*Exploring the emotional well-being and mental health of individuals who play fantasy football, 2021 – methodology document*

Participants: 2,026 individuals began the questionnaire, with 31 individuals subsequently removed for either reporting their age to be under 18 (as required by the study criteria) or for not completing at least 75% of the questionnaire. Thus, data from 1,995 participants were included for analysis in the study. Of the 1,995 participants, 1,914 (95.9%) were male and 71 (3.6%) were female, with 10 (0.5%) opting not to say. Participants ranged in age from 18 to 87, with a mean age of 33.07 (SD = 10.18). Whilst the majority were British (53.0%), 96 different nationalities were represented in total. All participants reported that they currently played FF, with the average years of experience in FF being 7.46 (SD = 4.94) and the average number of leagues currently involved in being 6.13 (SD = 6.02). Ethical approval was obtained from the first author’s institution.

Procedure: the online questionnaire was advertised through two mediums: Twitter (www.Twitter.com) and the well-known FF website, ‘Fantasy Football Scout’ (Fantasy Football Scout, 2021). Prospective participants were able to read a brief synopsis of the study through these adverts and could access the questionnaire via an accompanying link. Once clicked, participants were presented with an information sheet detailing the study in full, and then a consent statement. Participants gave their consent by selecting ‘yes’ to the statement: ‘I agree to participate in the present study’. Only by selecting ‘yes’ could the individual proceed to the main part of the questionnaire. In accordance with ethical guidelines, participants were informed that they did not have to answer a question if they did not wish to. As such, there were a number of instances of missing responses. However, even the question with the most missed responses (“On average, how many minutes per day do you spend thinking about FF?”) still had an extremely high response rate of 97.64%. The whole questionnaire took between 15 and 20 minutes to complete.

Measures: given that the present study is (to the authors’ knowledge) the first of its kind exploring the emotional experiences of individuals who play FF, it was felt necessary to use pre-existing questionnaires with proven validity and reliability and to adapt them such that they address FF, as opposed to creating a custom-made questionnaire. Adaptations were made by the first and second authors, who were both experienced FF players, before being corroborated by the third and fourth authors, who each have expertise in clinical psychology. The questionnaire was then pilot tested by independent individuals to ascertain whether all question and response items were easily understandable, and to ensure that the full questionnaire was of a suitable length. In total, the questionnaire consisted of six sections aiming to investigate both emotional experiences and mental health, of which the present paper will only report the findings from the former. The results from the remaining sections will be presented and discussed in a subsequent paper due to the large number of significant findings. The full questionnaire can be found in Appendix C.

1) Demographic Information.

The demographic information obtained was age, sex, nationality, current country of residence, and ethnicity.

2) FF Experience and Behaviour. Five questions were designed to measure experience and behaviour in FF. Specifically, individuals were asked to report: a) how many seasons they had played FF for (“FF Experience”); b) how many FF leagues they were currently participating in (“FF Leagues”); c) the average number of minutes they spent per day on FF sites/apps (“Time Playing”); d) the average number of minutes they spent per day on other FF-related activities such as listening to podcasts or browsing specific social media (“Time Researching”); and, e) the average number of minutes they spent per day thinking about FF (“Time Thinking”).

Given the lack of existing research to use as a guideline, questions required free text responses from participants as opposed to selecting from various discrete categories (e.g. “0-29 minutes”, “30-59 minutes”, etc.). Post-hoc categorisation of this data then allowed for participants to be assigned to logical and approximately equal groups for each FF experience/behaviour measure. For responses where the participant gave a lower or upper boundary (e.g. “10+” or “less than 5”), the number stated was used. For responses where participant gave a range rather than a single number (e.g. “60-90 minutes”) the mid-point of the response was used. Finally, for responses in which a reasonable number was not decipherable (e.g. the participant reported “too much” or “several times”) then the answer was deleted, and the response treated as missing data and not included in the statistical analysis. In an attempt to maintain consistency, all post-hoc categorisation and decisions regarding deletion were confirmed by the second author.

3) Multidimensional Emotion Questionnaire (MEQ). The MEQ was used to measure 10 emotions (five positive and five negative) across four subscales (frequency, intensity, duration, and regulation) in relation to FF. Specifically, participants were asked to rate how often they experience the stated emotion, how intense the emotion typically is when it occurs, how long-lasting the emotion typically is when it occurs, and how well they can regulate the emotion when it occurs. Separate five-point Likert scales were used for each subscale (e.g. ‘about once each month’ to ‘more than three times each day’ for frequency; ‘very low’ to ‘very high’ for intensity; ‘less than one minute’ to ‘longer than four hours’ for duration; and ‘very easy’ to ‘very difficult’ for regulation). Participants were reminded at the start of this section that all answers should be considered in relation to FF. The MEQ has been demonstrated to have adequate to excellent internal reliability (α = 0.61 to 0.85), strong test-retest reliability (r = .66 to .83), and strong validity with the PANAS-X measure of emotional affect (r = .69 to .72) (Klonsky et al., 2019).

