Social Networks and Health

Robert Eyre[a], Frances Griffiths[b], Thomas House[a]

[a] *Centre for Complexity Science, University of Warwick*
[b] *Warwick Medical School, University of Warwick*

Abstract

Depression forms a major contribution to the global disease burden, playing a debilitating part in the lives of millions of people across the world. It affects a significant percentage of adolescents. Depression is characterised by a set of symptoms affected by social networks. Previous work, utilising the Add Health data set, used a binary model of "depressed" versus "not depressed", and demonstrated that being "not depressed" can transmit over a friendship network. We aim to consider greater layers of complexity to the transmission of depressive symptoms. We do this by applying two changes to the previous work. First, we consider whether people change in their level of depression at all, and in what direction, i.e. getting worse or getting better over time. Second, we consider individual symptoms. We use data from 2194 young people aged 12-19 from the Add Health dataset. We examine changes in symptom levels and the dependency of the probabilities of worsening and improving of symptoms on the number of better off and worse off friends. We achieve this with empirical data analysis and parametric inference. We find that, for almost all symptoms, having more worse off friends makes it more likely for an individual to get worse, and less likely for them to improve, and vice versa for better off friends. This suggests that whether having more friends will make you emotionally healthier is dependent on the emotional state of the friends. We also find that individuals are more likely to change in symptoms than not at all, and no bias exists towards improving or worsening. Evidence also suggests that the change in symptom level follows an exponential distribution. This suggests that the change in symptom levels between time points occurs in one direction only, i.e. people keep getting worse or keep getting better between time points, instead of following a random walk. Therefore, the effects of the friendship network happen progressively.

1. Introduction

Depression is a major part of the global disease burden, affecting approximately 350 million people wordwide [1]. It can prevent people from functioning properly both socially and at work, and can at its worst lead to suicide, which is responsible for an estimated 1 million deaths per year [1]. There is worldwide concern about mental health of young people [2]. In fact, the Office of National Statistics Child and Adolescent Mental Health Survey found that in 2004 1.4% of 11-16 year olds in the UK were seriously depressed [3]. In the United States, the Substance Abuse and Mental Health Services Administration's National Survey on Drug Use and Health found that between 2004 and 2008 7.9-9.0% of 12-17 year olds suffered from depression [4]. Clearly, depression within young people is an issue which very much needs to be explored.

In recent years, there has been much interest in the connections between health and health behaviours of individuals, and the social networks (e.g. networks of friends or colleagues) to which these individuals belong. Previous work has looked into the spreading of disease epidemics over networks [5], the formation of patterns in friendship and sexual networks [6] [7], and the spreading of health behaviours such as smoking and obesity over social networks [8] [9] [10].

Depression and depressive symptoms are another set of health behaviours which may spread over social networks, and recent work suggests they do. Research has shown that mood and depressive symptoms are contagious [11] [12], and that greater social support and friendship can help mitigate depressive symptoms [13] [14] [15] [16]. Tools from epidemiology, such as the Susceptible-Infected-Susceptible model, have been used to model the spread of depression, and have provided formal evidence that both having less severe and more severe depressive symptoms can spread like an infectious disease [17].

More recent work, by Hill et. al. [18], has modelled the possession of severe depressive symptoms as a binary variable where the individual is either in a depressed state or a non-depressed state. Using data from the National Longitudinal Study of Adolescent Health (Add Health [19]), they have shown that, in this case, the non-depressed state, but not the depressed one, spreads over social networks.

We build on this previous work by introducing new levels of complexity in order to more accurately reflect the difficult nature of being able to declare an individual as being depressed or not. First, by considering various symptoms of depression individually, as well as a complete set as has been previously done, as these individual symptoms represent health behaviours which may be more easily observed. Secondly, by not restricting ourselves to a binary model where we observe whether individuals cross some artificial threshold of being depressed or not, and instead simply observing whether there is a positive or negative change in the level of depressive symptoms, or no change at all. We seek to answer the question of whether the worsening or improving of these individual symptoms spreads over social networks, and achieve this using data analysis and inference. We hope to provide even greater grounding for future research in this area.

2. Method

2.1. Add Health

The Add Health data set comes from a longitudinal study designed to examine the health and health behaviours of a nationally representative sample of adolescents in the United States [19]. The individuals questioned were all school students in grades 7-12 at the start of the study, which followed them through to adulthood in a series of "waves". We used data from the in-home interviews of wave 1, performed in 1994-95, and wave 2, performed in 1996, during which time the respondents were 12-19 years old. Wave 1 included 20745 students, and wave 2 included 14738. The main cause for the decrease in numbers was that the majority of 12th graders were removed from wave 2 due to exceeding the grade allowance in 1996.

The in-home interviews used in the survey contained over 2000 questions covering many diverse health behaviours. Though dishonest answers and errors surely exist within the data set, cautionary measures were taken by the interview administrators to keep the data secure and avoid interviewer and parental influence.

132 schools were selected, 80 of which were high schools. If a high school did not include all grades 7-12, then they were asked to identify feeder schools with a 7th grade, who sent at least 5 students to the high school. One feeder school for each high school was chosen for the survey with a probability proportional to the number of students contributed to the high school, and was labelled the sister school for that high school.

In most schools, only a selection of students were questioned, chosen in order to provide a nationally representative sample of areas such as ethnicity, disability, and gender. Only in 16 schools, labelled saturated schools, were all students interviewed and asked to provide friendship data, in order to allow the construction of social networks.

2.2. Friendship Data

Respondents in Add Health were all required to provide a list of friends. They did this by identifying the friend in the roster lists of their school and their school's sister school, and writing down the friend's AID (identity number) in a list. If the friend was a member of their school but could not be found on the roster, they were given a generic AID of 99999999. If they went to the sister school, but could not be found on the roster, they were given the generic AID 88888888. If they attended neither, they were given the generic AID 77777777.

In wave 1, 7106 respondents were allowed to list up to 5 male and up to 5 female friends, while the rest were only allowed up to 1 male and up to 1 female friend. Of these respodents, 3099 were from the saturated schools. In wave 2, 2729 of the respondents from the saturated schools were allowed to list up to 5 male and up to 5 female friends. The rest of the saturated school respondents, and all respondents not from saturated schools, were asked to list only up to 1 male and up to 1 female friend.

2.3. CES-D Scale

The Centre of Epidemiology Depression scale is a self-assessed questionnaire designed to give a numerical outcome measuring the severity of depressive symptoms [20]. It has been shown to be reliable for a period of up to 12 months [21].

As shown in the original article by Radloff [20], the questionnaire is composed of 20 questions, 18 of which were asked to respondents of Add Health. Questions 11 and 17 were not included. Respondents were asked to give an answer of 0-3, rating the intensity of the corresponding feeling over the past week. They were also allowed to give responses of “refused”, “don’t know”, and “not applicable”. If a respondent gave numerical answers to all questions, then they acquired a numerical score in 0-54, summing all the individual question scores (the scores from questions 4, 8, 12, and 16 were inverted, i.e. 3 minus the score, when summing as they represent positive emotions). The higher the score, the more severe the depressive symptoms.

In order to examine individual symptoms, we assigned each question to one symptom. Questions 12 and 16 contributed to anhedonia (with total possible score in 0-6). Question 2 to poor appetite (score 0-3). Question 5 to poor concentration (score 0-3). Questions 3, 6, 14, and 18 to dysphoria (score 0-12). Questions 1, 8, 10, 13, and 15 to helplessness (score 0-15). Questions 7 and 20 to tiredness (score 0-6). Questions 4, 9, and 19 to worthlessness (score 0-9).

2.4. Sample Dataset

Certain restrictions were used when choosing individuals from the complete data set to use in our analysis. These restrictions are shown as a flow diagram in figure 1.

First, from each wave, we chose individuals who were from saturated schools, were allowed to list up to 5 male and up to 5 female friends, and who gave numerical answers (i.e. not “refused”, “don’t know”, or “not applicable”) to all the CES-D questions. By only considering individuals from saturated schools, we avoid biases introduced by the weighted sampling of different minority groups done in order to make the study nationally representative. By only considering those who were allowed to list the maximum number of friends, we can create a full friendship network. Also, restricting to those who answered all CES-D questions with numerical answers allows us to give a score measuring the severity of depressive symptoms to each individual.

Then, in order to be able to look at the change of symptom levels over time, we only considered individuals who appeared and satisfied all the above criteria in both waves. This then left us with a dataset of 2194 individuals.

Lastly, when building the friendship network, friends with generic AIDs (identity numbers, see section 2.2) were ignored. This prevents the generic AIDs from becoming hub nodes and possibly affecting the structure of the network.

2.5. Empirical Data Analysis

It is informative to consider the empirical changes in score between the two waves. This can show what amount of change in score is most common, whether positive or negative changes in score are preferred, how these changes possibly occurred between the two time points, and if the number of friends has any effect on whether negative or positive changes occur.

We achieved this by examining first the distribution of the score changes. If the score changes are normally distributed, then this might indicate that between the two time points individuals change score repeatedly in different directions like a random walk. An exponential distribution would show that this is not true, and might indicate that individuals change score only positively or negatively. We used the `fitdistr()` function in R to find possible parameter values for normal and exponential distributions for the score changes, and compared qq (quantile-quantile) plots of the results.

We next explored a type of density plot of the wave 1 scores against the wave 2 scores, plotted as a grid of cells where each cell represents a combination of a wave 1 score and a wave 2 score and individuals are represented in the cell by points distributed uniformally across the cell. By colouring the points by number of friends, number of friends with higher scores, and number of friends with lower scores, we can see if any patterns emerge to imply that these variables have an effect.

These plots were performed for individual symptoms, as well as the total CES-D score, even though, due to the small changes in score allowed, not as much information can be gained from the symptom plots.

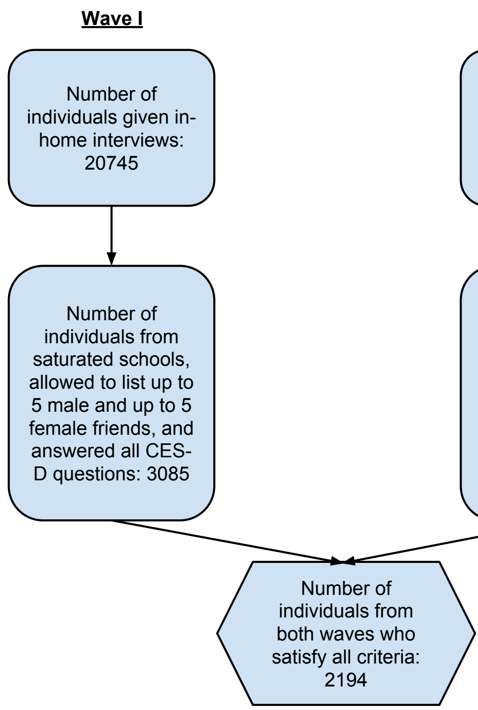


Figure 1: Sampling flow diagram for our data sample. At the top are all individuals given in-home interviews in both waves 1 and 2. At the bottom, in the hexagonal box, is the data we used.

2.6. Statistical Modelling

Most previous work made various assumptions that we do not. In looking at whether depression spreads over a network, they assumed that being not depressed (having a low score) does not, and they used static network measures to tell if transmission does occur [22] [17]. More recent work has also ignored these assumptions, but have still kept things relatively simple. They have treated depression as a binary variable, where if an individual is above a threshold CES-D score they are depressed, and if they are below it they are not. Instead of using static measures, we used different time points to confirm transmission, and instead of looking at whether individuals cross a score threshold between these time points, we considered whether they change in score at all and whether that change is positive or negative. As mentioned, unlike any other research before, we also did this for individual symptoms.

We used statistical modelling to examine the dependence (or lack of) of the likelihood of an individual worsening (increasing in score) and improving (decreasing in score) in the severity of their depressive symptoms on the number of friends they have who have a worse case (higher score) or better case (lower score) of those symptoms.

