stay in-game longer than you intended?
2. How often do you neglect household chores to spend more time in-game?
3. How often do you prefer the excitement of the game to intimacy with your partner?
4. How often do you form new relationships with fellow game users?
5. How often do others in your life complain to you about the amount of time you spend in-game?
6. How often do your grades or school work suffer because of the amount of time you spend in-game?
7. How often do you check your email before something else that you need to do?
8. How often does your job performance or productivity suffer because of the game?
9. How often do you become defensive or secretive when anyone asks you what you do in-game?
10. How often do you block out disturbing thoughts about your life with soothing thoughts about the game?
11. How often do you find yourself anticipating when you will go in-game again?
12. How often do you fear that life without the game would be boring, empty, and joyless?
13. How often do you snap, yell, or act annoyed if someone bothers you while you are in-game?
14. How often do you lose sleep due to late-night log-ins?
15. How often do you feel preoccupied with the game when offline, or fantasize about being in-game?
16. How often do you find yourself saying "just a few more minutes" when in-game?
17. How often do you try to cut down the amount of time you spend in-game and fail?
18. How often do you try to hide how long you've been in-game?
19. How often do you choose to spend more time in-game over going out with others?
20. How often do you feel depressed, moody, or nervous when you are offline, which goes away once you are back in-game?

XSL•FO
RenderX

Table 1. Internet Addiction Test item correspondence with the Diagnostic and Statistical Manual of Mental Disorders, fifth edition, gaming and gambling disorders criteria.

Internet gaming disorder: proposed DSM-5 criteria	Gambling disorder: DSM-5 criteria
Preoccupation with Internet games (11, 15)[a]	Preoccupation with gambling (11, 15)
Withdrawal symptoms when Internet gaming is taken away (20)	Withdrawal (restless or irritable when attempt to cut down or stop gambling) (20)
Tolerance (the need to spend increasing amounts of time engaged in Internet games)[b]	Tolerance (needs to gamble with an increasing amount of money)[b]
Unsuccessful attempts to reduce or stop Internet game participation (1, 16, 17)	Unsuccessful attempts to reduce or stop gambling (1, 16, 17)
Loss of interest in other activities (3, 7, 19)	N/A[c]
Continues excessive use of Internet games despite problems (2, 6, 8, 14)	N/A
Deceives others regarding the amount of Internet gaming (5, 9, 13, 18)	Lies to conceal the importance of gambling involvement (9, 18)
Use of Internet games to escape from adverse moods (10)	Escape (often gambles when feeling distressed) (10)
Has jeopardized or lost relationships or opportunities due to excessive Internet gaming (8)	Has jeopardized or lost relationships or opportunities due to excessive gambling (8)
N/A	Relies on others to provide money to relieve the financial situations caused by gambling[b]
N/A	After losing money gambling, often returns to get even ("chasing" one's losses)[b]
Number of criteria: 5 or more	Number of criteria: 4 or more
Time criteria: 12 months or more	Time criteria: 12 months or more
Exclusion criteria: Internet use not related to online games is not "analogous to Internet gaming disorder"	Exclusion criteria: The behavior is better explained by a manic episode

[a]The suggested IAT items from Textbox 1 for each criteria are shown in parentheses.

[b]Not associated with an IAT item.

[c]Not applicable (N/A).

Methods

Participants

Two samples were used in this study: a French-speaking sample of World of Warcraft (WoW) players and a French-speaking sample of Internet poker players. These two samples completed the same modified version of the IAT. The ethical committee of the Department of Psychology of the University of Geneva—for the WoW sample—and the ethical committee of the Geneva University Hospitals—for the poker players sample—approved the study.

World of Warcraft Players Sample

The WoW sample was taken from a larger study on the relationships between players' self-reported motives to play and their in-game behaviors [46]. To participate in this study, an individual had to be a French-speaking WoW player and at least 18 years old. Participants were recruited through advertisements posted in dedicated French-language forums: a guilds forum, an official Blizzard WoW forum, and more general online and video games forums. Some participants also joined the study after hearing about it through the local press or television interviews. All participants gave their online consent prior to starting the online survey. The sample included 920 subjects who completed the French-language translation of the IAT [32] modified for online gaming. The mean age of IAT completers was 26.0 years (SD 7.8) and 807 of the 920 subjects (87.7%) were men.

Poker Players Sample

Overview

The poker players sample was taken from a study on online gambling. Inclusion criteria included playing online poker, speaking French, and being at least 18 years old. Participants were recruited through advertisements posted in dedicated French-language forums on online gambling or poker. All participants gave their online consent. The sample of poker players included 442 participants, of whom 214 (48.4%) completed the IAT. The mean age of IAT completers was 31.9 years (SD 9.5) and 425 of the 442 participants (96.2%) were men.

Measurements

All participants—WoW gamers and poker players—completed the same modified IAT. The scale is a 20-item auto-questionnaire [23] rated on a 5-point Likert scale from 1 (rarely) to 5 (always) with a maximum total score of 100. The rating also includes a "not applicable" option that has a rating of 0. In this study, we used the validated French version [32], which was adapted by replacing words directly related to the

XSL•FO
RenderX

Internet with words describing the specific activity (eg, "How often do you find that you stay online longer than you intended?" was replaced with "How often do you find that you stay in-game longer than you intended?").

Data Analyses

Because no clear factor structure has emerged in the literature and because different studies that found the same number of factors were inconsistent regarding factor loadings, we decided to assess the factor structure underlying this questionnaire from scratch (ie, without imposing a specific model or number of factors). In order to achieve this goal, both samples were randomly split into two subsamples of half of the size of the original ones (ie, 107 subjects for the poker sample and 460 for the WoW sample). Two principal component analyses (PCAs) were first performed on the first subsamples separately. With the discrete nature of the IAT items, PCA is preferred over factor analysis since PCA does not assume any particular multivariate model, which is not the case for factor analysis [54]. Moreover, it is known that when the same number of factors or components are extracted, both techniques yield highly similar results [55]. The number of components to extract was determined by the scree test [56] and by Velicer's minimum average partial (MAP) test done on the correlation matrix [57]. The reliability of the questionnaire was assessed by using the Cronbach alpha coefficient [58], which is a measure of internal consistency.

In a second step, two confirmatory factor analyses (CFAs) were conducted to validate the structure that emerged from the PCA. The CFAs were run on the second subsamples. For the same reasons that PCA was preferred, the unweighted least-square method was chosen as the procedure for estimation. Four preestablished criteria were selected as indicators of the goodness of fit to the data: (1) goodness-of-fit index >.90 [59], (2) adjusted goodness-of-fit index >.80 [59], (3) normed-fit index >.90 [60], and (4) root-mean-square error <.08. The use and cutoff of the goodness-of-fit index and the adjusted goodness-of-fit index were recommended by Cole [61], the normed-fit index by Bentler and Bonnet [60], and the root-mean-square error by Hu and Bentler [62].

The PCA was done with R 3.1.0, using *psych* and *bootstrap* packages, and the CFA was done with AMOS 21.0.0 [63].

Results

Principal Component Analysis

The MAP test and the scree test clearly suggested in both subsamples that one component be extracted. In order to evaluate the stability of the PCA, a bootstrap technique [64] was performed first with the MAP test, which confirmed the one-factor solution—among the 1000 bootstrap samples, 60.10% (601) and 84.60% (846) suggested retaining one factor in the poker and in the WoW subsamples, respectively. The bootstrap was also applied to factor loadings of the PCA. Items 4 and 6 had a very low loading on the factor, confirmed by the confidence intervals based on the bootstrap in both subsamples, suggesting that these questions may not be well-suited for the questionnaire (Table 2).

The percentage explained variance (95% CI) was 41.6 (31.6-51.1) for poker players and 36.1 (32.6-39.8) for WoW players. The reliability, as reported by Cronbach alpha (95% CI), was .92 (.88-.95) for poker players and .90 (.88-.92) for WoW players.

XSL•FO
RenderX

Table 2. The 20-item Internet Addiction Test results from principal component analysis.

