Demonstrating opportunities of visualisation and network analysis in physical activity and its determinants: Baseline associations in the Let’s Move It trial

|  |
| --- |
| Matti T. J. Heino^1,2\*^, Eiko Fried3, Reijo Sund1,2, Ari Haukkala1, Keegan Knittle1, Katja Borodulin1, Antti Uutela1, Vera Araujo-Soares1, Falko Sniehotta1, Tommi Vasankari1, & Nelli Hankonen1 |
| 1 Department of Social Research, University of Helsinki |
| 2 Faculty of Social Sciences, University of Tampere |
| 3 Department of Psychology, Columbia University |
|  |

# Author note

Trial Registration Number: ISRCTN10979479. Registered retrospectively: 31.12.2015

Correspondence concerning this article should be addressed to Matti T. J. Heino, . E-mail: [matti.tj.heino@gmail.com](mailto:matti.tj.heino@gmail.com)

Abstract

Background: Let’s Move It is a cluster-randomised controlled trial evaluating a novel theory-based intervention, which aimed to reduce sedentary behaviour (SB) and increase physical activity (PA) among older adolescents in vocational schools, by targeting environmental and psychosocial determinants of the phenomena. This paper describes the characteristics of its baseline cohort in both arms, and explores the possibilities for visually presenting such data, making use of recent developments in software and network analyses. We provide a template for researchers to apply these tools to other data. Methods: At baseline, 1123 adolescents in 57 classes in 6 school units participated in the study. Data were gathered with 7-day accelerometry, bioimpedance measures and questionnaires, which were used to measure health status, activity behaviours, and potential constructs mediating the effect of the intervention on outcomes. Data were visualised e.g. combining ridge plots and diamond plots; network analysis was used to investigate relations between psychosocial variables and outcomes. Results: The participants were 16-49 years old (M = 18.8, Md = 17.0). On average, they engaged in moderate-to-vigorous PA 1h 29min (CI95: 1h 19min - 1h 40min) and SB 9h 34min (CI95: 8h 56min - 10h 12min) daily and interrupted sitting 28 times (CI95: 24.7 - 31.4) per day on average. Several variables differed among the four educational tracks, but not across intervention and control groups. Naturally occurring behaviour change technique (BCT) usage was reported as low for many but not all techniques. Conclusion: We have shown benefits of presenting such data visually, and encourage researchers to routinely make the extensive analyses and descriptions they have produced, available in website supplements. This practice has potential to increase the speed and quality of scientific communication, as well as to address recent concerns of reduced confidence towards research findings.

*Keywords:* exercise, physical activity, school-based intervention, behaviour change, sedentary behaviour

Word count: X

Demonstrating opportunities of visualisation and network analysis in physical activity and its determinants: Baseline associations in the Let’s Move It trial

# Background

Declining physical activity (PA) and increasing sedentary behaviour (SB) are costly and growing concerns for public health, especially for individuals with low socioeconomic status (SES) [1]. Patterns of low PA among adults begin earlier in life. As there is evidence that the declines in PA and increases in SB are already evident in childhood and adolescence [3, 4], there is a need for further research on to how to improve PA and SB among adolescents.

As adolescents spend a significant amount of their time in schools, schools provide a promising opportunity for PA interventions [5]. The Let’s Move It intervention aimed to reduce SB and increase PA among adolescents in vocational schools; developed using stakeholder input and co-creation with target group representatives, as well as behavioural science theory and empirical evidence [6, 7]. Contrary to typical school-based interventions with uniform study populations, this trial was carried out in vocational schools with different educational tracks, differences between which may be important to making accurate conclusions in the trial evaluation phase.

The programme theory (explicated in REF: DEVELOPMENT; see also [8]; [9]; <https://mrc.ukri.org/documents/pdf/mrc-phsrn-process-evaluation-guidance-final/>, p. 32) for changing PA and SB was hypothesised to be slightly different for PA and SB. In order to engage in PA, one needs to have a conscious effort and self-regulatory capacity to make use of the opportunities, such as planning for active times and overcoming barriers to exercise. In the intervention, one of the key emphases in helping adolescents change their PA was to help them understand and use techniques to manage their motivation and behaviour (see also [10]). To date, there is little systematic theorising on how the use of these techniques link to each other, and it would be important to understand these interlinkages empirically. The theoretical model for changing SB, on the other hand, is more driven by environmental opportunities and incentives, such as having the option of standing up during class.