Given a score for individual i at time t of $X_i(t)$ for some symptom, we examined models for the probabilities $p = Pr\left[X_i(t+1) > X_i(t)\right]$, i.e. the probability of worsening, and $q = Pr\left[X_i(t+1) < X_i(t)\right]$, i.e. the probability of improving. For each model we compared different possible forms, namely transmission (dependence on higher or lower scoring friends) and no transmission (independence on friends). Following the literature, we used an S-shaped (sigmoidal) dependency for the transmission forms [23] [24]. This was done for all symptoms, as well as the total CES-D score.

2.6.1. The Models

The models examined included three possible states. Model 1 has a probability p of worsening, a probability q of improving, and a probability $1 - p - q$ of not changing, and examines the dependency of p and q on higher scoring friends. Model 2 is the same, except it examines the dependency of p and q on lower scoring friends. Models with two states were also examined (see Appendix A).

For each model, four forms were examined: both worsening and improving trasmitting, neither transmitting, only worsening transmitting, and only improving transmitting. The both transmitting form was

$$p_k = \alpha + \beta \sum_{l=0}^{k} \binom{10}{l} \gamma^l \left(1-\gamma\right)^{10-l} \tag{1}$$

$$q_k = \delta + \epsilon \sum_{l=0}^{k} \binom{10}{l} \zeta^l \left(1-\zeta\right)^{10-l} \tag{2}$$

where k is the number of higher scoring friends for model 1, and is the number of lower scoring friends for model 2. The neither transmitting form was

$$p_k = \alpha \tag{3}$$

$$q_k = \delta \quad . \tag{4}$$

The worsening transmitting only form was

$$p_k = \alpha + \beta \sum_{l=0}^{k} \binom{10}{l} \gamma^l \left(1-\gamma\right)^{10-l} \tag{5}$$

$$q_k = \delta \quad . \tag{6}$$

The improving transmitting only form was

$$p_k = \alpha \tag{7}$$

$$q_k = \delta + \epsilon \sum_{l=0}^{k} \binom{10}{l} \zeta^l \left(1-\zeta\right)^{10-l} \quad . \tag{8}$$

2.6.2. Parametric Inference

Maximum likelihood extimates were found for the parameters α, β, γ, δ, ϵ, and ζ for each model, for each symptom.

The likelihood function for model 1 was

$$L\left(\mathbf{x}, \mathbf{y} | \mathbf{p}, \mathbf{q}, \mathbf{M}\right) = \prod_k \binom{M_k}{x_k, y_k, M_k - x_k - y_k} p_k^{x_k} q_k^{y_k} \left(1 - p_k - q_k\right)^{M_k - x_k - y_k} \tag{9}$$

where x_k was the number of individuals with k higher scoring friends who worsened, y_k was the number of individuals with k higher scoring friends who improved, and M_k was the total number of individuals with k higher scoring friends.

The likelihood function for model 2 was

$$L\left(\mathbf{u}, \mathbf{v} | \mathbf{p}, \mathbf{q}, \mathbf{N}\right) = \prod_k \binom{N_k}{u_k, v_k, N_k - u_k - v_k} p_k^{u_k} q_k^{v_k} \left(1 - p_k - q_k\right)^{N_k - u_k - v_k} \tag{10}$$

where u_k was the number of individuals with k lower scoring friends who worsened, v_k was the number of individuals with k lower scoring friends who improved, and N_k was the total number of individuals with k lower scoring friends.

The MLEs were found by minimising the negative log-likelihood using functions in R. For models 1 and 2, the `nmk()` function from the `dfoptim` package was used, chosen over the standard function due to its ability to handle non smooth functions. In this case, the function was evaluated from five different random initial values, and, where results disagreed, the result with the lowest minimal value was chosen.

The different forms for each model were compared by their AIC (Akaike Information Criterion) values (see Appendix B).

3. Results

3.1. Distribution of changes in symptom levels

We considered the distribution of the change in total CES-D score between waves 1 and 2 for our dataset. Figure 2a shows the empirical distribution of the score changes, which appears very sharply peaked at 0. Only less than 10% of individuals did not change in score at all, so we find that individuals are more likely to change than not. Also, almost all members of the dataset had a change in score of no greater than ± 20. Together with how steeply the distribution rises towards 0, this implies that small changes in score occur much more, and large changes are incredibly rare. We also note that the distribution appears symmetrical, so that neither worsening nor improving are preferred over the other.

The shape of the distribution implies that, though the data is technically discrete, it may follow a two-sided exponential distribution, rather than a normal distribution. This would then suggest that, between the two waves, each individual changed score in only one direction, i.e. they got progressively worse or better over the course of time. If it was a normal distribution, then this would suggest a random walk in score changes, where individuals got better and worse several times at random between the two time points. The log-linear plot of the empirical distribution, shown in figure 2b, is difficult to interpret, but arguably shows the data falling on two straight lines, consistent with an exponential distribution.

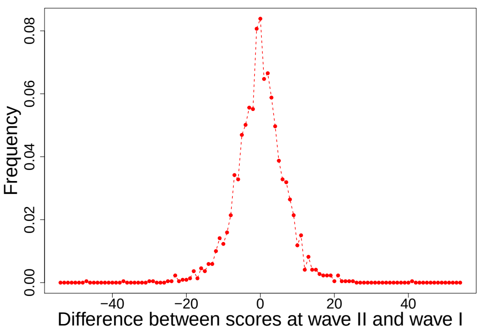
We used the R function `fitdistr()` to fit the parameters of the distributions from the data. The exponential distribution of positive score changes had a rate parameter of 0.2205. For the negative score changes, the rate parameter was 0.2159. The normal distribution of score changes had a mean of -0.0975, and a standard deviation of 6.7612. Quantile-quantile (qq) plots were made of the data against these fitted distributions, as shown in figure 2. Though close to 0, both appear to give a good fit, at the tails the normal distribution diverges away more, supporting the conclusion that the score changes follow an exponential distribution. Using `fitdistr()` again to give the log likelihoods of the distributions with the fitted parameters, we can calculate the AIC values (see Appendix B). For the exponential distribution of positive changes, the AIC is 5920. For the exponential distribution of negative changes, it is 6081. For the normal distribution of changes, it is 14604. The exponential distributions have smaller AIC values than the normal distribution, and are therefore preferred.

Figure 3 shows the empirical distributions for the individual symptoms. Due to the much smaller range in possible score changes, these are not as informative, and comparisons to continuous distributions are pointless. However, we can see that these distributions share characteristics with that of the overall CES-D score. Small changes in score are favoured, though individuals are more likely to experience some change than none at all. There appears to be little to no bias to either improving or worsening. All the distributions appear quite sharply peaked at 0, again implying a possible exponential distribution. However, certain symptoms, such as concentration, tiredness, and helplessness, do not appear to rise to 0 so steeply or be peaked quite as sharply, possibly implying they follow exponential distributions of very different rate parameters from the overall CES-D score.

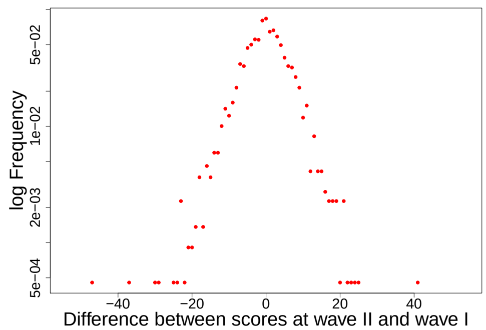
3.2. Patterns in changes in symptom levels

Still considering the score changes, we plotted the wave 1 scores against the wave 2 scores as a grid of cells, each cell corresponding to a particular wave 1 and wave 2 score combination. Points randomly distributed across the cell represent the number of individuals in that cell. The results of this for the total CES-D score can be seen in figure 4a. The points form one cluster centred at combinations of low wave 1 and wave 2 scores, so most individuals have and continue to have low depressive symptom scores. This is unsurprising, as the dataset is taken from a representative, and therefore largely healthy, group of people. It is also noteworthy that the points are very strongly clustered along the central diagonal going from the (0,0) cell up to the (54,54) cell, which indicates that small changes in score occur most often. Large changes in score are clearly quite rare. This supports the results found in section 3.1. Also, again like in the previous section, there does not appear to be any bias towards improving or worsening.

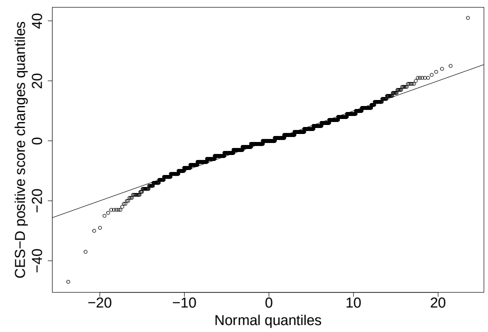
The remaining subfigures in figure 4 are coloured by number of friends (figure 4b), number of worse off (higher scoring) friends (figure 4c), and number of better off (lower scoring) friends (figure 4d). There is no discernible pattern from the number of friends irrespective of score. However, in both of the figures concerned with friends of either higher or lower scores, a pattern does emerge. Individuals with more higher scoring friends tend to fall below the diagonal, so that they have got worse between the two time points.



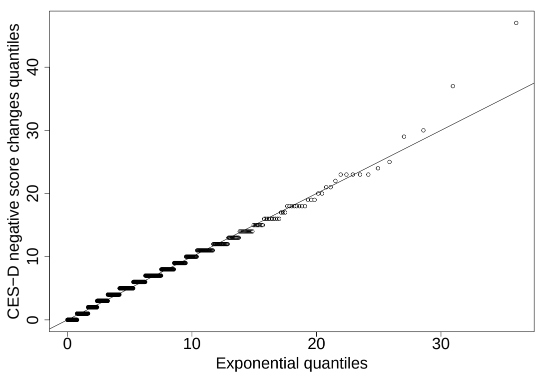
(a) Empirical distribution of score changes.



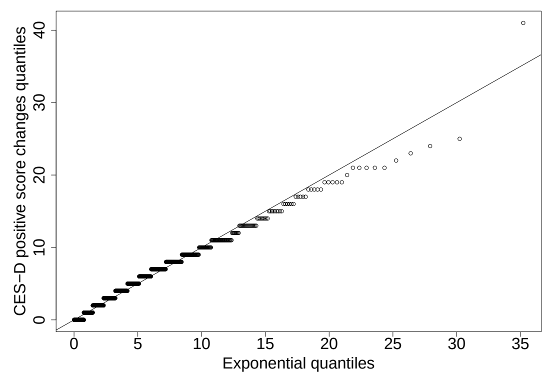
(b) Log plot of the empirical distribution of score changes.



(c) QQ plot of the empirical distribution of score changes against the fitted normal distribution of mean -0.0975 and standard deviation 6.7612.

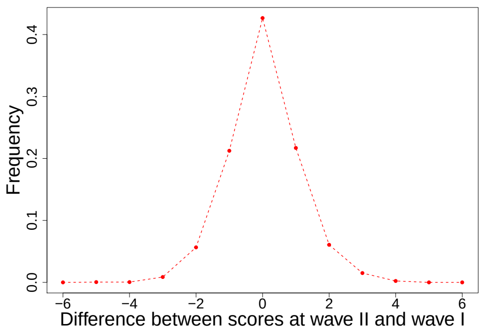


(d) QQ plot of the empirical distribution of negative score changes against the fitted exponential distribution of rate 0.2159.

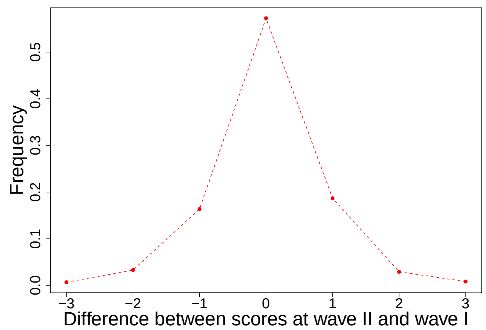


(e) QQ plot of the empirical distribution of positive score changes against the fitted exponential distribution of rate 0.2205.