Items	Estimated factor loadings (95% bootstrap CI)	
	Poker players (n=107)	WoW players (n=460)
1. How often do you find that you stay in-game longer than you intended?	.65 (.52-.75)	.49 (.38-.58)
2. How often do you neglect household chores to spend more time in-game?	.67 (.54-.77)	.73 (.68-.77)
3. How often do you prefer the excitement of the game to intimacy with your partner?	.71 (.54-.84)	.41 (.29-.52)
4. How often do you form new relationships with fellow game users?	0 (-.20 to .20)	.26 (.14-.37)
5. How often do others in your life complain to you about the amount of time you spend in-game?	.73 (.59-.86)	.61 (.53-.67)
6. How often do your grades or school work suffer because of the amount of time you spend in-game?	.29 (.04-.55)	.50 (.41-.59)
7. How often do you check your email before something else that you need to do?	.60 (.35-.76)	.67 (.60-.73)
8. How often does your job performance or productivity suffer because of the game?	.63 (.38-.80)	.66 (.57-.73)
9. How often do you become defensive or secretive when anyone asks you what you do in-game?	.65 (.42-.82)	.52 (.42-.61)
10. How often do you block out disturbing thoughts about your life with soothing thoughts about the game?	.72 (.57-.84)	.68 (.62-.74)
11. How often do you find yourself anticipating when you will go in-game again?	.57 (.39-.71)	.69 (.64-.73)
12. How often do you fear that life without the game would be boring, empty, and joyless?	.56 (.32-.73)	.64 (.54-.71)
13. How often do you snap, yell, or act annoyed if someone bothers you while you are in-game?	.62 (.38-.77)	.63 (.55-.70)
14. How often do you lose sleep due to late-night log-ins?	.68 (.48-.80)	.64 (.58-.69)
15. How often do you feel preoccupied with the game when offline, or fantasize about being in-game?	.64 (.46-.77)	.69 (.62-.74)
16. How often do you find yourself saying "just a few more minutes" when in-game?	.65 (.45-.78)	.57 (.50-.63)
17. How often do you try to cut down the amount of time you spend in-game and fail?	.77 (.62-.87)	.52 (.41-.61)
18. How often do you try to hide how long you've been in-game?	.80 (.64-.89)	.55 (.45-.63)
19. How often do you choose to spend more time in-game over going out with others?	.69 (.53-.81)	.64 (.56-.71)
20. How often do you feel depressed, moody, or nervous when you are offline, which goes away once you are back in-game?	.79 (.67-.86)	.72 (.64-.77)

Reliability

Cronbach alpha was above .90 in both subsamples, which was found to be excellent. It is worth noting that when Item 4 or Item 6 were removed, Cronbach alpha increased from .92 to .93 for poker players and from .90 to .91 for WoW players.

Confirmatory Factor Analysis

According to the cutoff defined above, all four goodness-of-fit indices were considered excellent in both subsamples (Table 3).

Shorter Version of the Internet Addiction Test: 17-Item Questionnaire

Because some items had low loadings and some questions had more missing values than occurred in the rest of the questionnaire, we performed additional investigations. In particular, Question 4—"How often do you form new relationships with fellow game users?"—seemed somewhat outdated and thus no longer relevant. Moreover, it had a low loading and decreased Cronbach alpha. Therefore, we decided to remove it.

Since Question 6—"How often do your grades or school work suffer because of the amount of time you spend in-game?"—is more suitable for school-aged persons, whereas Question 8—"How often does your job performance or productivity suffer because of the game?"—is more adapted to adults, we decided to merge the two questions into one. This new question addresses the consequences for the participant's principal occupation, either school or work, preventing the participant from omitting the answer because it is not applicable. For the same reasons, we also merged Question 3—"How often do you prefer the excitement of the game to intimacy with your partner?"—and Question 19—"How often do you choose to spend more time in-game over going out with others?"

These modifications led to a 17-item questionnaire. Despite the fact that this version had not been tested on new subjects, we performed the same analyses as we did for the original questionnaire—randomly split both samples into two

XSL•FO
RenderX

subsamples, running MAP, PCA, and CFA—by using the data at hand. For the merged questions, we decided to create two new items as follows: use the maximum mark of the IAT Item 3 and Item 19, as well as of Item 6 and Item 8, for each participant when both questions have been answered, or use the mark of only a single answered question.

In accordance with these modifications, the WoW sample size of IAT completers increased from 920 to 942 subjects, and the French-speaking poker sample size increased from 214 to 232. As expected, the results with this new method of coding the questionnaire led to the same conclusions regarding the number of components to extract—the bootstrapped MAP test suggested retaining only one factor in 74.2% and 95.6% of the poker and the WoW subsamples, respectively—regarding the factorial solution with the benefit of avoiding low loadings (Table 4), and regarding Cronbach alpha. Moreover, every fit index from the CFA was the same or slightly better (Table 3).

The percentage explained variance (95% CI) was 48.4 (37.5-58.1) for poker players and 37.3 (33.4-41.0) for WoW players. The reliability, as reported by Cronbach alpha (95% CI), was .93 (.90-.96) for poker players and .89 (.87-.91) for WoW players.

Table 3. Results from unweighted least-square confirmatory factor analysis.

Fit indices	Unweighted least squares			
	Poker players		WoW players	
	IAT20[a] (n=107)	IAT17[b] (n=116)	IAT20 (n=460)	IAT17 (n=471)
Root-mean-square residual	.08	.07	.08	.08
Goodness-of-fit index	.97	.97	.97	.97
Adjusted goodness-of-fit index	.97	.97	.96	.97
Normed-fit index	.96	.96	.95	.96

[a]The 20-item Internet Addiction Test (IAT20).

[b]The 17-item Internet Addiction Test (IAT17).

Table 4. The 17-item Internet Addiction Test results from principal component analysis.

Items	Estimated factor loadings (95% bootstrap CI)	
	Poker players (n=107)	WoW players (n=460)
1. How often do you find that you stay in-game longer than you intended?	.63 (.49-.73)	.51 (.41-.60)
2. How often do you neglect household chores to spend more time in-game?	.75 (.64-.84)	.70 (.65-.75)
Item 3 plus Item 19	.78 (.68-.86)	.64 (.57-.70)
5. How often do others in your life complain to you about the amount of time you spend in-game?	.68 (.51-.79)	.54 (.47-.62)
Item 6 plus Item 8	.65 (.53-.75)	.64 (.56-.71)
7. How often do you check your email before something else that you need to do?	.67 (.47-.79)	.69 (.62-.75)
9. How often do you become defensive or secretive when anyone asks you what you do in-game?	.70 (.48-.82)	.46 (.35-.55)
10. How often do you block out disturbing thoughts about your life with soothing thoughts about the game?	.76 (.60-.87)	.67 (.60-.73)
11. How often do you find yourself anticipating when you will go in-game again?	.59 (.42-.72)	.68 (.63-.73)
12. How often do you fear that life without the game would be boring, empty, and joyless?	.67 (.49-.80)	.59 (.49-.68)
13. How often do you snap, yell, or act annoyed if someone bothers you while you are in-game?	.61 (.41-.76)	.57 (.48-.65)
14. How often do you lose sleep due to late-night log-ins?	.63 (.43-.77)	.66 (.59-.71)
15. How often do you feel preoccupied with the game when offline, or fantasize about being in-game?	.69 (.52-.80)	.67 (.61-.73)
16. How often do you find yourself saying "just a few more minutes" when in-game?	.64 (.46-.76)	.55 (.47-.62)
17. How often do you try to cut down the amount of time you spend in-game and fail?	.73 (.55-.85)	.50 (.39-.60)
18. How often do you try to hide how long you've been in-game?	.79 (.63-.89)	.54 (.45-.62)
20. How often do you feel depressed, moody, or nervous when you are offline, which goes away once you are back in-game?	.81 (.63-.88)	.69 (.63-.75)

XSL•FO
RenderX

Discussion

Principal Findings

This study is the first to assess, to our knowledge, the psychometric characteristics, and specifically the factorial structure, of the IAT in online samples of WoW gamers and poker players. The main finding is that the one-factor model of the IAT has good psychometric properties and fits the data well in these samples.

In consideration to both the important discrepancies in the factorial solutions found in the previous studies on the IAT and the inconsistencies in the items included in a given factor in studies with a similar number of factors (Multimedia Appendix 1), the study at hand used an exploratory approach (ie, PCA) rather than an assessment of multiple competitive CFA models.

Although heterogeneous results were found regarding the one-factor solution in previously reported studies, it was considered the best factor solution, or as possibly an acceptable factor solution (Multimedia Appendix 1) [32-34,37-39,44,45].

As suggested elsewhere [39], the instability of the IAT factorial structure is possibly linked to its multiple strongly correlated facets. The variability of the samples studied, as well as some specific item-related aspects, may have contributed to the phenomenon. This consideration led some authors to propose modifications of the IAT [27,28,30,31,34,39].

The modified 17-item IAT scale proposed here (Table 4) also has a one-factor structure. The modification and combination of items related to a specific consequence on lifestyle (eg, students, workers) may reduce the sample variability of the proposed tool. The changes related to Item 6—impact on school—and Item 8—impact on occupation—are concordant with other findings [29,30,32,33], which previously led some authors to remove one of the two items [28,34,44,45]. Removing Item 6 or 8, however, may cause variable results, depending on the population studied (ie, students or workers). The modification proposed here may allow the answer to the combined item to depend less on the population studied. Similarly, observations about Item 3 and Item 19 of the IAT led some authors to discard or modify one of these items [28]. Again, the 17-item IAT may reduce sample-related changes in the scale characteristics.