In order to change moderate-to-vigorous-intensity PA, a central component of the intervention targeted autonomous motivation, social cognitions, as well as participants’ skills to use behaviour change techniques to self-regulate motivation and behaviour (Hankonen et al, unpublished manuscript). To change SB, or specifically, to reduce total SB as well as introduce breaks in SB, the program aimed to change the school environment by training teachers in providing more active teaching and altering physical choice architecture in classrooms (Köykkä et al, accepted) . The intervention included also poster campaign in schools and a website, as well as materials to target community actors and parents [11]. More information of the content of the intervention and the development of it is reported elsewhere (Hankonen et al 2017 NELLI?), Hankonen et al unpublished manuscript). The mediators postulated by the program theory included behavioural beliefs (outcome expectations, descriptive norms, intention, self-efficacy/perceived behavioural control), autonomous and controlled motivation, environmental opportunities, action and coping planning, and behaviour change technique (BCT) use. Key hypotheses regarding students’ PA change have been registered in OSF (<https://osf.io/tb8fu/>).

It has long been a standard recommendation for quantitative analyses to investigate data visually as a core precursor of conducting statistical analyses [12, 13]. However, in social and life sciences, such visualisations have rarely been shared in publications. Information about data are usually limited to means and standard deviations, which presents at best limited information about the variables of interest. Medians, modes, skewness and kurtosis provide helpful additional information, but can still hide important distributional properties.

Successful communication of important information includes using ways to avoid flooding readers with unintuitive numerical point values or easily reified bound estimates, such as confidence intervals. Visualising data is imperative, because it allows for communicating large amounts of information-and the associated uncertainty-in an accessible format, without requiring extensive mathematical background from the reader. When researchers extend work on previous findings of others, they may not know that they are working with inadequately complete information; this has direct implications to the recent crisis of confidence in the reproducibility and replicability of research findings [14–23].

Three recent developments allow a different approach. First, many journals now allow for publication of supplementary online materials, which circumvents both word and figure restrictions of traditional manuscripts. Second, statistical software such as R [24] has recently become increasingly mainstream among applied researchers, opening the door for a wide variety of data visualisation techniques. Third, novel statistical methods in social and health psychology, such as psychological network analysis, may help in understanding relationships of variables by making better use of visual representations of their associations.

In summary, by describing the characteristics of the Let’s Move It baseline cohort, the current paper aims to 1) provide a strong rationale for the urgency of data visualisation, discuss its advantages and recent developments in scientific publishing, statistical software, and statistical models that enable researchers to use data visualisation tools more easily and efficiently; 2) provide a detailed visualisation of the LMI trial baseline data, with focus on psychosocial correlates and hypothesised mediators of the intervention effect on moderate-to-vigorous physical activity; 3) describe, in particular, the relationships between physical activity, quality of motivation and behaviour change technique usage; and finally, 4) it provides all code to use as a template for other researchers to get started with, and delivers the information in a format which only necessitates a web browser to access.

All conducted analyses and visualisations with accompanying code, can be found in the supplementary website at <https://git.io/fNHuf> (permalink at [REF]). Source code to reproduce this manuscript itself can be found at [REF].

# Methods

The study has been described earlier in [11]. In brief, the study was a cluster-randomised controlled trial of a complex multi-level intervention based in Finnish vocational schools. The consenting participants answered an electronic survey, underwent bioimpedance measurement and were instructed to wear an accelerometer for seven consecutive days.

Five schools providing four educational tracks; 1. Practical Nurse (Nur), 2. Hotel, Restaurant and Catering (HRC), 3. Business and Administration (BA), and 4. Information and Communications Technology (IT) were recruited. Schools were paired so that there would be matching numbers of students from each educational track for both members of the pair. Blinded randomization by a statistician without knowledge of pairs, track or schools was then conducted so that a random member of each pair was selected as case school, the other as control school (details reported in [11]). Participants were blind to randomisation at baseline.