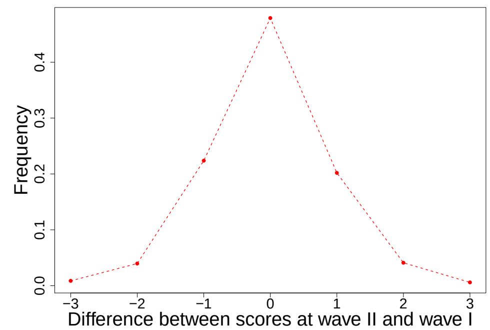
Figure 2: Distribution of change in total CES-D score between waves 1 and 2.



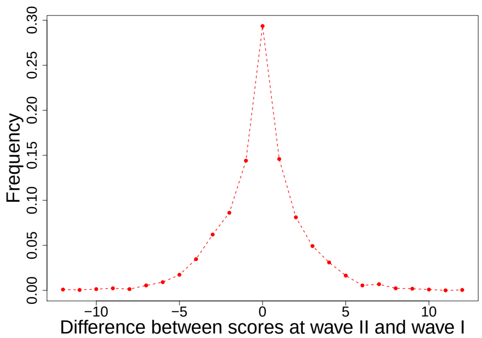
(a) Anhedonia.



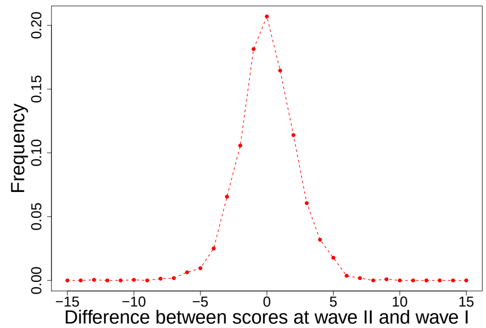
(b) Appetite.



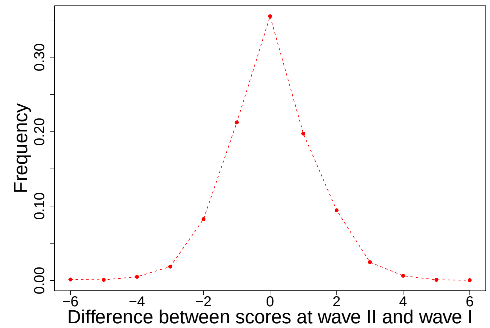
(c) Concentration.



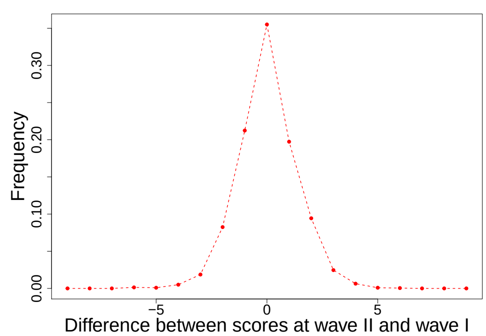
(d) Dysphoria



(e) Helplessness.



(f) Tiredness.

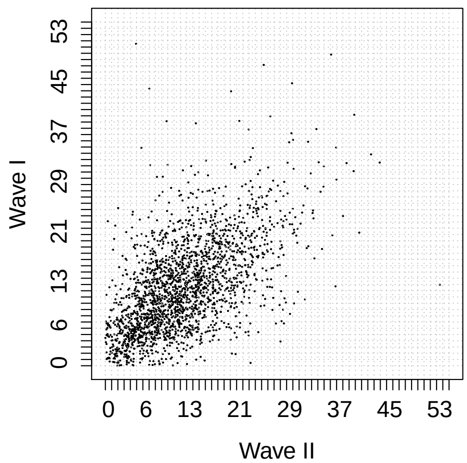


(g) Worthlessness.

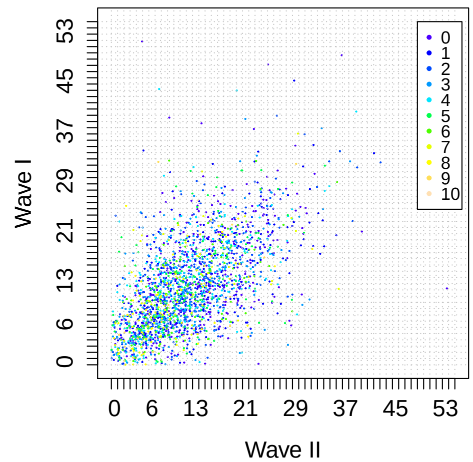
Figure 3: Empirical distributions of changes in score between waves 1 and 2 for individual symptoms.

Those with more lower scoring friends tend to fall above the diagonal, so that they have improved between the two time points. This supports the natural hypothesis that having more worse off friends will make an individual more likely to worsen, and having more better off friends will make them more likely to improve.

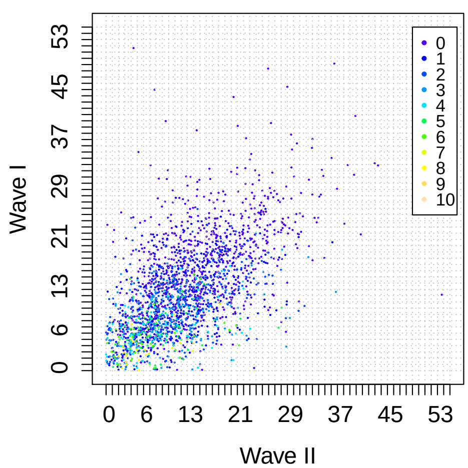
Figure 5 shows similar plots made for the individual symptoms. They each show very similar results to those for the total CES-D score. Most are clustered at the lower score combinations. Especially helplessness, which corresponds to the greatest number of CES-D questions and therefore gives the greatest contribution to the total score. The others have much more spread out points, such that a greater proportion of individuals possess higher scores and experience greater changes in score. However, it must be noted that some of these symptoms do not have a very great possible range of scores, such as appetite and concentration. Of interest is the fact that anhedonia clusters at score combinations of 3 and 4, just higher than the middle values. This implies that young people experience anhedonia more than the other symptoms and maintain this over time. All symptoms seem to focus along the central diagonal, so small changes in score are preferred, and do not show bias to either worsening or improving, like with the total CES-D score.



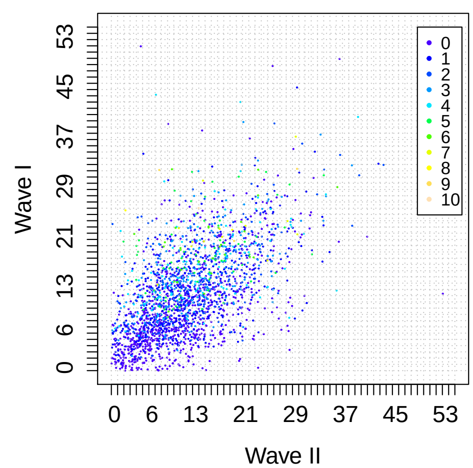
(a) Standard density plot.



(b) Individual points coloured by number of friends.

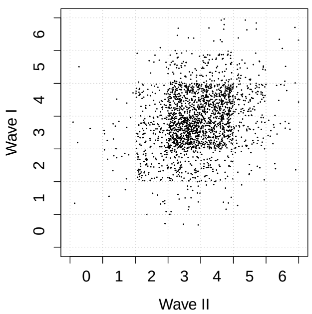


(c) Individual points coloured by number of worse off (higher scoring) friends.

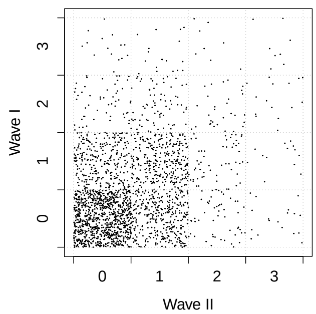


(d) Individual points coloured by number of better off (lower scoring) friends.

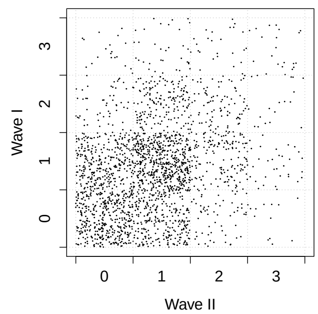
Figure 4: Density plots of wave 1 and 2 score combinations for total CES-D score. The greater the amount with a particular score combination, the more points appear in the corresponding cell.



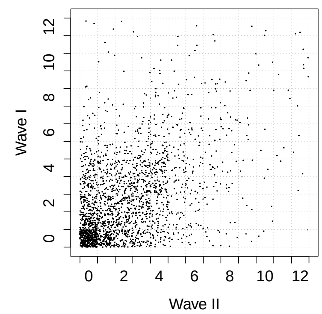
(a) Anhedonia.



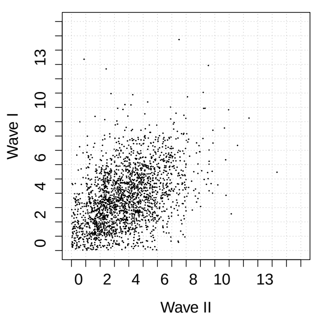
(b) Appetite.



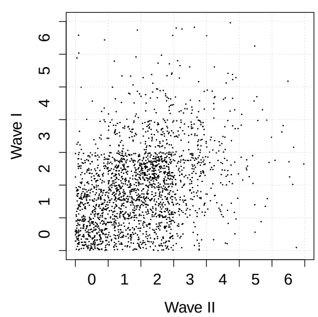
(c) Concentration.



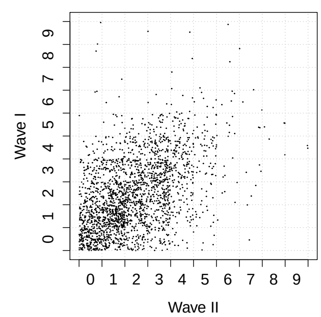
(d) Dysphoria



(e) Helplessness.



(f) Tiredness.



(g) Worthlessness.

Figure 5: Density plots of wave 1 and 2 score combinations for individual symptom scores. The greater the amount with a particular score combination, the more points appear in the corresponding cell.

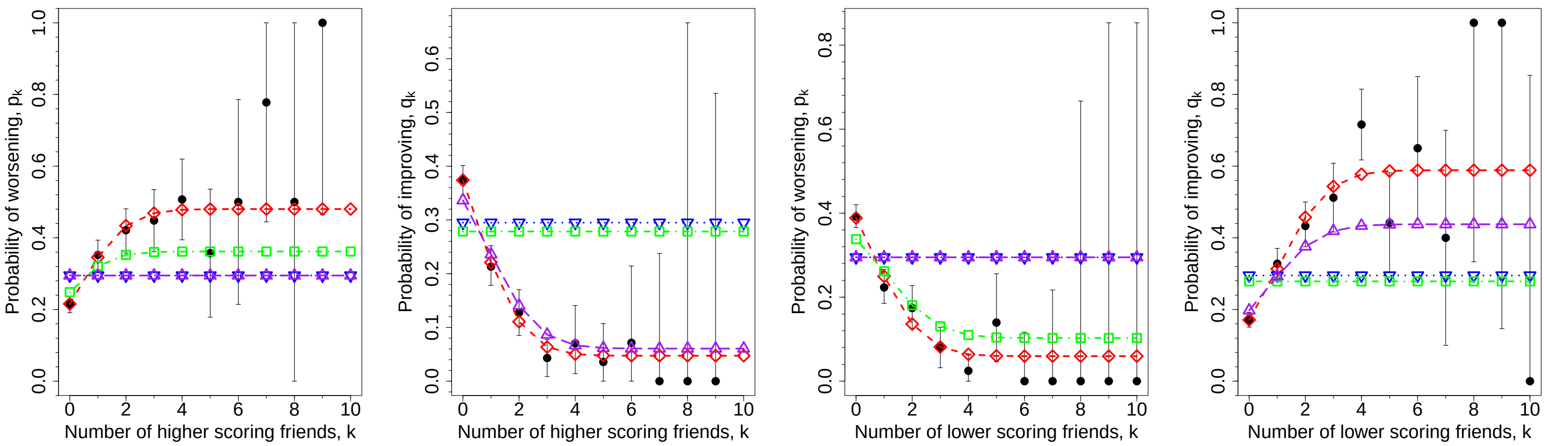
3.3. Spread of worsening and improving of symptoms

Maximum likelihood estimates were found for the parameters of the various forms of the models discussed in section 2.6.1. The resulting forms of p_k and q_k for all the different symptoms, as well as the total CES-D score, are shown in figures 6, 7, and 8. The observed frequencies of individuals improving and worsening with certain numbers of higher and lower scoring friends are shown as black points. In order to compare the models to the observed frequencies, 95% confidence intervals were included as error bars about the frequencies. These were found by bootstrapping, using the `boot` package in R. In the few cases where bootstrapping failed to produce a confidence interval, a 95% Jeffreys interval was used (see Appendix C). For higher numbers of friends, the error bars are very large, due to the very small amount of individuals with that number of friends. The shapes of the probabilities are therefore mostly informed by the individuals with 7 or less higher or lower scoring friends.