The problematic loading of Item 4 was also reported in some studies [27,28,31,36]. In consideration of the emergence and wide dissemination of online social networks, this item appears to be less pertinent for the assessment of Internet addiction. In contrast with the results of previous studies, however, those of this study did not indicate specific problems related to Item 7—"How often do you check your email before something else that you need to do?" [27,28,35,41,44]—or Item 14—"How often do you lose sleep due to being online late at night?" [28]. Because of the concerns related to Item 7, some authors chose to delete it [28,30,41], whereas others proposed replacing "emails" with "Internet" [31]. In relation to this, the modification proposed by Lee and colleagues [31] seems relevant for further studies. The word "Internet" could be replaced with the specific Internet usage studied (ie, game). An interesting rewording of Item 14 was also proposed elsewhere [28] as follows: "How often do you lose sleep due to being online late at night?" The development of the Internet has led to possible permanent access without a requirement to log in. The lack of a specific problem with Items 7 and 14 in this study was possibly due to the age of the sample participants, for whom checking email or logging in were more familiar concepts than they were to younger people. The 17-item IAT could be adapted to include the modification proposed by Lee et al [31] for Item 7 and to that proposed by Pawlikowski et al [28] for Item 14 (Table 4).

The 17-item IAT scale (Textbox 2) and the 20-item IAT (Textbox 1, Table 1) offer interesting coverage of the main items of the DSM-5 criteria for IGD, with the exception of tolerance and time frame, which were not covered by these tools. As shown in Textbox 2, it is easy to add a time frame using a question such as "During the last year, how often…?"

As there is an absence of questions related to tolerance in the original IAT, it is then also the case for the proposed data-driven 17-item IAT. A further study may add questions specifically related to tolerance in order to assess the full range of symptoms proposed by the DSM-5 criteria for IGD.

The relatively good coverage of the 17-item IAT of the DSM-5 criteria for IGD is a possible advantage in comparison to other shorter forms of the IAT that lead, for example, to withdrawal of the escape-related item. Unsurprisingly, the 17-item IAT scale and the original IAT do not cover items related to the financial conflict-related items of the DSM-5 criteria for gambling disorder.

One of the strengths of this paper is related to the design of a 17-item IAT scale through a data-driven approach, rather than a priori choices.

Despite the possible lack of items linked to specific Internet use, it appears from this study that the original IAT and the 17-item IAT are interesting assessment tools for disorders related to excessive Internet use. In particular, the results assessed in this study add to the validity of IAT use with specific rewording, such as replacing "Internet" with "game."

XSL•FO
RenderX

Textbox 2. Proposed new 17-item Internet Addiction Test gaming questionnaire. a: Possible modification to: “How often do you stay in-game before something else that you need to do?“ b: Possible modification to: “How often do you lose sleep due to being in-game late at night?“

IAT items modified for the 17-item IAT (a time frame should be added to the top of the scale, such as “During the last year, how often….?”):

1. How often do you find that you stay in-game longer than you intended?
2. How often do you neglect household chores to spend more time in-game?
3. How often do you prefer the excitement of the game to intimacy with your partner, or to spend more time in-game over going out with others?
4. How often do others in your life complain to you about the amount of time you spend in-game?
5. How often do your grades or your school work, or your job performance or productivity suffer because of the amount of time you spend in-game?
6. How often do you check your email[a] before something else that you need to do?
7. How often do you become defensive or secretive when anyone asks you what you do in-game?
8. How often do you block out disturbing thoughts about your life with soothing thoughts about the game?
9. How often do you find yourself anticipating when you will go in-game again?
10. How often do you fear that life without the game would be boring, empty, and joyless?
11. How often do you snap, yell, or act annoyed if someone bothers you while you are in-game?
12. How often do you lose sleep due to late-night log-ins?[b]
13. How often do you feel preoccupied with the game when offline, or fantasize about being in-game?
14. How often do you find yourself saying "just a few more minutes" when in-game?
15. How often do you try to cut down the amount of time you spend in-game and fail?
16. How often do you try to hide how long you've been in-game?
17. How often do you feel depressed, moody, or nervous when you are offline, which goes away once you are back in-game?

Limitations

The main limitations of this study include the representativeness of self-selected samples [65], the lack of direct assessment with the modified 17-item IAT version, and the lack of concomitant diagnostic interviews.

Conclusions

Further studies may test the 17-item IAT among various samples in parallel with other psychopathological and Internet-related behavior assessments, including the use of the DSM-5 proposed criteria and other assessment tools, with possibly different or complementary coverage of the concept of Internet addiction [20,66-70]. The use of the original IAT, and possibly the 17-item IAT, may stimulate further research on IGD and Internet addiction. Due to the wide variety of Internet activities, such as Internet gambling and Internet gaming, it should be necessary to specify the assessed behavior (eg, WoW, other game, gambling) for each use of the 17-item IAT. Research in these fields seems to be of particular interest, as suggested by the recent inclusion of IGD in the DSM-5, Section III. The relevance of the proposed criteria in the DSM-5 needs to be confirmed, however, as some criteria, such as deception and tolerance, continue to be debated [17].

Conflicts of Interest

None declared.

Multimedia Appendix 1

Summary of previous studies on the Internet Addiction Test factorial structure.

[PDF File (Adobe PDF File), 111KB - mental_v2i2e12_app1.pdf]

References

1. Rodda S, Lubman DI, Dowling NA, Bough A, Jackson AC. Web-based counseling for problem gambling: exploring motivations and recommendations. J Med Internet Res 2013;15(5):e99 [FREE Full text] [doi: 10.2196/jmir.2474] [Medline: 23709155]
2. Powell J, Hamborg T, Stallard N, Burls A, McSorley J, Bennett K, et al. Effectiveness of a web-based cognitive-behavioral tool to improve mental well-being in the general population: randomized controlled trial. J Med Internet Res 2013;15(1):e2 [FREE Full text] [doi: 10.2196/jmir.2240] [Medline: 23302475]