4) Patient Health Questionnaire Depression Scale (PHQ-9) The PHQ-9 was used to measure low mood due to FF. The PHQ-9 has been validated across a wide range of ages and settings (Kroenke, Spitzer, & Williams, 2001) and is the most frequently used measure of depression/low mood globally (Wang et al., 2021). It asks individuals to report how often they have been bothered by various problems over the last two weeks using a four-point Likert scale ranging from 0 (‘not at all’) to 3 (‘nearly every day’). Thus, higher scores indicate greater severity of low mood. To ensure relevance to FF, the wording of the questions was amended. For example, in the original PHQ-9, the individual is asked: “Over the last two weeks, how often have you been bothered by any of the following problems?”, with the subsequent first item being: “little interest or pleasure in doing things”. Instead, in the adapted version, the individual is asked: “In the last two weeks, how often has fantasy football left you feeling that you have little interest or pleasure in doing other things?” Additionally, the traditional version contains nine items, however, it was felt that two (“Moving or speaking so slowly that other people could have noticed? Or the opposite – being so fidgety or restless that you have been moving around a lot more than usual?” and “Thoughts that you would be better off dead, or of hurting yourself in some way?”) were not appropriate, and therefore were removed from the questionnaire.

5) Generalised Anxiety Disorder Questionnaire (GAD-7) The GAD-7 was used to measure feelings of anxiety due to FF. The GAD-7 has been found to have good validity and reliability across a range of populations and settings (Löwe et al., 2008; Spitzer et al., 2006). Similar to the PHQ-9, individuals are asked to report how often they have been bothered by various problems over the last two weeks using a four-point Likert scale ranging 0 (‘not at all’) to 3 (‘nearly every day’), with higher scores reflecting greater severity of anxiety. Again, the wording of questions was amended to ensure relevance to FF. For example, in the original GAD-7, the individual is asked: “Over the last two weeks, how often have you been bothered by any of the following problems?”, with the subsequent first item being: “feeling nervous, anxious, or on edge”. Instead, in the adapted version, the individual is asked: “In the last two weeks, how often has fantasy football left you feeling nervous, anxious or on edge?” There are seven items in the GAD-7, and all were used in the current study.

6) Work and Social Adjustment Scale (WSAS) The WSAS was used to measure everyday functional impairment due to FF. Developed by Marks (1986), the WSAS has been shown to have high reliability and validity (Cella et al., 2011; Mundt et al., 2002; Zahra et al., 2014). It contains five items that address an individual’s ability to function with regard to work activities, home management, social leisure activities, private leisure activities, and relationships. Participants respond using a nine-point Likert scale ranging from 0 (‘not at all’) to 8 (‘very severely’), with higher scores indicating greater everyday functional impairment. As before, the wording of questions was amended to ensure relevance to FF. For instance, rather than asking: “Because of my [problem] my ability to work is impaired”, the current study asked: “Because of fantasy football my ability to work is impaired”.

7) Problematic Online Gaming Questionnaire Short-Form (POGQ-S). The POGQ-S was used to measure problematic FF behaviour across six dimensions: preoccupation, immersion, withdrawal, overuse, interpersonal conflicts, and social isolation. It was developed by Demetrovics and colleagues (2012) for use in the online gaming domain and the short-form used in the current study has been found by Pápay et al. (2013) to have satisfactory psychometric properties. Participants are required to respond to 12 items using a five-point Likert scale ranging from 1 (‘never’) to 5 (‘always’), with higher scores reflecting more problematic behaviours. Again, the wording of questions was amended so that they were relevant to FF. For example, in place of the traditional POGQ-S item which asks: “How often do you daydream about gaming?”, the question in the current study asked: “How often do you daydream about FF?”. One question was also added to the POGQ-S. In addition to asking: “How often do you get restless or irritable if you are unable to play Fantasy Football for a few days?”, it was thought that it would be necessary to also ask this question in relation to being unable to play for “over a week”. Finally, participants were notified that for all questions, the term “playing” FF referred to any time spent on a FF website/app, as well as other activities such as reading and listening to podcasts that have the primary aim of benefiting FF performance.