The various forms for models 1 and 2 were compared using their AIC (see Appendix B). It was found that, for all symptoms except appetite and for the total CES-D score, both improving and worsening being dependent on both the number of higher and the number of lower scoring friends (shown by the red dashed line) was preferred over the other possibilities. This means that, the more worse off friends an individual has, the more likely they are to get worse, and the less likely they are to get better, and vice versa for better off friends. The fact that this is true for 6 of the 7 symptoms, including ones that contribute little to the overall score, lends support to the validity of the CES-D scale in properly measuring the severity of these symptoms.

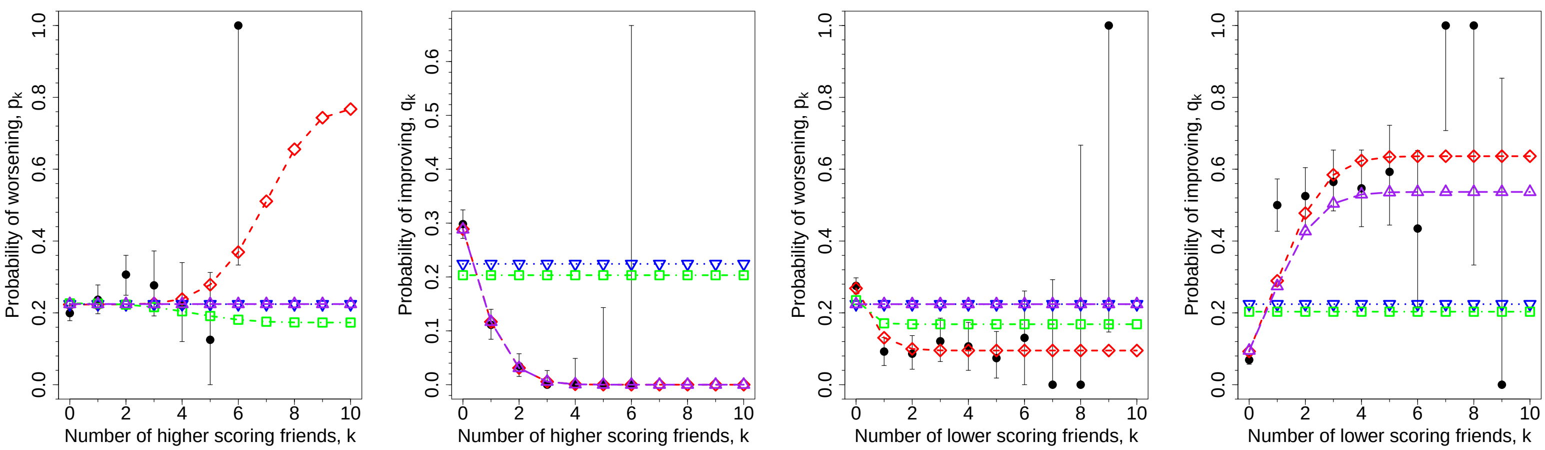
The only symptom to not follow this trend was appetite. For model 1, the form where improving is dependent and worsening is independent on the number of higher scoring friends (the purple dashed line) was preferred, as shown in figure 6c. The observed frequencies certainly support this result. It is not a surprise that this is not reflected in the total score, as appetite contributed little to it, which may be a reason for the discrepency. The allowed score range is very small. However, as this does not affect concentration, there may be deeper differences between appetite and the other symptoms.

One final interesting point is in the shape of the transmission forms of the probabilities. For all preferred cases, the probability begins increasing or decreasing at 0 friends, and changes by a magnitude of 0.1-0.3, before plateauing at a value of around 4 friends. Therefore, having an increasing number of worse off or better off friends does make a difference, but there comes an amount of around 4 where having any more no longer has an effect. We note that, for higher numbers of friends, some of the points do appear much further away from the models. Though the error bars are large enough to not be concerned by this, it is possible that a more complicated shape for the probabilities could fit better, though it would have to be shown as physically justifiable to avoid adding unnecessary complexity to the model.



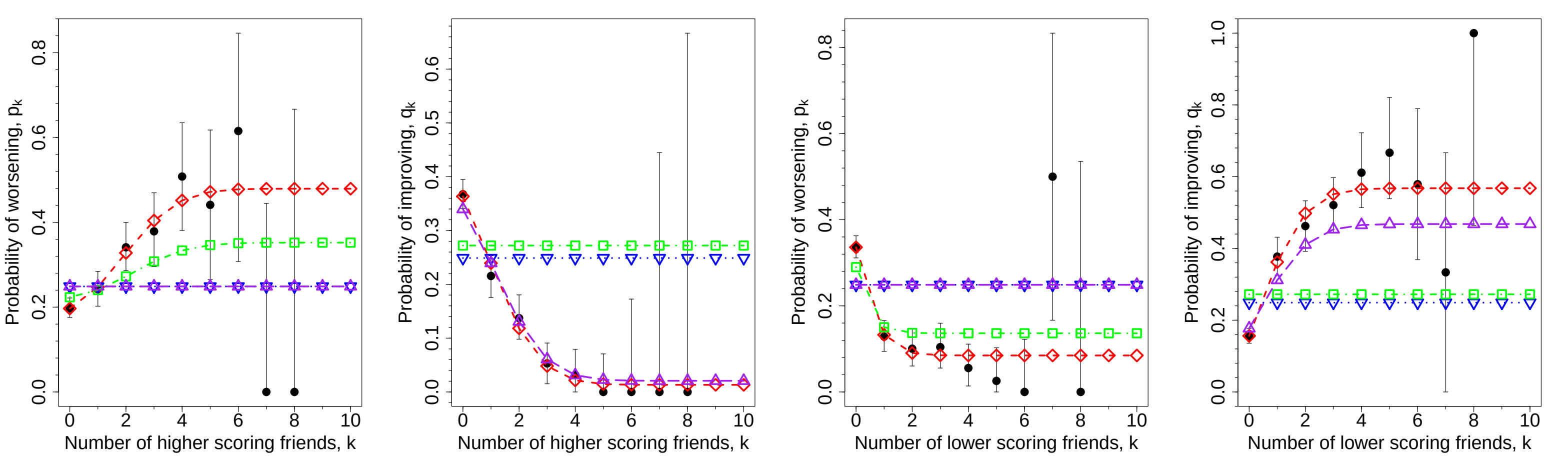
(a) Anhedonia model 1. Both transmitting is preferred.

(b) Anhedonia model 2. Both transmitting is preferred.



(c) Appetite model 1. Improving transmitting only is preferred.

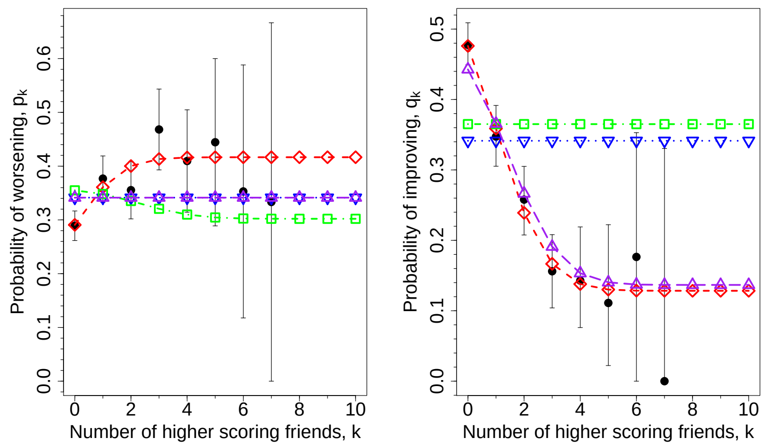
(d) Appetite model 2. Both transmitting is preferred.



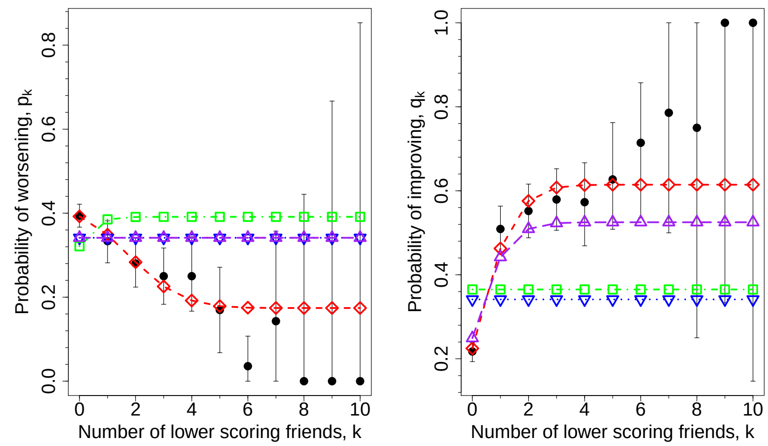
(e) Concentration model 1. Both transmitting is preferred.

(f) Concentration model 2. Both transmitting is preferred.

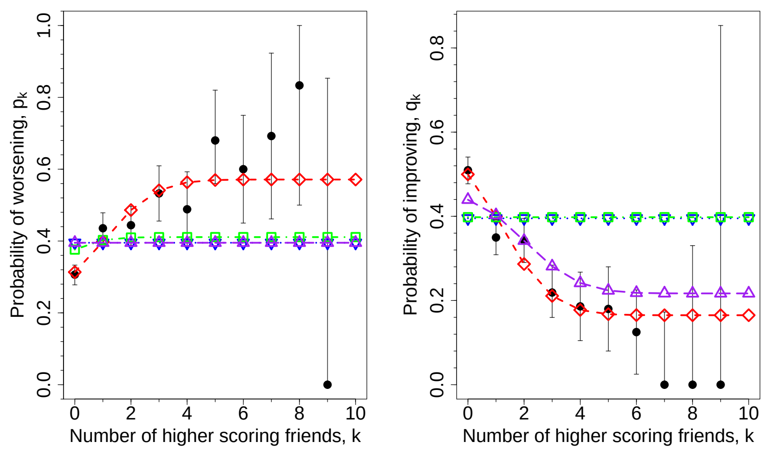
Figure 6: Inferred probabilities of worsening and improving from three state likelihood. Black points - observed frequencies. Red dashed line - both improving and worsening transmit. Blue dashed line - neither worsening or improving transmit. Green dashed line - worsening only transmits. Purple dashed line - improving only transmits. Error bars given by 95% confidence intervals. Preferred forms found by minimum AIC.