XSL•FO
RenderX

3. Bolier L, Haverman M, Kramer J, Westerhof GJ, Riper H, Walburg JA, et al. An Internet-based intervention to promote mental fitness for mildly depressed adults: randomized controlled trial. J Med Internet Res 2013;15(9):e200 [FREE Full text] [doi: 10.2196/jmir.2603] [Medline: 24041479]
4. Harris IM, Roberts LM. Exploring the use and effects of deliberate self-harm websites: an Internet-based study. J Med Internet Res 2013;15(12):e285 [FREE Full text] [doi: 10.2196/jmir.2802] [Medline: 24362563]
5. van Gaalen JL, Beerthuizen T, van der Meer V, van Reisen P, Redelijkheid GW, Snoeck-Stroband JB, et al. Long-term outcomes of internet-based self-management support in adults with asthma: randomized controlled trial. J Med Internet Res 2013;15(9):e188 [FREE Full text] [doi: 10.2196/jmir.2640] [Medline: 24028826]
6. Choi NG, Dinitto DM. Internet use among older adults: association with health needs, psychological capital, and social capital. J Med Internet Res 2013 May;15(5):e97 [FREE Full text] [doi: 10.2196/jmir.2333] [Medline: 23681083]
7. Ramo DE, Prochaska JJ. Broad reach and targeted recruitment using Facebook for an online survey of young adult substance use. J Med Internet Res 2012;14(1):e28 [FREE Full text] [doi: 10.2196/jmir.1878] [Medline: 22360969]
8. Schaub M, Sullivan R, Stark L. Snow control - an RCT protocol for a web-based self-help therapy to reduce cocaine consumption in problematic cocaine users. BMC Psychiatry 2011;11:153 [FREE Full text] [doi: 10.1186/1471-244X-11-153] [Medline: 21943294]
9. Morel V, Chatton A, Cochand S, Zullino D, Khazaal Y. Quality of web-based information on bipolar disorder. J Affect Disord 2008 Oct;110(3):265-269. [doi: 10.1016/j.jad.2008.01.007] [Medline: 18280578]
10. Zanetta Dauriat F, Zermatten, Billieux J, Thorens G, Bondolfi G, Zullino D, et al. Motivations to play specifically predict excessive involvement in massively multiplayer online role-playing games: evidence from an online survey. Eur Addict Res 2011;17(4):185-189. [doi: 10.1159/000326070] [Medline: 21494046]
11. Khazaal Y, Chatton A, Cochand S, Coquard O, Fernandez S, Khan R, et al. Brief DISCERN, six questions for the evaluation of evidence-based content of health-related websites. Patient Educ Couns 2009 Oct;77(1):33-37. [doi: 10.1016/j.pec.2009.02.016] [Medline: 19372023]
12. Spada MM. An overview of problematic internet use. Addict Behav 2014 Jan;39(1):3-6. [doi: 10.1016/j.addbeh.2013.09.007] [Medline: 24126206]
13. Koo C, Wati Y, Lee CC, Oh HY. Internet-addicted kids and South Korean government efforts: boot-camp case. Cyberpsychol Behav Soc Netw 2011 Jun;14(6):391-394. [doi: 10.1089/cyber.2009.0331] [Medline: 21204693]
14. Aboujaoude E. Problematic Internet use: an overview. World Psychiatry 2010 Jun;9(2):85-90 [FREE Full text] [Medline: 20671890]
15. Kuss DJ, Griffiths MD, Karila L, Billieux J. Internet addiction: a systematic review of epidemiological research for the last decade. Curr Pharm Des 2014;20(25):4026-4052. [Medline: 24001297]
16. Block JJ. Issues for DSM-V: internet addiction. Am J Psychiatry 2008 Mar;165(3):306-307. [doi: 10.1176/appi.ajp.2007.07101556] [Medline: 18316427]
17. Tao R, Huang X, Wang J, Zhang H, Zhang Y, Li M. Proposed diagnostic criteria for internet addiction. Addiction 2010 Mar;105(3):556-564. [doi: 10.1111/j.1360-0443.2009.02828.x] [Medline: 20403001]
18. Cho H, Kwon M, Choi JH, Lee SC, Choi SW, Choi SW, et al. Development of the Internet addiction scale based on the Internet Gaming Disorder criteria suggested in DSM-5. Addict Behav 2014 Sep;39(9):1361-1366. [doi: 10.1016/j.addbeh.2014.01.020] [Medline: 24880884]
19. Petry NM, Rehbein F, Gentile DA, Lemmens JS, Rumpf HJ, Mößle T, et al. An international consensus for assessing internet gaming disorder using the new DSM-5 approach. Addiction 2014 Sep;109(9):1399-1406. [doi: 10.1111/add.12457] [Medline: 24456155]
20. Demetrovics Z, Urbán R, Nagygyörgy K, Farkas J, Griffiths MD, Pápay O, et al. The development of the Problematic Online Gaming Questionnaire (POGQ). PLoS One 2012;7(5):e36417 [FREE Full text] [doi: 10.1371/journal.pone.0036417] [Medline: 22590541]
21. King DL, Delfabbro PH, Griffiths MD. Trajectories of problem video gaming among adult regular gamers: an 18-month longitudinal study. Cyberpsychol Behav Soc Netw 2013 Jan;16(1):72-76. [doi: 10.1089/cyber.2012.0062] [Medline: 23098213]
22. Lortie CL, Guitton MJ. Internet addiction assessment tools: dimensional structure and methodological status. Addiction 2013 Jul;108(7):1207-1216. [doi: 10.1111/add.12202] [Medline: 23651255]
23. Young KS. Caught in the Net: How to Recognize the Signs of Internet Addiction--and a Winning Strategy for Recovery. New York, NY: John Wiley & Sons; Feb 1998.
24. King DL, Haagsma MC, Delfabbro PH, Gradisar M, Griffiths MD. Toward a consensus definition of pathological video-gaming: a systematic review of psychometric assessment tools. Clin Psychol Rev 2013 Apr;33(3):331-342. [doi: 10.1016/j.cpr.2013.01.002] [Medline: 23396015]
25. Meerkerk GJ, Van Den Eijnden RJ, Vermulst AA, Garretsen HF. The Compulsive Internet Use Scale (CIUS): some psychometric properties. Cyberpsychol Behav 2009 Feb;12(1):1-6. [doi: 10.1089/cpb.2008.0181] [Medline: 19072079]
26. Guertler D, Rumpf HJ, Bischof A, Kastirke N, Petersen KU, John U, et al. Assessment of problematic internet use by the Compulsive Internet Use Scale and the Internet Addiction Test: a sample of problematic and pathological gamblers. Eur Addict Res 2014;20(2):75-81 [FREE Full text] [doi: 10.1159/000355076] [Medline: 24080838]

XSL•FO
RenderX

27. Chong Guan N, Isa SM, Hashim AH, Pillai SK, Harbajan Singh MK. Validity of the Malay version of the internet addiction test: a study on a group of medical students in Malaysia. Asia Pac J Public Health 2015 Mar;27(2):NP2210-NP2219. [doi: 10.1177/1010539512447808] [Medline: 22652253]
28. Pawlikowski M, Altstotter-Gleich C, Brand M. Validation and psychometric properties of a short version of Young's Internet Addiction Test. Comput Human Behav 2013 May;29(3):1212-1223. [doi: 10.1016/j.chb.2012.10.014]
29. Barke A, Nyenhuis N, Kröner-Herwig B. The German version of the internet addiction test: a validation study. Cyberpsychol Behav Soc Netw 2012 Oct;15(10):534-542. [doi: 10.1089/cyber.2011.0616] [Medline: 23002984]
30. Lai CM, Mak KK, Watanabe H, Ang RP, Pang JS, Ho RC. Psychometric properties of the internet addiction test in Chinese adolescents. J Pediatr Psychol 2013 Aug;38(7):794-807 [FREE Full text] [doi: 10.1093/jpepsy/jst022] [Medline: 23671059]
31. Lee K, Lee HK, Gyeong H, Yu B, Song YM, Kim D. Reliability and validity of the Korean version of the internet addiction test among college students. J Korean Med Sci 2013 May;28(5):763-768 [FREE Full text] [doi: 10.3346/jkms.2013.28.5.763] [Medline: 23678270]
32. Khazaal Y, Billieux J, Thorens G, Khan R, Louati Y, Scarlatti E, et al. French validation of the internet addiction test. Cyberpsychol Behav 2008 Dec;11(6):703-706. [doi: 10.1089/cpb.2007.0249] [Medline: 18954279]
33. Hawi NS. Arabic validation of the Internet addiction test. Cyberpsychol Behav Soc Netw 2013 Mar;16(3):200-204. [doi: 10.1089/cyber.2012.0426] [Medline: 23441685]
34. Panayides P, Walker MJ. Evaluation of the psychometric properties of the Internet Addiction Test (IAT) in a sample of Cypriot high school students: the Rasch Measurement Perspective. EJOP 2012 Aug 29;8(3):327-351. [doi: 10.5964/ejop.v8i3.474]
35. Tsimtsiou Z, Haidich AB, Kokkali S, Dardavesis T, Young KS, Arvanitidou M. Greek version of the Internet Addiction Test: a validation study. Psychiatr Q 2014 Jun;85(2):187-195. [doi: 10.1007/s11126-013-9282-2] [Medline: 24307176]
36. Jelenchick LA, Becker T, Moreno MA. Assessing the psychometric properties of the Internet Addiction Test (IAT) in US college students. Psychiatry Res 2012 Apr 30;196(2-3):296-301 [FREE Full text] [doi: 10.1016/j.psychres.2011.09.007] [Medline: 22386568]
37. Pontes HM, Patrão IM, Griffiths MD. Portuguese validation of the Internet Addiction Test: An empirical study. J Behav Addict 2014 Jun;3(2):107-114 [FREE Full text] [doi: 10.1556/JBA.3.2014.2.4] [Medline: 25215221]
38. Korkeila J, Kaarlas S, Jääskeläinen M, Vahlberg T, Taiminen T. Attached to the web--harmful use of the Internet and its correlates. Eur Psychiatry 2010 May;25(4):236-241. [doi: 10.1016/j.eurpsy.2009.02.008] [Medline: 19556111]
39. Teo T, Kam C. Validity of the Internet Addiction Test for Adolescents and Older Children (IAT-A): tests of measurement invariance and latent mean differences. J Psychoeduc Assess 2014 May 07;32(7):624-637. [doi: 10.1177/0734282914531708]
40. Ferraro G, Caci B, D'Amico A, Di Blasi M. Internet addiction disorder: an Italian study. Cyberpsychol Behav 2007 Apr;10(2):170-175. [doi: 10.1089/cpb.2006.9972] [Medline: 17474832]
41. Chang MK, Law SPM. Factor structure for Young's Internet Addiction Test: A confirmatory study. Comput Human Behav 2008;24(6):2597-2619.
42. Widyanto L, Griffiths MD, Brunsden V. A psychometric comparison of the Internet Addiction Test, the Internet-Related Problem Scale, and self-diagnosis. Cyberpsychol Behav Soc Netw 2011 Mar;14(3):141-149. [doi: 10.1089/cyber.2010.0151] [Medline: 21067282]
43. Widyanto L, McMurran M. The psychometric properties of the internet addiction test. Cyberpsychol Behav 2004 Aug;7(4):443-450. [doi: 10.1089/cpb.2004.7.443] [Medline: 15331031]
44. Faraci P, Craparo G, Messina R, Severino S. Internet Addiction Test (IAT): which is the best factorial solution? J Med Internet Res 2013;15(10):e225 [FREE Full text] [doi: 10.2196/jmir.2935] [Medline: 24184961]
45. Watters CA, Keefer KV, Kloosterman PH, Summerfeldt LJ, Parker JDA. Examining the structure of the Internet Addiction Test in adolescents: A bifactor approach. Comput Human Behav 2013;29(6):2294-2302.
46. Billieux J, Van der Linden M, Achab S, Khazaal Y, Paraskevopoulos L, Zullino D, et al. Why do you play World of Warcraft? An in-depth exploration of self-reported motivations to play online and in-game behaviours in the virtual world of Azeroth. Comput Human Behav 2013 Jan;29(1):103-109.
47. Brand M, Laier C, Pawlikowski M, Schächtle U, Schöler T, Altstötter-Gleich C. Watching pornographic pictures on the Internet: role of sexual arousal ratings and psychological-psychiatric symptoms for using Internet sex sites excessively. Cyberpsychol Behav Soc Netw 2011 Jun;14(6):371-377. [doi: 10.1089/cyber.2010.0222] [Medline: 21117979]
48. King D, Delfabbro P, Griffiths M. The convergence of gambling and digital media: implications for gambling in young people. J Gambl Stud 2010 Jun;26(2):175-187. [doi: 10.1007/s10899-009-9153-9] [Medline: 19756979]
49. Dufour M, Gendron A, Brunelle N, Leclerc D, Cousineau MM. Adolescent technology use: Profiles of distinct groups and associated risky behaviors. J Addict Res Ther (in press) 2015 (forthcoming).
50. Billieux J, Chanal J, Khazaal Y, Rochat L, Gay P, Zullino D, et al. Psychological predictors of problematic involvement in massively multiplayer online role-playing games: illustration in a sample of male cybercafé players. Psychopathology 2011;44(3):165-171. [doi: 10.1159/000322525] [Medline: 21372629]
51. Yau YH, Pilver CE, Steinberg MA, Rugle LJ, Hoff RA, Krishnan-Sarin S, et al. Relationships between problematic internet use and problem-gambling severity: findings from a high-school survey. Addict Behav 2014 Jan;39(1):13-21 [FREE Full text] [doi: 10.1016/j.addbeh.2013.09.003] [Medline: 24140304]