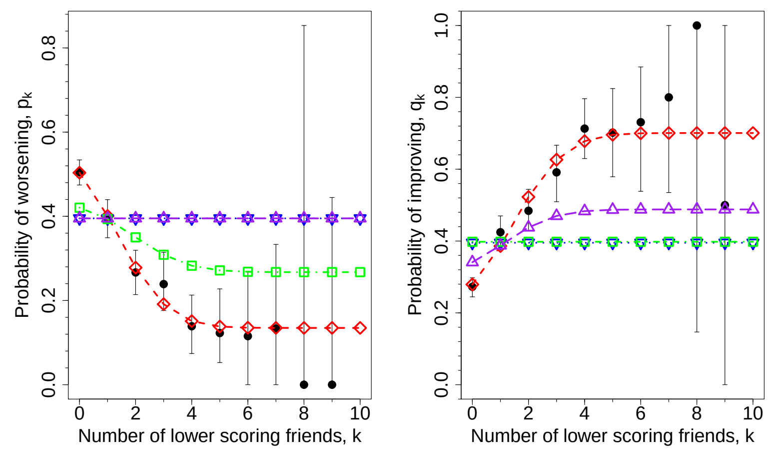
(a) Dysphoria model 1. Both transmitting is preferred.



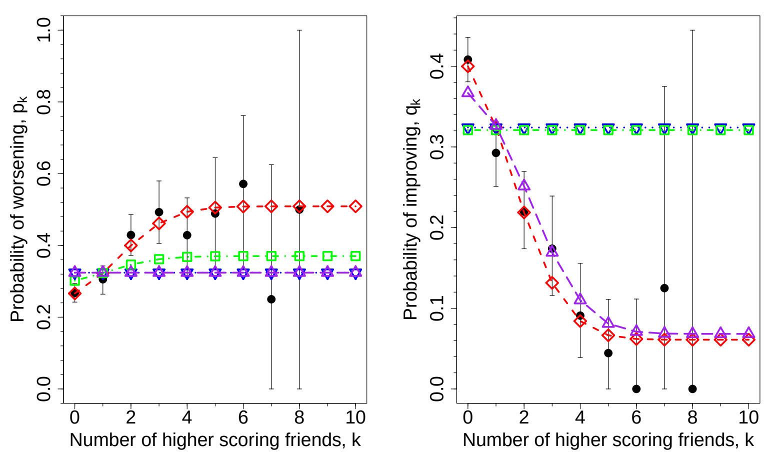
(b) Dysphoria model 2. Both transmitting is preferred.



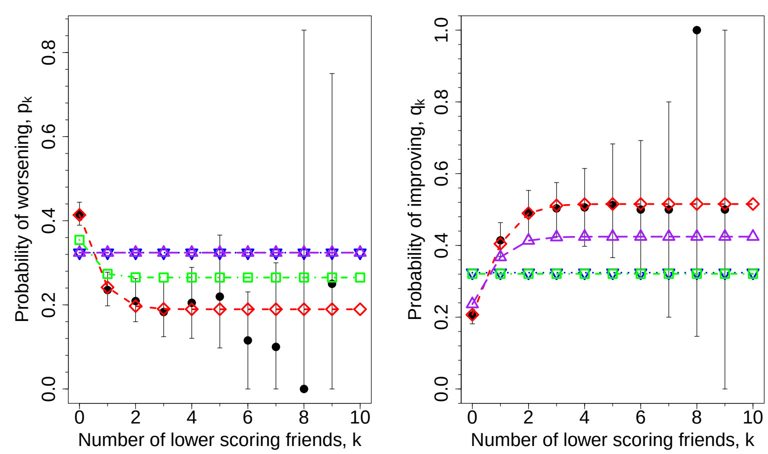
(c) Helplessness model 1. Improving transmitting only is preferred.



(d) Helplessness model 2. Both transmitting is preferred.

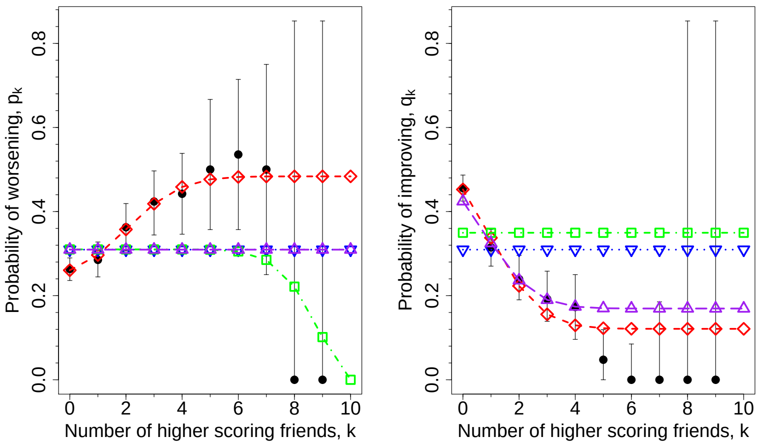


(e) Tiredness model 1. Both transmitting is preferred.

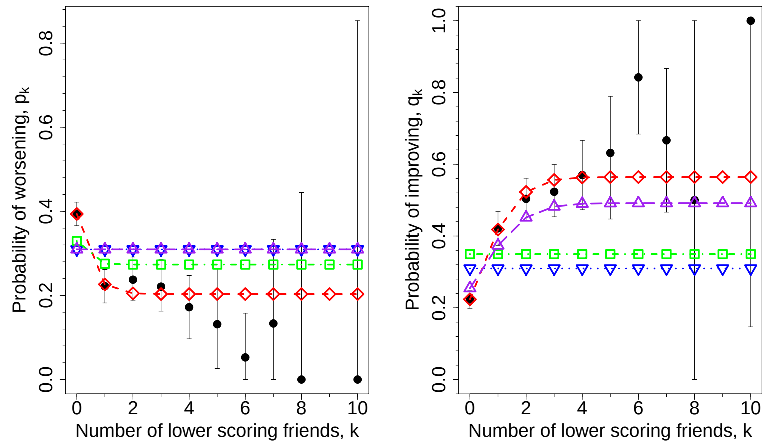


(f) Tiredness model 2. Both transmitting is preferred.

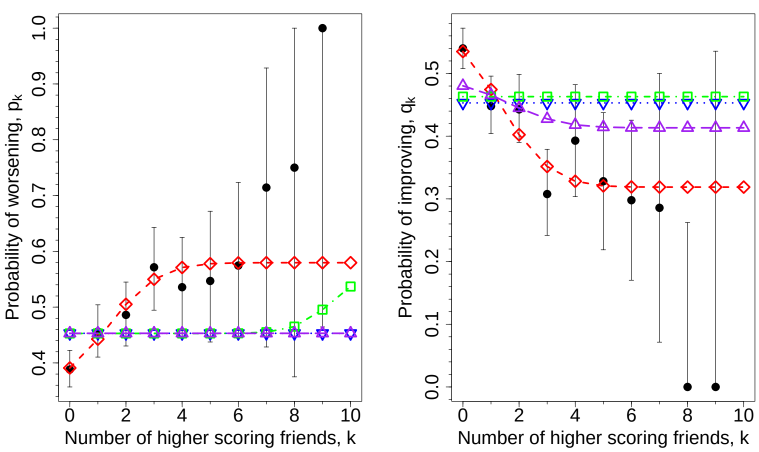
Figure 7: Inferred probabilities of worsening and improving from three state likelihood. Black points - observed frequencies. Red dashed line - both improving and worsening transmit. Blue dashed line - neither worsening or improving transmit. Green dashed line - worsening only transmits. Purple dashed line - improving only transmits. Error bars given by 95% confidence intervals. Preferred forms found by minimum AIC.



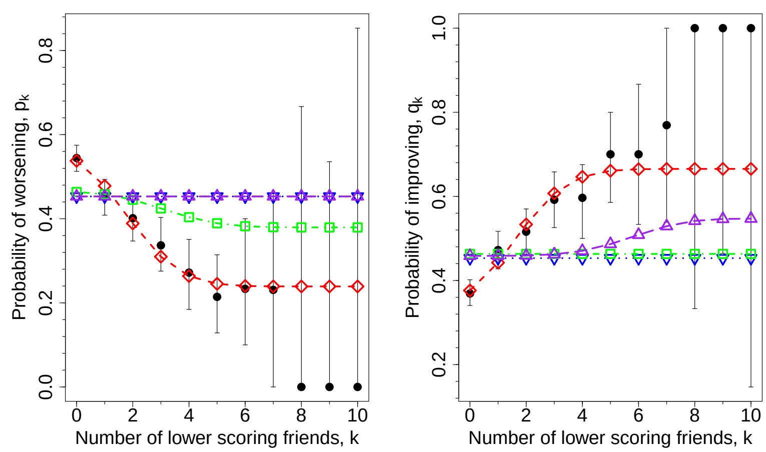
(a) Worthlessness model 1. Both transmitting is preferred.



(b) Worthlessness model 2. Both transmitting is preferred.



(c) Total CES-D score model 1. Improving transmitting only is preferred.



(d) Total CES-D score model 2. Both transmitting is preferred.

Figure 8: Inferred probabilities of worsening and improving from three state likelihood. Black points - observed frequencies. Red dashed line - both improving and worsening transmit. Blue dashed line - neither worsening or improving transmit. Green dashed line - worsening only transmits. Purple dashed line - improving only transmits. Error bars given by 95% confidence intervals. Preferred forms found by minimum AIC.

4. Discussion

4.1. Summary of Findings

The core result in our study has been that, for the symptoms of anhedonia, poor concentration, dysphoria, helplessness, tiredness, and worthlessness, and for depressive symptoms as a whole, the probabilities of an individual worsening and improving are dependent on the number of worse off friends and the number of better off friends. In other words, the more worse off friends an individual has, the more likely they are to get worse, and the less likely they are to get better, and vice versa for better off friends, for each of these symptoms. However, there is a cutoff to this effect at around 4 (as a rule of thumb) better off or worse off friends, where having any more than that will not make a difference. It would be interesting to see if this could be quantified. The fact that the total CES-D score showed this as well, shows that the mechanics of the spread of each individual symptom is reflected in the spread of depression as a whole. Support for this result was also shown in figure 4, where definite patterns can be observed in the data for the total CES-D score that reflect the impact of better off and worse off friends on the change in score.

The only symptom to go against this is poor appetite. The effect from better off friends is the same. However, having more friends with poorer appetite than an individual may make them less likely to improve in appetite, but will have no effect on whether they worsen in appetite. This could simply be due to the much more severe limitations in how much an individual could vary in appetite score, but this did not effect concentration. Therefore, it may be due to deeper dissimilarities between appetite and the other symptoms, and in how others observe their presence in an individual. Appetite has a more physical basis than the other more emotional symptoms, but could also be less evident to individuals due to how it is most often only displayed during set times of day when people eat.

Further conclusions come from sections 3.2 and 3.3. In all symptoms, and the depressive symptoms as a whole, there is no bias within the healthy population of the study towards either improving or worsening. Also, small changes in the severity of each symptom are preferred, though some change is more likely to occur than no change. Individuals tended to have low levels of each symptom, and maintained similar levels over time. Anhedonia shows the only difference, where most individuals have and maintain higher levels of severity. Therefore, this healthy population have a tendency to lack interest in enjoyable activities, though they maintain healthy levels in all other areas. This could perhaps be due to the scale not being calibrated appropriately for this particular population. The scale is used for people of all ages, not just adolescents. Perhaps it is normal for adolescents to experience a certain level of anhedonia.

Evidence from our study also suggests that the change in score follows an exponential distribution. This means that, between the two time points, individuals only got better or got worse, and did not repeatedly switch between the two. However, this was only analysed using qq plots and by comparing AIC values. More rigorous statistical tests should be used to confirm this result.

Our results are useful in the context of design of public-health efforts to improve the moods of adolescents. It is beneficial for an individual to gain more friends only if those friends are in a healthier state than the individual. If an individual is experiencing depressive symptoms this can impact on the emotional state of their friends. The same is true if an individual is not experiencing depressive symptoms. This should be taken into account when considering the encouragement of socialisation as a preventative measure for depression. A positive result occurs from this as well, however, in that knowing the effect of worse off friends allows for better tailoring of preventions to avoid a cascade in the transmission of worsening.

4.2. Limitations of the Study

The greatest limitations of the study come from the data. First off, in the friendship data. Though allowing the listing of up to 5 male and up to 5 female friends did allow for the generation of a substantial friendship network, it still enforced an unnatural upper bound on the number of friends an individual could have. Also, it did not consider friends individuals may have outside of their school or their sister school, or those who they could not find on the roster. These were listed under generic AIDs, and were ignored by our study.

Further limitations come from the CES-D scale. It has been shown to be reliable, as mentioned in section 2.3. However, it is also limiting in two ways. First, it assigns a discrete measure of mood to each individual. A continuous scale would possibly be more appropriate, but would be very difficult to implement, and as yet does not exist. Second, it is self-assessed, and therefore depends on an individuals ability to recognise changes of symptoms within themselves. A third point would be that it gives an upper bound to the mood score, but as very few get anywhere near this upper bound, this does not appear to be a limiting factor.

As per the nature of all data, errors and ommissions exist within the dataset used, but efforts were made to minimise these, as detailed in section 2.1.

Further limitations were introduced by the method of the study. The splitting of the scale into symptoms was performed by common sense, as no official method exists. Though unlikely, it is certainly possible that others may disagree with the assignment of questions to symptoms. However, all symptoms considered are ones represented in the scale [20].

A final limitation is in the consideration of only three states, i.e. improving, worsening, and not changing. Many more states could be studied, one for each possible change in score, but this would be unwieldy and introduce possibly unneeded complexity.

4.3. Findings in the Context of Previous Work

As it is upon their work we build, it is most informative to consider our results in the context of the findings of Hill et. al. [18]. They showed that healthy moods transmit, whilst unhealthy moods do not. Our results appear different. They also show probabilities of becoming not depressed that have more of a step shape, where having a middling amount of friends had an impact, but low numbers of friends and high numbers of friends did not. Our probabilities increased even for low numbers, but plateaued well before reaching high numbers.

All of this can be explained by the differences in our methodology. They considered the crossing of individuals over a threshold CES-D score, while we considered individuals changing score by any amount in either a positive or negative way. The different methodologies produce different results. It is of note, though, that the conclusions are different in that we found that worsening in depressive symptoms also transmits, while their work concluded there was no strong evidence that depression does transmit. A possible reason for this can be seen in figure 4, the density plots coloured by higher and lower scoring friends in section 3.2. We note that, for lower scoring friends, those who have more appear above the diagonal, and so have gotten better. In fact many appear quite far above it, above the wave 1 threshold used in Hill's work. For higher scoring friends, those who have more appear below the diagonal, and therefore have gotten worse. However, unlike with those who got better from having lower scoring friends, very few individuals appear above the threshold wave 2 score, and therefore it seems that due to the methodology of the previous work, the individuals who did get worse were missed because they did not worsen enough to cross the threshold.

As referenced in 1, much other work has concluded that increased socialisation is nothing but beneficial. Our work shows that this is not always the case, and socialising with individuals of worse symptoms can increase the likelihood of an individual getting worse. Though, the benefits of socialising with individuals of better symptoms still exist.

4.4. Future Work

Much further work can be done in this area. Leading directly on from this work, further analysis of appetite compared to other symptoms is warranted. Further confirmation of the distribution of score changes is also needed. More complicated shapes for the inferred probabilities could also be considered, if they can be justified. A model with all possible changes in score making up all the possible states could also be considered.

With the data set, more work could be done in looking at other risk factors for depression, other than the symptoms that make up the CES-D scale, and how they spread across social networks. Goudie et. al. have already looked into which areas covered in the Add Health study impact on depression [25]. This then identifies risk factors that could be considered.

A final interesting area is in creating models of changes in mood. The majority of the work done now, including ours, is comprised of statistical data analysis, where we infer results from the data. The example of Physics could be followed, in making a model first to later be confirmed or denied by data. Various Markov type models have already been considered, but have obvious limitations [26]. It might be beneficial to design a deterministic dynamical model of mood, incorporating features seen in the literature, such as resilience [27].

5. Acknowledgements

Thanks goes to the Engineering and Physical Sciences Research Council for supporting this work, to Add Health for allowing us access to their data, to Rob Goudie for supplying R code, and to Ed Hill for supplying R code and invaluable advice.

References

[1] W. H. Organisation, Depression - fact sheet no. 369 (2014).
URL `http://www.who.int/mediacentre/factsheets/fs369/en/`

[2] V. Patel, A. J. Flisher, S. Hetrick, P. McGorry, Mental health of young people: a global public-health challenge, The Lancet 369 (9569) (2007) 1302–1313.

[3] H. Green, et al., Mental health of children and young people in Great Britain, 2004, Palgrave MacMillan Basingstoke, 2005.

[4] N. I. of Mental Health, Major depressive disorder in children (2013).
URL `http://www.nimh.nih.gov/statistics/1mdd_child.shtml`

[5] M. J. Keeling, K. T. Eames, Networks and epidemic models, Journal of the Royal Society Interface 2 (4) (2005) 295–307.

[6] S. M. Goodreau, J. A. Kitts, M. Morris, Birds of a feather, or friend of a friend? using exponential random graph models to investigate adolescent social networks*, Demography 46 (1) (2009) 103–125.

[7] P. S. Bearman, J. Moody, K. Stovel, Chains of affection: The structure of adolescent romantic and sexual networks1, American Journal of Sociology 110 (1) (2004) 44–91.

[8] N. A. Christakis, J. H. Fowler, The collective dynamics of smoking in a large social network, New England journal of medicine 358 (21) (2008) 2249–2258.

[9] N. A. Christakis, J. H. Fowler, The spread of obesity in a large social network over 32 years, New England journal of medicine 357 (4) (2007) 370–379.

[10] M. M. Ali, A. Amialchuk, S. Gao, F. Heiland, Adolescent weight gain and social networks: is there a contagion effect?, Applied Economics 44 (23) (2012) 2969–2983.

[11] R. Neumann, F. Strack, " mood contagion": the automatic transfer of mood between persons., Journal of personality and social psychology 79 (2) (2000) 211.

[12] T. E. Joiner, J. Katz, Contagion of depressive symptoms and mood: Meta-analytic review and explanations from cognitive, behavioral, and interpersonal viewpoints, Clinical Psychology: Science and Practice 6 (2) (1999) 149–164.

[13] A. Dean, W. M. Ensel, Modelling social support, life events, competence, and depression in the context of age and sex, Journal of Community Psychology 10 (4) (1982) 392–408.

[14] K. Ueno, The effects of friendship networks on adolescent depressive symptoms, Social Science Research 34 (3) (2005) 484–510.

[15] S. Y. Rueger, C. K. Malecki, M. K. Demaray, Relationship between multiple sources of perceived social support and psychological and academic adjustment in early adolescence: Comparisons across gender, Journal of Youth and Adolescence 39 (1) (2010) 47–61.

[16] N. Mead, H. Lester, C. Chew-Graham, L. Gask, P. Bower, Effects of befriending on depressive symptoms and distress: systematic review and meta-analysis, The British Journal of Psychiatry 196 (2) (2010) 96–101.

[17] A. L. Hill, D. G. Rand, M. A. Nowak, N. A. Christakis, Emotions as infectious diseases in a large social network: the sisa model, Proceedings of the Royal Society B: Biological Sciences 277 (1701) (2010) 3827–3835.

[18] E. M. Hill, F. E. Griffiths, T. House, Spreading of healthy mood in adolescent social networks, to appear.

[19] K. Harris, et al., The national longitudinal study of adolescent health: Research design (2009).
URL `http://www.cpc.unc.edu/projects/addhealth/design`

[20] L. S. Radloff, The ces-d scale a self-report depression scale for research in the general population, Applied psychological measurement 1 (3) (1977) 385–401.

[21] R. E. Roberts, P. M. Lewinsohn, J. R. Seeley, Screening for adolescent depression: a comparison of depression scales, Journal of the American Academy of Child & Adolescent Psychiatry 30 (1) (1991) 58–66.

[22] N. A. Christakis, J. H. Fowler, Social contagion theory: examining dynamic social networks and human behavior, Statistics in Medicine 32 (4) (2013) 556–577.

[23] D. Centola, M. Macy, Complex contagions and the weakness of long ties1, American Journal of Sociology 113 (3) (2007) 702–734.

[24] D. Centola, The spread of behavior in an online social network experiment, science 329 (5996) (2010) 1194–1197.

[25] R. J. B. Goudie, S. Mukherjee, F. E. Griffiths, Exploratory network analysis of large social science questionnaires, to appear.

[26] A. Le Lay, N. Despiegel, C. François, G. Duru, Can discrete event simulation be of use in modelling major depression?, Cost Effectiveness and Resource Allocation 4 (1) (2006) 19.

[27] C. Dowrick, R. Kokanovic, K. Hegarty, F. Griffiths, J. Gunn, Resilience and depression: perspectives from primary care, Health: 12 (4) (2008) 439–452.

[28] H. Akaike, A new look at the statistical model identification, Automatic Control, IEEE Transactions on 19 (6) (1974) 716–723.

[29] L. D. Brown, T. T. Cai, A. Dasgupta, Confidence intervals for a binomial proportion and asymptotic expansions, Annals of Statistics (2002) 160–201.

Appendix A. Two state models

Appendix A.1. Models

Unlike models 1 and 2, these models included two state, where we have a probability p of worsening and a probability $1-p$ of not worsening, or we have a probability q of improving and a probability $1-q$ of not improving. Model 3 examined the dependency of p on the number of higher scoring friends, model 4 the dependency of p on lower scoring friends, model 5 the dependency of q on higher scoring friends, and model 6 the dependency of q on lower scoring friends.

For each model, two forms were examined: transmission and no transmission. For models 3 and 4, the no transmission form was

$$p_k = \alpha \tag{A.1}$$

and the transmission form was

$$p_k = \alpha + \beta \sum_{l=0}^{k} \binom{10}{l} \gamma^l \left(1-\gamma\right)^{10-l} \tag{A.2}$$

where k is the number of higher scoring friends and the number of lower scoring friends for each model respectively.

Similarly, for models 5 an 6, the no transmission form was

$$q_k = \delta \tag{A.3}$$

and the transmission form was

$$q_k = \delta + \epsilon \sum_{l=0}^{k} \binom{10}{l} \zeta^l \left(1-\zeta\right)^{10-l} \tag{A.4}$$

where again k is the number of higher scoring friends and the number of lower scoring friends for each model respectively.

Appendix A.2. Parametric Inference

The likelihood function used for model 3 was

$$L\left(\mathbf{x}|\mathbf{p},\mathbf{M}\right) = \prod_k \binom{M_k}{x_k} p_k^{x_k} \left(1-p_k\right)^{M_k - x_k} \tag{A.5}$$

where x_k was the number of individuals with k higher scoring friends who worsened, and M_k was the total number of individuals with k higher scoring friends.

The likelihood function used for model 4 was

$$L\left(\mathbf{u}|\mathbf{p},\mathbf{N}\right) = \prod_k \binom{N_k}{u_k} p_k^{u_k} \left(1-p_k\right)^{N_k - u_k} \tag{A.6}$$

where u_k was the number of individuals with k lower scoring friends who worsened, and N_k was the total number of individuals with k lower scoring friends.

The likelihood function used for model 5 was

$$L\left(\mathbf{y}|\mathbf{p},\mathbf{M}\right) = \prod_k \binom{M_k}{y_k} p_k^{y_k} \left(1-p_k\right)^{M_k - y_k} \tag{A.7}$$

where y_k was the number of individuals with k higher scoring friends who improved.

The likelihood function for model 6 was

$$L\left(\mathbf{v}|\mathbf{p},\mathbf{N}\right) = \prod_k \binom{N_k}{v_k} p_k^{v_k} \left(1-p_k\right)^{N_k - v_k} \tag{A.8}$$

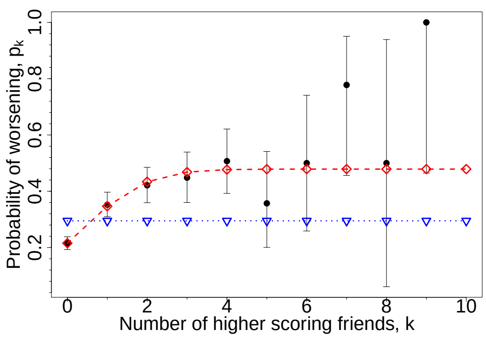
where v_k was the number of individuals with k lower scoring friends who improved.