XSL•FO
RenderX

52. Starcevic V. Is Internet addiction a useful concept? Aust N Z J Psychiatry 2013 Jan;47(1):16-19. [doi: 10.1177/0004867412461693] [Medline: 23293309]
53. Durkee T, Kaess M, Carli V, Parzer P, Wasserman C, Floderus B, et al. Prevalence of pathological internet use among adolescents in Europe: demographic and social factors. Addiction 2012 Dec;107(12):2210-2222. [doi: 10.1111/j.1360-0443.2012.03946.x] [Medline: 22621402]
54. Schneeweiss H, Mathes H. Factor analysis and principal components. J Multivar Anal 1995 Oct;55(1):105-124. [doi: 10.1006/jmva.1995.1069]
55. Velicer WF, Jackson DN. Component analysis versus common factor analysis: some issues in selecting an appropriate procedure. Multivariate Behav Res 1990 Jan;25(1):1-28. [doi: 10.1207/s15327906mbr2501_1]
56. Cattell RB. The scree test for the number of factors. Multivariate Behav Res 1966 Apr;1(2):245-276. [doi: 10.1207/s15327906mbr0102_10]
57. Velicer WF. Determining the number of components from the matrix of partial correlations. Psychometrika 1976;41:321-327.
58. Cronbach LJ, Meehl PE. Construct validity in psychological tests. Psychol Bull 1955 Jul;52(4):281-302. [doi: 10.1037/h0040957]
59. Joreskog KG, Sorbom D. LISREL 8: User's Reference Guide. Skokie, IL: Scientific Software International; 1996.
60. Bentler PM, Bonett DG. Significance tests and goodness of fit in the analysis of covariance structures. Psychol Bull 1980;88(3):588-606. [doi: 10.1037/0033-2909.88.3.588]
61. Cole DA. Utility of confirmatory factor analysis in test validation research. J Consult Clin Psychol 1987 Aug;55(4):584-594. [Medline: 3624616]
62. Hu LT, Bentler P. Cutoff criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. Struct Equ Modeling 1999;6:1-55. [doi: 10.1080/10705519909540118]
63. Arbuckle JL. IBM Corporation. Crawfordville, FL: Amos Development Corporation; 2012. IBM SPSS Amos 21: User's Guide URL: ftp://public.dhe.ibm.com/software/analytics/spss/documentation/amos/21.0/en/Manuals/IBM_SPSS_Amos_Users_Guide.pdf [accessed 2015-04-16] [WebCite Cache ID 6XpjVO9eA]
64. Efron B. The Jackknife, the Bootstrap, and Other Resampling Plans. Philadelphia, PA: Society for Industrial and Applied Mathematics; 1987.
65. Khazaal Y, van Singer M, Chatton A, Achab S, Zullino D, Rothen S, et al. Does self-selection affect samples' representativeness in online surveys? An investigation in online video game research. J Med Internet Res 2014;16(7):e164 [FREE Full text] [doi: 10.2196/jmir.2759] [Medline: 25001007]
66. Demetrovics Z, Szeredi B, Rózsa S. The three-factor model of Internet addiction: the development of the Problematic Internet Use Questionnaire. Behav Res Methods 2008 May;40(2):563-574. [Medline: 18522068]
67. Caplan SE. Theory and measurement of generalized problematic Internet use: A two-step approach. Comput Human Behav 2010;26(5):1089-1097.
68. Meerkerk GJ, Van Den Eijnden RJ, Vermulst AA, Garretsen HF. The Compulsive Internet Use Scale (CIUS): some psychometric properties. Cyberpsychol Behav 2009 Feb;12(1):1-6. [doi: 10.1089/cpb.2008.0181] [Medline: 19072079]
69. Lemmens JS, Valkenburg PM, Peter J. Development and validation of a game addiction scale for adolescents. Media Psychol 2009 Mar 05;12(1):77-95. [doi: 10.1080/15213260802669458]
70. Mak KK, Lai CM, Ko CH, Chou C, Kim DI, Watanabe H, et al. Psychometric properties of the Revised Chen Internet Addiction Scale (CIAS-R) in Chinese adolescents. J Abnorm Child Psychol 2014 Oct;42(7):1237-1245. [doi: 10.1007/s10802-014-9851-3] [Medline: 24585392]

Abbreviations

CFA: confirmatory factor analysis
DSM: Diagnostic and Statistical Manual of Mental Disorders
DSM-5: Diagnostic and Statistical Manual of Mental Disorders, fifth edition
IAT: Internet Addiction Test
IAT17: 17-item Internet Addiction Test
IAT20: 20-item Internet Addiction Test
IGD: Internet gaming disorder
MAP: minimum average partial
N/A: not applicable
PCA: principal component analysis
WoW: World of Warcraft

XSL•FO
RenderX

Edited by G Eysenbach; submitted 25.08.14; peer-reviewed by Z Demetrovics, G Craparo, G Floros; comments to author 10.01.15; revised version received 24.02.15; accepted 06.03.15; published 22.04.15.

Please cite as:
Khazaal Y, Achab S, Billieux J, Thorens G, Zullino D, Dufour M, Rothen S
Factor Structure of the Internet Addiction Test in Online Gamers and Poker Players
JMIR Mental Health 2015;2(2):e12
URL: http://mental.jmir.org/2015/2/e12/
doi:10.2196/mental.3805
PMID:26543917

©Yasser Khazaal, Sophia Achab, Joel Billieux, Gabriel Thorens, Daniele Zullino, Magali Dufour, Stéphane Rothen. Originally published in JMIR Mental Health (http://mental.jmir.org), 22.04.2015. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Original Paper

Twitter: A Novel Tool for Studying the Health and Social Needs of Transgender Communities

Evan A Krueger[1], MPH; Sean D Young[2], MS, PhD

[1]Department of Community Health Sciences, University of California, Los Angeles, Los Angeles, CA, United States

[2]UC Institute for Prediction Technology, Department of Family Medicine, University of California, Los Angeles, Los Angeles, CA, United States

Corresponding Author:
Evan A Krueger, MPH
Department of Community Health Sciences
University of California, Los Angeles
650 Charles E. Young Drive South, 36-071 CHS
Box 951772
Los Angeles, CA, 90095
United States
Phone: 1 310 794 0619
Fax: 1 N/A
Email: eakrueger@ucla.edu

Abstract

Background: Limited research has examined the health and social needs of transgender and gender nonconforming populations. Due to high levels of stigma, transgender individuals may avoid disclosing their identities to researchers, hindering this type of work. Further, researchers have traditionally relied on clinic-based sampling methods, which may mask the true heterogeneity of transgender and gender nonconforming communities. Online social networking websites present a novel platform for studying this diverse, difficult-to-reach population.