As before, the minus log likelihood functions were minimised in R using the numerical optimising functions described in 2.6.2. The only difference was for the no transmission forms of models 3 - 6 where only a single parameter value was inferred. Then the `optimize()` function was used. The different possible forms were compared using their AIC.

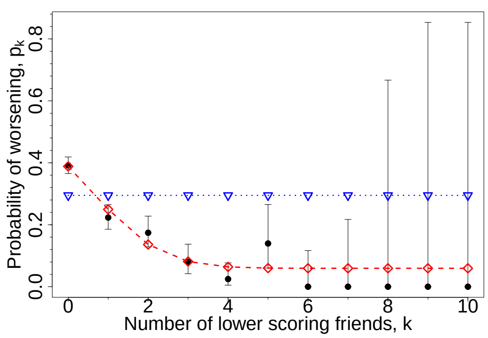
Appendix A.3. Results and Discussion

Results for the two state models can be seen in figures A.9, A.10, A.11, and A.12. Largely, they are the same as the results for the three state models, shown in section 3.3. Transmission is preferred above no transmission in all cases, allowing for the same conclusions to be drawn.

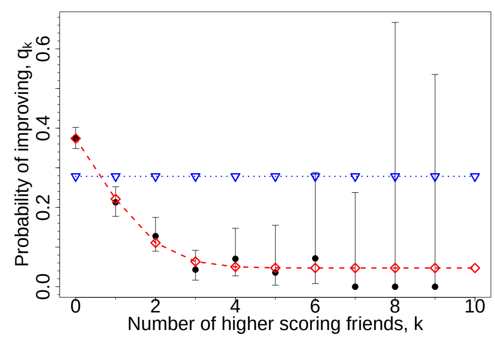
The only difference is in appetite, where transmission is preferred above no transmission for the probability of worsening being dependent on the number of worse off friends, as shown in figure A.9e. This is different from the results of the three state model, where improving was dependent on worse off friends but worsening was not. It is interesting to note that, when not contrained by also having to consider the individuals improving, the numerical optimisation of the likelihood could find a better transmission model. However, when comparing to the no transmission form, we see that the transmission form is only slightly different, incorporating the slight increase in observed frequencies between 0 and 2 friends. Therefore, this difference from the three state model does not seem significant.



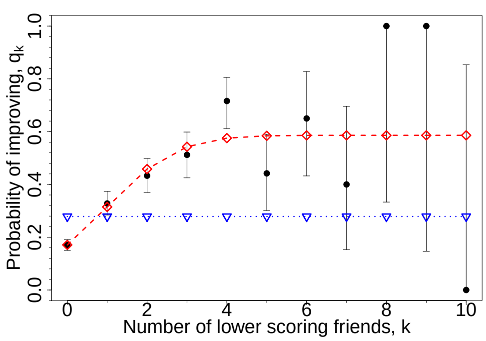
(a) Anhedonia model 3. Transmission is preferred.



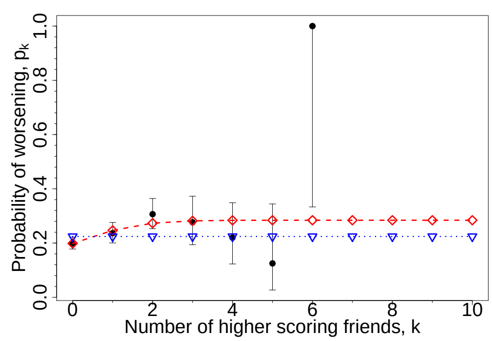
(b) Anhedonia model 4. Transmission is preferred.



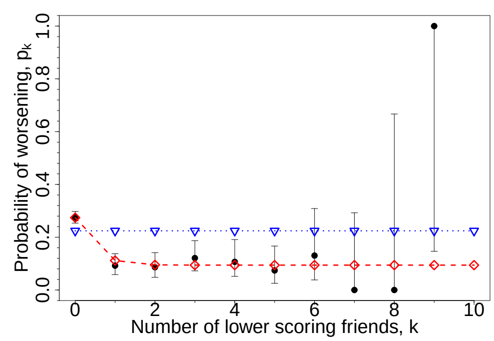
(c) Anhedonia model 5. Transmission is preferred.



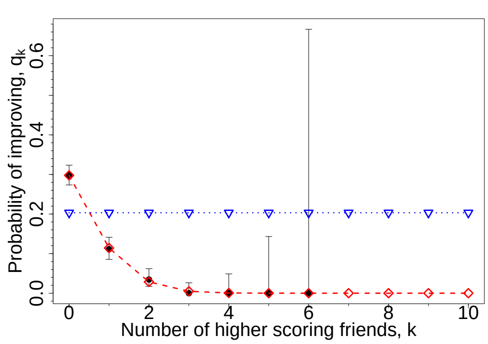
(d) Anhedonia model 6. Transmission is preferred.



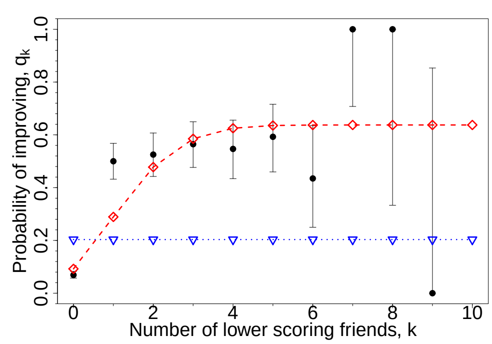
(e) Appetite model 3. Transmission is preferred.



(f) Appetite model 4. Transmission is preferred.

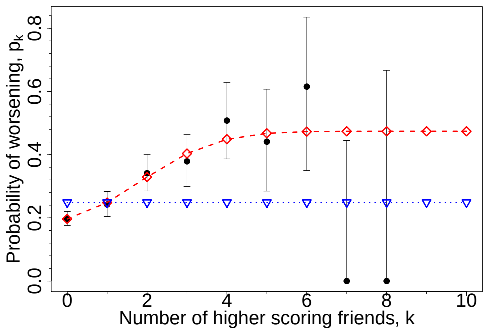


(g) Appetite model 5. Transmission is preferred.

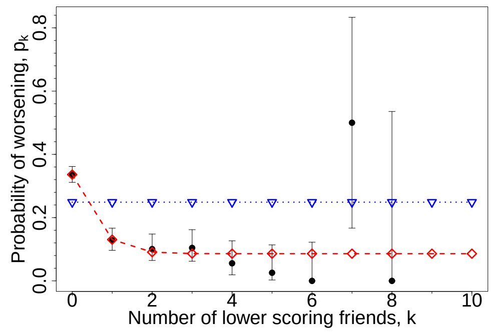


(h) Appetite model 6. Transmission is preferred.

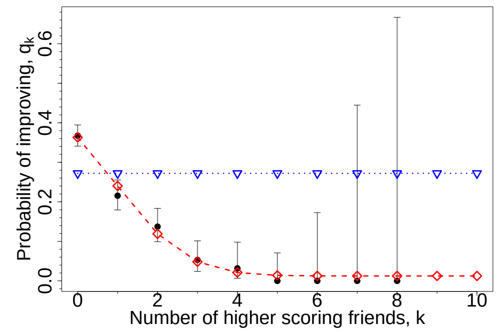
Figure A.9: Inferred probabilities of improving and worsening in level of anhedonia and appetite from two state models (i.e. improve versus not, worsen versus not). Red dashed lines - transmission form. Blue dashed lines - no transmission form. Black points - observed frequencies. Error bars are given by 95% confidence intervals. Preferred forms found by minimum AIC.



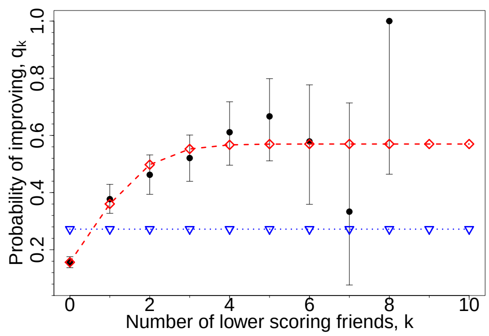
(a) Concentration model 3. Transmission is preferred.



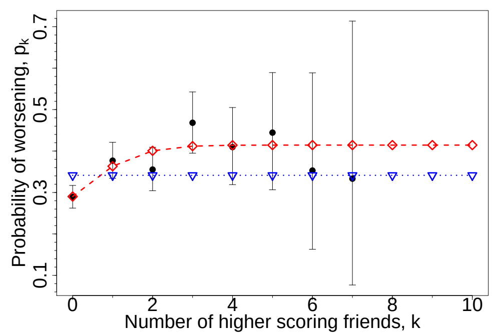
(b) Concentration model 4. Transmission is preferred.



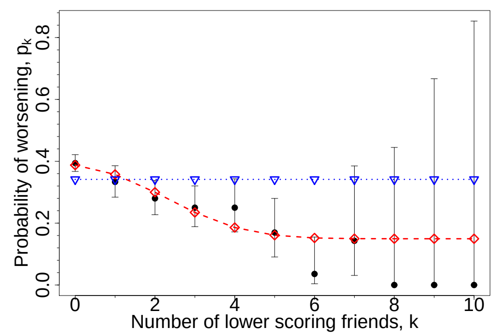
(c) Concentration model 5. Transmission is preferred.



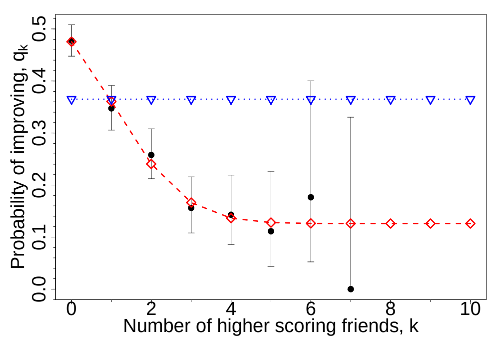
(d) Concentration model 6. Transmission is preferred.



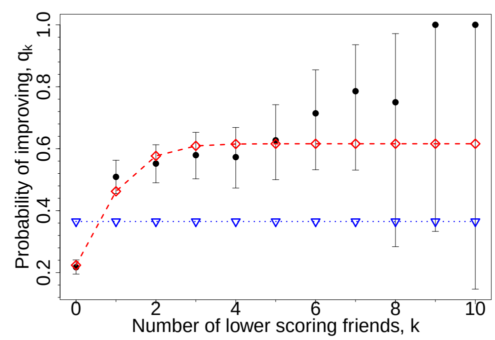
(e) Dysphoria model 3. Transmission is preferred.



(f) Dysphoria model 4. Transmission is preferred.

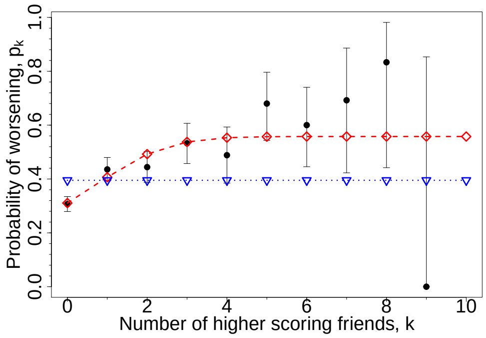


(g) Dysphoria model 5. Transmission is preferred.

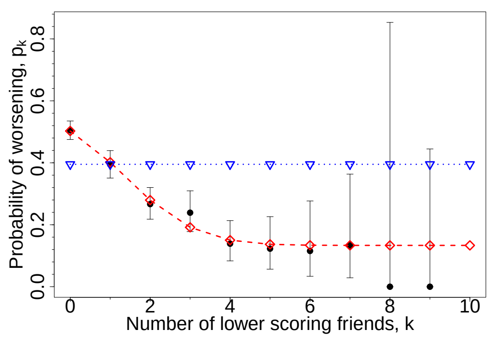


(h) Dysphoria model 6. Transmission is preferred.