Objective: The objective of this study was to attempt to examine the perceived health and social needs of transgender and gender nonconforming communities by examining messages posted to the popular microblogging platform, Twitter.

Methods: Tweets were collected from 13 transgender-related hashtags on July 11, 2014. They were read and coded according to general themes addressed, and a content analysis was performed. Qualitative and descriptive statistics are presented.

Results: There were 1135 tweets that were collected in total. Both “positive” and “negative” events were discussed, in both personal and social contexts. Violence, discrimination, suicide, and sexual risk behavior were discussed. There were 34.36% (390/1135) of tweets that addressed transgender-relevant current events, and 60.79% (690/1135) provided a link to a relevant news article or resource.

Conclusions: This study found that transgender individuals and allies use Twitter to discuss health and social needs relevant to the population. Real-time social media sites like Twitter can be used to study issues relevant to transgender communities.

(JMIR Mental Health 2015;2(2):e16) doi:10.2196/mental.4113

KEYWORDS

Twitter; social media; transgender; health

Introduction

Difficulties in Studying Transgender Communities

Limited research has been conducted that aims to better understand the shared experiences of transgender and gender nonconforming individuals. Despite an increasing interest in transgender health issues, researchers and practitioners have struggled to conduct studies that are representative of this diverse and heterogeneous population. Transgender individuals may not be identifiable, and may take action to remain hidden in public spaces [1], impeding study recruitment through traditional means. Much research to date has relied on clinic-based sampling methods. This approach may oversample transsexual individuals, who commonly seek medical intervention [2]. Doing so may bias findings about the most salient health needs of transgender communities as medical in nature.

XSL•FO
RenderX

Despite the diversity inherent to the transgender population, individuals often share common experiences of violence, harassment, and discrimination [3-5]. Transgender youths are more likely to drop out of school, report difficulty finding employment, and report engagement in sex work, compared to the general population [6-8]. Such experiences lead to elevated rates of mental and psychological distress [9,10], as well as decreased rates of health care utilization [11], likely due to perceived stigma or ignorance of transgender health issues from providers [12,13]. Further, transgender people are at significantly elevated risk of suicide ideation and attempt [14,15].

It is imperative to better understand the nature of these experiences in order for public health and social work practitioners to implement effective intervention strategies. With notable exceptions [2,9], few studies have employed Web-based approaches to studying transgender populations. To our knowledge, no other studies have used online social media websites for such a purpose.

Transgender Communities and Online Social Networks

The Internet, and particularly social media websites, have shaped the way that people from marginalized, and otherwise stigmatized, groups communicate by allowing them to organize in large numbers around common interests and goals [2]. Evidence suggests that sexual minority (lesbian, gay, bisexual) individuals are more likely to use social media technologies than heterosexual individuals [16]. Online social networks allow users a certain degree of anonymity, possibly allowing researchers access to population groups in which privacy and secrecy are important. This may include some transgender people who feel obligated to publically hide their gender identities out of fear, shame, and desire to change [1].

Prior qualitative work has demonstrated that men who have sex with men (MSM) willingly enrolled in an online social media-based intervention and voluntarily had conversations with each other about common areas of interest and perceived health needs. This work demonstrated the unique potential to understand this population through a social media-based approach [17]. Statistics on social media usage are not available for transgender communities, but youth report having little access to information on transgender health concerns, and that this type of knowledge is commonly sought through the Internet [7].

This study attempts to determine the feasibility of accessing communities of transgender individuals and allies through the online microblogging platform, Twitter. We performed a content analysis of tweets (messages) posted to several transgender-related hashtags (analogous to discussion boards) in order to classify the types of information being communicated, the ways in which they are communicated, and to gain a community-level perspective on the perceived health and social needs faced by transgender communities. This approach represents a novel, effective way of accessing information about the needs of a population that has traditionally been difficult to reach.

Methods

Institutional Review Board

This is an observational study, which used publically available data. As such, the study does not meet the definition of human subjects research and was exempted from review by the Office of the Human Research Protection Program at the University of California, Los Angeles (UCLA).

Twitter

Twitter is a popular social networking platform that allows users to post brief messages, limited to 140 characters, called “tweets”. Tweets can be shared with other users in a number of ways. These include posting the message to the original tweeter’s or another user’s or organization’s home page, where it can also be viewed easily by all other users who view those pages. Users have the option to “follow” other users, automatically updating them whenever a change is made to the followed page. In addition, users may retweet (repost) tweets of interest and affix tweets with a “hashtag”, denoted with a “#” applied before a keyword (for example, “#transgender”). Hashtagging a tweet places the message on a communal discussion board specific to that hashtag. This function is especially useful for allowing users to organize around a topic of interest, by placing all tweets affixed with a given hashtag in a central location that is easily searchable [18].

Data Collection and Processing

Tweets were collected through the applied use of hashtags. All tweets posted to 13 different transgender and gender nonconforming-specific hashtags were collected over a one-day period on July 11, 2014 (Table 1). Due to a software malfunction, only a subset of tweets was collected from the hashtag, “ftm”. Therefore, this hashtag was excluded from analysis. Due to the large number of tweets made daily to these hashtags, one 24-hour period was considered sufficient for this feasibility study. A thorough Internet search was performed to identify hashtags that were currently in use, and designed for transgender community usage. Given the rapidly evolving nature by which social media platforms are used, as well as the identifying terminology used by transgender populations, only hashtags that appeared to be specific to transgender communities on the date of collection were selected for inclusion in the study.

Tweets were collected using the Twitter Archiving Google Spreadsheet (TAGS) version 5, a programmable Google Spreadsheet developed for time-based tweet collection [18,19]. The spreadsheet was programmed to collect all new tweets posted to an individual hashtag on an hourly basis. There were 5860 tweets that were collected in total. After compiling all tweets and removing duplicates (occurred when multiple hashtags were applied to a single tweet, resulting in the TAGS software collecting the same tweet from both hashtags), 5454 tweets remained. Despite the consideration with which hashtags were chosen for inclusion in the study, a considerable number of tweets irrelevant to transgender communities were posted to the hashtags, and 3161 such tweets were removed from analysis. Finally, non-English language tweets, and tweets containing sexually explicit material were also removed, resulting in a final

XSL•FO
RenderX

dataset of 1135 analyzed tweets. Tweets containing sexual health-related text were relevant to this analysis, and so were left in the dataset.

Table 1. Hashtags from which tweets were collected.

Hashtags	Total	Proportion of total
#trans	507	0.45
#transgender	457	0.40
#girlslikeus	161	0.14
#mtf	47	0.04
#genderqueer	14	0.01
#nonbinary	12	0.01
#genderbender	11	0.01
#f2m	5	0.00
#m2f	2	0.00
#transguy	2	0.00
#transproblems	2	0.00
#dysphoria	1	0.00
#transgirl	1	0.00

Data Analysis

A content analysis was performed. Tweets were coded and analyzed in a Web-based mixed-methods data analysis software, Dedoose [20]. Tweets were coded along multiple dimensions: (1) hashtag(s) from which a tweet originated, (2) general themes or topics addressed within the message, and (3) whether or not information was shared in a distinctive way (whether the tweet was an original post or a retweet, whether a link to a relevant news article was included in the message, and whether the message was specifically posted on an organization's home page).

Codes were developed as general themes emerged from the data, and were assigned to all tweets to which they applied. The coding process, completed by a single researcher with expertise in transgender topics, was completed twice. No new codes were applied during this second review of the data; the purpose was to ensure that all tweets were coded in a consistent way. Care was taken to code tweets according to the text presented within them, and to avoid inferring meaning as much as possible. As such, it was possible for different tweets written about the same issue to be coded in slightly different ways. Table 2 shows examples of tweets, as they were coded.

Table 2. Topics discussed among transgender and gender nonconforming individuals and allies on transgender-specific hashtags on Twitter.

Category	Example tweet	Codes applied[a]
Social (positive)	Atherton High School #Kentucky Finalizes #Anti-Discrimination Policies For #Transgender Students [link to news article] via @wfplnews Nice!	6
	we are all #equal === sending my love to #gay #lesbian #bisexual #transgender peeps xoxoxo … #LGBTPrideMonth #equal #rights #marriage	3, 23
Social (negative)	Hobby Lobby takes stand against "Transgender Restrooms" #Transgender #HobbyLobby [link to news article]	11
	W/out accurate ID #transgender people face tremendous difficulty fully participating in society, finding jobs or accessing benefits & services	12, 15, 18
Personal (positive)	One year hrt #transgender #transpride #over40 #happy [link to image]	23, 24
	Last wknd in "boy mode" around my kids. Finally coming out to them as my authentic self later this month #transgender [link to image]	24
Personal (negative)	The worst part about having a female body is the fact that when I swim I have to wear a top #transgender #ftm #swimming #bathingsuit	27
	RT @cns_health: Press Conf #AIDS2014: #MSM & #transgender need to protect themselves from anal STIs & #HIV [link to article]	25

[a] Codes were applied as numbered in Table 3

XSL•FO
RenderX

Results

Hashtags Applied to Tweets

Hashtags were applied 1222 times to 1135 independent tweets. Further, the following three hashtags represented 92.06% (1125/1222) of the total hashtags applied: “#transgender”, “#trans”, and “#girlslikeus” (Table 1). There were 94.45% (1072/1135) of tweets that addressed topical areas that fell into at least one of four major categories, each containing several subcategories, applied as codes: “social (positive)”, “social (negative)”, personal (positive)”, and “personal (negative)” (Table 3).

Contents of Tweets

Social (positive) tweets were represented by messages that described socially progressive ideas, events, or actions taken by others beyond the original tweeter. Commonly occurring themes in this category included messages of ally affirmation or support for transgender individuals and causes, discussions around improving social conditions for transgender individuals, and messages related to research conducted that was specific to transgender needs. Social (positive) tweets accounted for 54.71% (621/1135) of all tweets in the dataset. Social (negative) tweets accounted for 26.34% (299/1135) of all tweets, and were characterized by messages related to socially conservative ideas, or actions that harm transgender individuals or communities. Examples of such themes include descriptions of discriminatory policies, violence experienced by transgender individuals, or accounts of ignorance about transgender-specific issues. Personal (positive) tweets, accounting for 11.10% (126/1135) of all tweets, included messages of self-affirmation, self-affirming comments regarding one’s physical appearance, or declarations of pride. Personal (negative) tweets accounted for 2.29% (26/1135) of all tweets, and included messages related to body image dysphoria, prior suicide attempts, and sexual risk behaviors.

There were five current events pertinent to transgender communities that were heavily discussed, accounting for 34.36% (390/1135) of all tweets analyzed, and were included in the major categories described above. Laverne Cox, a transgender actress and activist had recently been nominated for an Emmy award for her role in a television series, accounting for 19.82% (225/1135) of all tweets. Tweets related to Sparkle, a transgender pride festival held annually in July in Manchester, United Kingdom accounted for 4.58% (52/1135) of tweets analyzed. An Internet video of a Fox News clip, in which interviewees failed to become upset upon learning about a local gender inclusive restroom policy was also shared in 0.70% (8/1135) of tweets. Finally, information was shared about two organizations’ discriminatory policies, both having recently garnered media attention. George Fox University, a Christian university, had recently been granted a Title IX exemption allowing the university to deny gender-appropriate housing to transgender students, accounting for 5.81% (66/1135) of tweets. Hobby Lobby, a popular chain craft store, was highlighted in news stories for refusing a transgender employee access to a gender-appropriate restroom, accounting for 3.44% (39/1135) of tweets.

Despite the positive or negative themes present within the tweets, the ideological orientations/attitudes of the individuals posting them were overwhelmingly positive and affirming of transgender communities. Of all tweets analyzed, four were coded as having come directly from individuals espousing clearly “negative” views toward transgender people, for example,

> *This "#Baptist" #Church thinks it's OK to have a #Transgender pastor. (link to news article) - Satan has hijacked the Baptist name!*

In addition, 60.79% (690/1135) of tweets contained a URL linking the reader to a news article or resource directly related to the central message of the tweet. Retweets consisted of 46.43% (527/1135) of tweets, whereby a user copies and reposts a tweet originating from another user. Finally, 33.13% (376/1135) of tweets linked the message text to an organization’s Twitter account.

XSL•FO
RenderX

Table 3. Four major categories of conversation, each consisting of several subcategories.

Categories[a]		Total	Proportion of total
Social (positive)			621/1135 (54.71%)
	1. Laverne Cox Emmy nomination	225	0.20
	2. Advertisement of an event, service, or job/solicitation of help at an event/on a project	120	0.11
	3. Ally affirmation/support	84	0.07
	4. Sparkle festival	52	0.05
	5. Transgender celebrity news (not including Laverne Cox Emmy nomination)	48	0.04
	6. Descriptions of improvements in social condition for transgender individuals	36	0.03
	7. Transgender-specific research findings	31	0.03
	8. Recognition of correct gender identity on legal documents	17	0.01
	9. Video of a Fox News interview, depicting interviewees who failed to become upset about enactment of a gender inclusive restroom policy	8	0.01
Social (negative)			299/1135 (26.34%)
	10. George Fox University (housing discrimination)	66	0.06
	11. Hobby Lobby discriminatory restroom policy	39	0.03
	12. Discrimination (general)	37	0.03
	13. Violence	31	0.03
	14. Lack of understanding, ignorance toward transgender issues or individuals	29	0.03
	15. Workplace discrimination	24	0.02
	16. Descriptions of other individuals holding negative views toward transgender people[b]	24	0.02
	17. Police mistreatment or brutality	17	0.01
	18. Refusal to recognize one's gender identity, either interpersonally or legally	13	0.01
	19. Discriminatory treatment of prison inmates	8	0.01
	20. Tokenizing of transgender population, generalizations applied to whole population	6	0.01
	21. Insulting remark about transgender celebrity	5	0.00
Personal (positive)			126/1135 (11.10%)
	22. Interest in appearance	51	0.04
	23. Pride	38	0.03
	24. Self affirmation	37	0.03
Personal (negative)			26/1135 (2.29%)
	25. Sexual risk behavior	10	0.01
	26. Suicide	7	0.01
	27. Body image dysphoria	3	0.00
	28. Depression	3	0.00
	29. Eating disorders	3	0.00

[a] Topics presented represent major categories and themes of conversation presented within tweets. Individual messages marked with a specific code, for example, violence, may not pertain to the same event.

[b] Does not indicate a negatively held view by the original tweeter; instead, indicates the negative views of others, as acknowledged within the text of a tweet.

XSL•FO
RenderX

Discussion

Online Social Media

Online social media provides users with platforms to interact and share information in increasingly innovative ways, and provides researchers with new points of access to information. Twitter is one such platform, in which short messages, limited to 140 characters, may be composed and shared with a wide audience. Additionally, such platforms allow users to share as much or as little personal information with other users as is desired. Users are able to create a username of their choice and omit additional identifying information. As such, Twitter provides a platform for discussions about controversial or taboo topics with relative anonymity, that many may be uncomfortable discussing in a less anonymous setting.

Information is shared commonly, and with relative ease among Twitter users. This is evidenced both by the large number of tweets providing links to news articles and the high rates of retweeting and posting of tweets to organizational Twitter accounts. Organizational pages may be seen as central locations in which Twitter users can easily find and access information. The high degree of information sharing and the methods of sharing employed by users may be of importance to consider when designing effective community-based prevention strategies. Online prevention strategies, specifically those delivered through social networking websites, may benefit by encouraging peer-to-peer information sharing.

Transgender Support on Twitter

Transgender individuals and allies use Twitter to discuss a large range of topics relevant to the social conditions they experience. Additionally, over one-third of tweets addressed relevant current events, demonstrating that Twitter provides a platform for sharing news about such topics with other members of the community in real-time, and potentially for mobilizing around an issue of interest. It is important, and rather interesting to note, the disproportionately large number of "positive" posts over "negative" ones. While it is possible that the ratio of positive to negative tweets is unique to the date on which data were collected (July 11, 2014), or to the specific hashtags surveyed, it may also be indicative of the purpose that Twitter serves to transgender individuals and allies, as a place of camaraderie and support. The finding that only 4 tweets were posted by users who openly espoused negative views of transgender communities further supports this notion. Indeed, while many tweets certainly did address the very real health and social issues faced by transgender communities, the Twitter hashtags surveyed appear to serve as a place of optimism for social progress.

Limitations

This method of data collection may represent a novel, effective way of accessing information about a population that has traditionally been difficult to reach. However, our study presents several limitations to researchers. First, and perhaps most limiting, tweeters utilize the platform in an uncontrolled manner. They are free to post messages to the hashtags from which data are collected, to other hashtags, or to no hashtag at all. Indeed, a substantial number of tweets initially collected were dropped from analysis because they were irrelevant to transgender issues. The nature of this research is purely observational, and so significant data cleaning and processing was necessary to abstract relevant information. Further, without careful selection of hashtags from which to collect tweets, large portions of relevant data, and potentially from distinct demographic subgroups, may go uncollected. Additionally, tweets are written and posted on a voluntary basis. Our dataset includes tweets from users who were interested in talking about transgender-related issues on the date of collection, and is not specific to transgender-identified individuals. However, given the overwhelmingly "positive" natures of those who posted, it is clear that this sample includes largely transgender-identified individuals and allies.