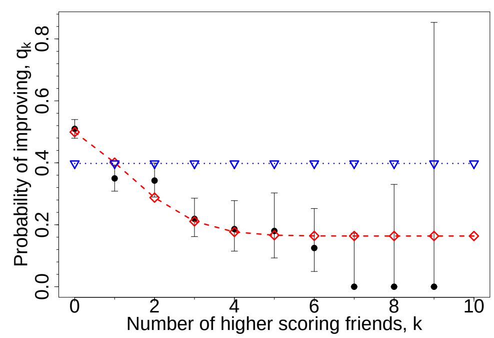
Figure A.10: Inferred probabilities of improving and worsening in level of concentration and dysphoria from two state models (i.e. improve versus not, worsen versus not). Red dashed lines - transmission form. Blue dashed lines - no transmission form. Black points - observed frequencies. Error bars are given by 95% confidence intervals. Preferred forms found by minimum AIC.



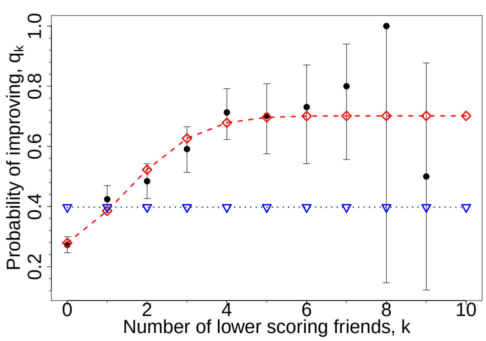
(a) Helplessness model 3. Transmission is preferred.



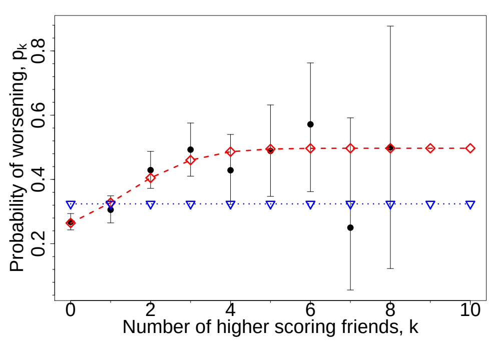
(b) Helplessness model 4. Transmission is preferred.



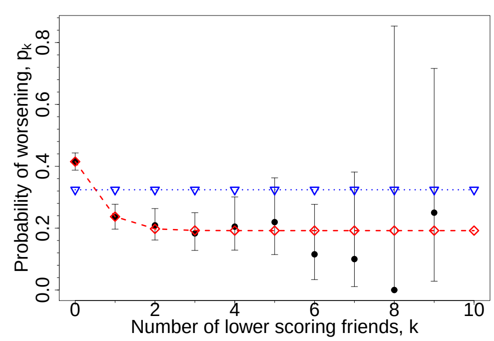
(c) Helplessness model 5. Transmission is preferred.



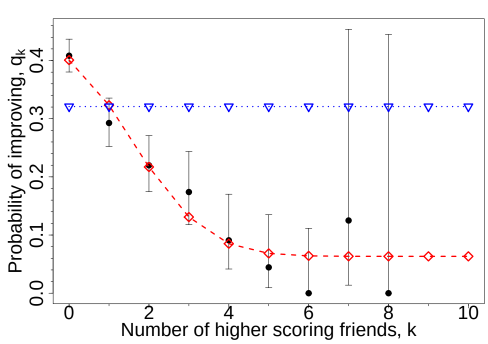
(d) Helplessness model 6. Transmission is preferred.



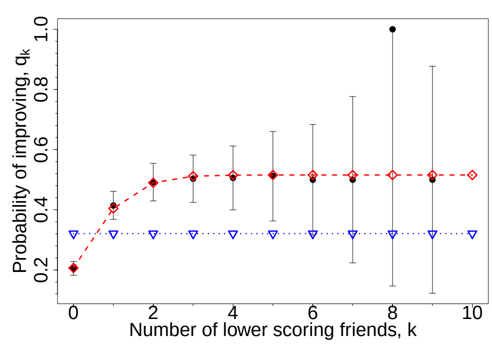
(e) Tiredness model 3. Transmission is preferred.



(f) Tiredness model 4. Transmission is preferred.

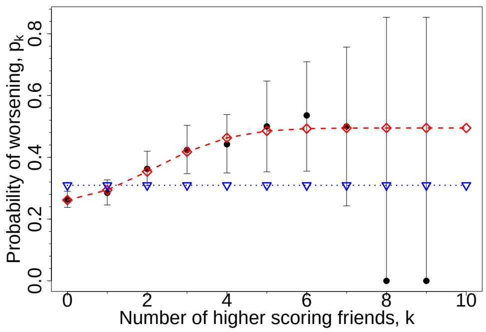


(g) Tiredness model 5. Transmission is preferred.

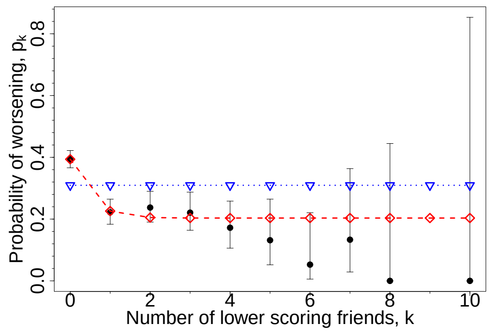


(h) Tiredness model 6. Transmission is preferred.

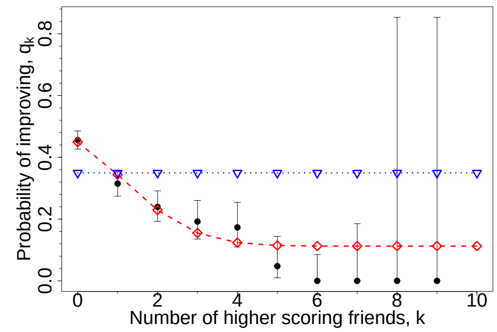
Figure A.11: Inferred probabilities of improving and worsening in level of helplessness and tiredness from two state models (i.e. improve versus not, worsen versus not). Red dashed lines - transmission form. Blue dashed lines - no transmission form. Black points - observed frequencies. Error bars are given by 95% confidence intervals. Preferred forms found by minimum AIC.



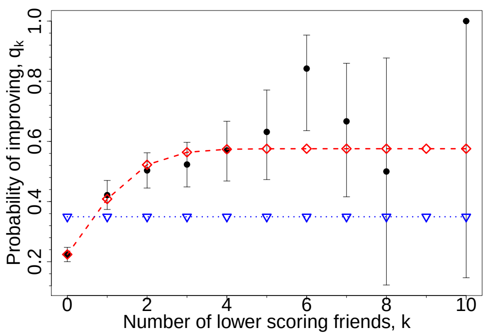
(a) Worthlessness model 3. Transmission is preferred.



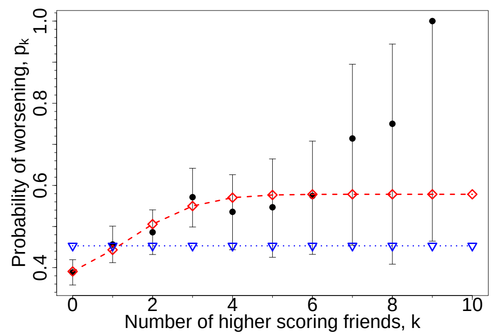
(b) Worthlessness model 4. Transmission is preferred.



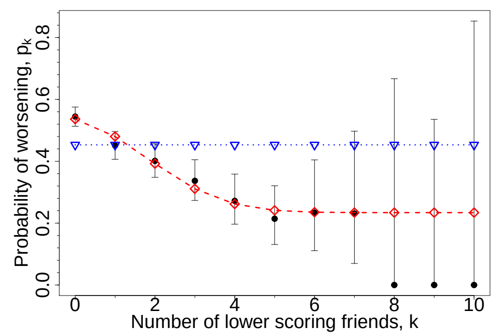
(c) Worthlessness model 5. Transmission is preferred.



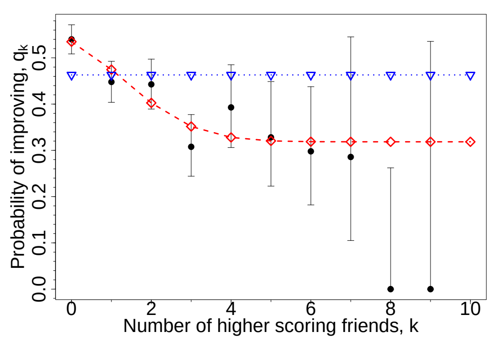
(d) Worthlessness model 6. Transmission is preferred.



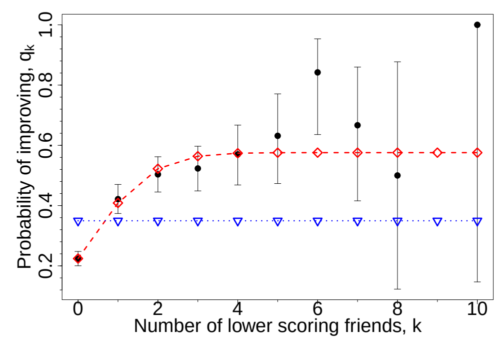
(e) Total CES-D score model 3. Transmission is preferred.



(f) Total CES-D score model 4. Transmission is preferred.



(g) Total CES-D score model 5. Transmission is preferred.



(h) Total CES-D score model 6. Transmission is preferred.

Figure A.12: Inferred probabilities of improving and worsening in level of worthlessness and total CES-D score from two state models (i.e. improve versus not, worsen versus not). Red dashed lines - transmission form. Blue dashed lines - no transmission form. Black points - observed frequencies. Error bars are given by 95% confidence intervals. Preferred forms found by minimum AIC.

Appendix B. Akaike Information Criterion

In order to compare the different forms discussed in section 2.6.1, we required a comparable measure of their quality. A common one is the Akaike Information Criterion [28]. It gives a trade off between the informational content of the model considered, and a penalty for the complexity involved. The precise definition is

$$\mathrm{AIC} = 2n_p - 2\log L_{\max} \tag{B.1}$$

where n_p is the number of parameters in the model and $L_{\max}$ is the maximised value of the likelihood function $\mathscr{L}$ of the model.

Though the AIC gives no absolute measure of the quality of the model, so reasoned comparisons to observed frequencies must still be used, it does give a way to compare different possible models. The preferred model, i.e. the model with the highest informational content and least overfitting, is that which has the minimum AIC value. So, to compare different models, we simply calculated their AIC values, and then found the one with the smallest.

Appendix C. Jeffreys Interval

We wanted to be able to compare the inferred models to the observed frequencies. To do so, we wished to attach error bars to the observed frequencies using 95% confidence intervals. While this could be done using bootstrapping for most of the three state frequencies, we decided to use Jeffreys intervals for the two state models.

The Jeffreys interval provides a Bayesian motivated confidence interval for binomial proportions [29]. Using a non-informative Jeffreys prior (i.e. a Beta distribution of parameters (1/2,1/2)) for the binomial proportion p, after x successes in n trials we will be left with a posterior Beta distribution of parameters $(x + 1/2, n - x + 1/2)$. For $x \neq 0$ and $x \neq n$, the Jeffreys interval is the $100\,(1-\alpha)\,\%$ equal tailed posterior probability interval with lower bound

$$L = \mathscr{B}^{-1}\left\{\alpha/2, x + 1/2, n - x + 1/2\right\} \tag{C.1}$$

and upper bound

$$U = \mathscr{B}^{-1}\left\{1 - \alpha/2, x + 1/2, n - x + 1/2\right\} \tag{C.2}$$

where $\mathscr{B}^{-1}\left\{y, p_1, p_2\right\}$ is the quantile function at y of the Beta distribution of parameters p_1 and p_2. When $x = 0$, the lower limit is set to 0 and the upper limit remains as before. When $x = n$, the upper limit is set to 1 and the lower limit remains as before.