Second, given the relative anonymity of such networks, it is impossible to obtain accurate demographic information on those being studied. As such, we do not know the ages, races, ethnicities, educational, or employment backgrounds, or even the identified genders of participants, unless explicitly stated by the user. Additionally, only tweets written in English, and on the single day of data collection are represented in this sample. These findings may not remain stable, as pertinent issues vary over time, and across different language and cultural groups. While Twitter may represent a novel approach to understanding transgender-related needs, it is not able to provide representative population estimates. Instead, it is able to provide meaningful and real-time information about a specific population of people, Twitter-using individuals who are interested in transgender-related topics.

Third, we attempted to access tweets from a wide range of transgender-related hashtags. As such, our results paint a global picture of transgender-related issues. Given the heterogeneity of individuals and communities that qualify as "transgender", it is important to recognize that despite some common experiences, different subpopulations will likely communicate different needs and concerns. Future studies should attempt to disentangle the unique needs of specific transgender communities, such as male-to-female, female-to-male, and gender nonconforming populations.

Finally, social acceptance and usage of different social networking platforms may vary over time. The ways in which users utilize a platform (eg, for political discourse vs socializing) will directly affect the data that are abstracted [21]. Researchers will need to keep abreast of such changes in order to effectively collect data. However, the approach demonstrated here provides an innovative means of understanding the unique needs of an understudied population, both cheaply and in a timely manner [22].

XSL•FO
RenderX

Acknowledgments

EAK conceived the topic and analysis plan, conducted analyses, and drafted and edited the manuscript and SDY conceived the topic and analysis plan, advised in manuscript preparation, and revised and critiqued the manuscript.

The project described was supported by a grant from the UCLA Graduate Division, award number T32AG033533 from the National Institute on Aging, award number R24HD041022 from the Eunice Kennedy Shriver National Institute of Child Health and Human Development, award number K01MH090884 from the National Institute of Mental Health, and an award from the University of California Office of the President. The content is solely the responsibility of the authors and does not necessarily represent the official views of the funding agencies.

Conflicts of Interest

None declared.

References

1. Maguen S, Shipherd JC, Harris HN, Welch LP. Prevalence and predictors of disclosure of transgender identity. International Journal of Sexual Health 2007 Feb 08;19(1):3-13. [doi: 10.1300/J514v19n01_02]
2. Rosser BRS, Oakes JM, Bockting WO, Miner M. Capturing the social demographics of hidden sexual minorities: An internet study of the transgender population in the United States. Sex Res Soc Policy 2007 Jun;4(2):50-64. [doi: 10.1525/srsp.2007.4.2.50]
3. Kenagy GP. Transgender health: Findings from two needs assessment studies in Philadelphia. Health Soc Work 2005 Feb;30(1):19-26. [Medline: 15847234]
4. Stotzer RL. Violence against transgender people: A review of United States data. Aggression and Violent Behavior 2009 May;14(3):170-179. [doi: 10.1016/j.avb.2009.01.006]
5. Xavier JM, Bobbin M, Singer B, Budd E. A needs assessment of transgendered people of color living in Washington, DC. International Journal of Transgenderism 2005 Oct 11;8(2-3):31-47. [doi: 10.1300/J485v08n02_04]
6. Grant J, Mottet L, Tanis J, Harrison J, Herman J, Keisling M. Transgender Discrim. Surv. Injustice at Every Turn: A Report of the National Transgender Discrimination Survey. Natl. Transgender Discrim. Surv. Washington; 2011. URL: http://www.thetaskforce.org/static_html/downloads/reports/reports/ntds_full.pdf [accessed 2015-04-24] [WebCite Cache ID 6Y15eTzt0]
7. Grossman AH, D'Augelli AR. Transgender youth: Invisible and vulnerable. Journal of Homosexuality 2006;51(1):111-128. [doi: 10.1300/J082v51n01_06]
8. Garofalo R, Deleon J, Osmer E, Doll M, Harper GW. Overlooked, misunderstood and at-risk: Exploring the lives and HIV risk of ethnic minority male-to-female transgender youth. J Adolesc Health 2006 Mar;38(3):230-236. [doi: 10.1016/j.jadohealth.2005.03.023] [Medline: 16488820]
9. Bockting WO, Miner MH, Swinburne Romine Rebecca E, Hamilton A, Coleman E. Stigma, mental health, and resilience in an online sample of the US transgender population. Am J Public Health 2013 May;103(5):943-951. [doi: 10.2105/AJPH.2013.301241] [Medline: 23488522]
10. Coleman E, Bockting W, Botzer M, Cohen-Kettenis P, DeCuypere G, Feldman J, et al. Standards of care for the health of transsexual, transgender, and gender-nonconforming people, Version 7. International Journal of Transgenderism 2012 Aug;13(4):165-232. [doi: 10.1080/15532739.2011.700873]
11. Mayer KH, Bradford JB, Makadon HJ, Stall R, Goldhammer H, Landers S. Sexual and gender minority health: What we know and what needs to be done. Am J Public Health 2008 Jun;98(6):989-995. [doi: 10.2105/AJPH.2007.127811] [Medline: 18445789]
12. Sperber J, Landers S, Lawrence S. Access to health care for transgendered persons: Results of a needs assessment in Boston. International Journal of Transgenderism 2005 Oct 11;8(2-3):75-91. [doi: 10.1300/J485v08n02_08] [Medline: 23153142]
13. Mallon GP. Knowledge for practice with transgendered persons. Journal of Gay & Lesbian Social Services 2000 May 26;10(3-4):1-18. [doi: 10.1300/J041v10n03_01]
14. Clements-Nolle K, Marx R, Katz M. Attempted suicide among transgender persons: The influence of gender-based discrimination and victimization. J Homosex 2006;51(3):53-69. [doi: 10.1300/J082v51n03_04] [Medline: 17135115]
15. Kenagy GP, Bostwick W. Health and social service needs of transgender people in Chicago. International Journal of Transgenderism 2005 Oct 11;8(2-3):57-66. [doi: 10.1300/J485v08n02_06] [Medline: 15847234]
16. Pew Research Center. 2013 Jun 13. A survey of LGBT Americans: Attitudes, experiences and values in changing times URL: http://www.pewsocialtrends.org/2013/06/13/a-survey-of-lgbt-americans/ [accessed 2015-03-27] [WebCite Cache ID 6XWJ9H63v]
17. Young SD, Jaganath D. Online social networking for HIV education and prevention: A mixed-methods analysis. Sex Transm Dis 2013 Feb;40(2):162-167 [FREE Full text] [doi: 10.1097/OLQ.0b013e318278bd12] [Medline: 23324979]
18. Bruns A, Stieglitz S. Towards more systematic analysis: Metrics for tweeting activities. International Journal of Social Research Methodology 2013 Mar;16(2):91-108. [doi: 10.1080/13645579.2012.756095]

XSL•FO
RenderX

19. Hawksey M. 2013. Twitter archiving Google spreadsheet TAGS internet URL: https://mashe.hawksey.info/2013/02/twitter-archive-tagsv5/ [accessed 2015-03-28] [WebCite Cache ID 6XNCltTrY]
20. Dedoose V5. 0.11, web application for managing, analyzing, and presenting qualitative and mixed method research data Internet. Los Angeles: SocioCultural Research Consultants, LLC; 2014. URL: http://www.dedoose.com/ [accessed 2015-04-24] [WebCite Cache ID 6Y15p1pJ1]
21. Ruths D, Pfeffer J. Social media for large studies of behavior: Large-scale studies of human behavior in social media need to be held to higher methodological standards. Science 2014;346(6213):1063-1064. [doi: 10.1126/science.1257756]
22. McFarlane M, Kachur R, Klausner JD, Roland E, Cohen M. Internet-based health promotion and disease control in the 8 cities: Successes, barriers, and future plans. Sexually Transmitted Diseases 2005;32(supplement):S60-S64. [doi: 10.1097/01.olq.0000180464.77968.e2]

Abbreviations

MSM: men who have sex with men
TAGS: Twitter Archiving Google Spreadsheet
UCLA: University of California, Los Angeles

Edited by G Eysenbach; submitted 05.12.14; peer-reviewed by I Holloway; comments to author 17.02.15; revised version received 02.03.15; accepted 14.03.15; published 07.05.15.

Please cite as:
Krueger EA, Young SD
Twitter: A Novel Tool for Studying the Health and Social Needs of Transgender Communities
JMIR Mental Health 2015;2(2):e16
URL: http://mental.jmir.org/2015/2/e16/
doi:10.2196/mental.4113
PMID:26082941

©Evan A Krueger, Sean D Young. Originally published in JMIR Mental Health (http://mental.jmir.org), 07.05.2015. This is an open-access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on http://mental.jmir.org/, as well as this copyright and license information must be included.

XSL•FO
RenderX

Publisher:
JMIR Publications
130 Queens Quay East.
Toronto, ON, M5A 3Y5
Phone: (+1) 416-583-2040
Email: support@jmir.org

XSL•FO
RenderX