## Measures

The measurements have been previously described in [11], and all individual items of the scales are available in the supplementary file [MATTI TODO]. Thus, we will present these baseline measures only briefly.

### Primary outcome variables

The primary outcome for PA was moderate-to-vigorous intensity physical activity (MVPA). It was measured by accelerometry and self-reports. Primary outcomes for sedentary behaviour (SB) were measured by the accelerometer. They included time spent sitting or lying down, and the number of times sitting was interrupted during a day.

*Self-reported MVPA.* Self-reported MVPA was measured with two questions in accordance with the NordPAQ measurement [25]. The first question asked participants about the number of days in which they did more than 30 minutes of MVPA during the last week, the other asked about the overall amount of MVPA (in hours) during the previous week.

*Accelerometer-measured MVPA.* Within one week after responding to the questionnaire, students were given an accelerometer to be worn for seven consecutive days. The hip-worn accelerometer (Hookie AM 20, Traxmeet Ltd, Espoo, Finland) using a digital triaxial acceleration sensor (ADXL345; Analog Devices, Norwood MA) was attached to a flexible belt and participants were instructed to wear the belt around their right hip for seven consecutive days during waking hours, except during shower and other water activities. The acceleration signal was collected at 100 Hz sampling frequency, ±16 g acceleration range and 0.004 g resolution. PA-parameters were based on mean amplitude deviation (MAD) of the resultant acceleration analysed in 6s epochs [26]. The MAD values were then converted to metabolic equivalent (MET) values [27]. The epoch-wise MET values were further smoothed by calculating 1min exponential moving average [**TOMMI CHECK if 1min moving avg or 6s epoch exponential moving average**]. Using the smoothed MET values total PA was classified in terms of energy consumption covering MET values higher than 1.5 and moderate-to-vigorous PA (MVPA) covering MET values equal to or higher than 3 [26, 27].

*Sedentary behaviour.* According to the definition of SB [28], time spent in sitting and reclining positions were combined to indicate SB, whereas standing was analysed separately as another form of stationary behaviour. Body postures were recognized from the raw acceleration data by employing both direction and intensity information from all three measurement axes. The recognition was based on the low intensity of movement (<1.5 MET) and the accelerometer orientation in relation to identified upright position (angle for posture estimation, APE) calculated at the end of each 6 s epoch [29].

### Theoretical predictors of PA

The mediators postulated by the program theory included behavioural beliefs (outcome expectations, descriptive norms, intention, self-efficacy/perceived behavioural control), autonomous and controlled motivation, opportunities, action- and coping planning, and behaviour change technique (BCT) use. Participants were allowed to skip questions, and scales were computed as means of all items where responses were available. Items and scales are available in the supplementary website.

### Statistical analysis

We used RStudio [30] running R [31] for all our analyses and figures.

We used psychological network analysis to estimate and visualise relations among items. Such networks contain nodes (variables) and edges (statistical relationships between variables). We used state-of-the-art network models that estimate conditional dependence relations among a set of items, which can be interpreted akin to partial correlations. An edge between two variables implies that they are related after controlling for all other items; the absence of an edge implies that the two items are conditionally independent.

Network analysis has recently shown promise in many fields such as social psychology [32, 33], personality [34], intelligence [35], psychopathology [36], and empathy research [37], and is beginning to be applied for health behaviours on a broader scale. Several helpful tutorial papers aimed at empirical researchers working in psychology are available ([38]; [39]; [40]; [41]; [42]].

Network models applied to between-subjects data at one time-point can be useful for describing health psychological data, as well as facilitating group-level hypothesis generation regarding which parts of the system are central for a problem at hand [43]. Identifying these determinants of importance can thus supplement traditional structural equation modeling (SEM) approaches, while dealing with some possibly problematic approaches to SEM [44, 45].

Network analysis naturally entails its own set of assumptions. As with any model, it does not make sense to include variables which can be thought to be embedded in each other. For example, it is difficult to argue that there is no conceptual overlap between positive outcome expectations and autonomous motivation. In this regard, behaviour change technique usage and the quality of one’s motivation (as posited by self-determination theory) seem less problematic. In the case of the model presented here, we dichotomised heavily skewed controlled forms of motivation (introjected & extrinsic) such that if a person answered 3 or higher (at least partly, or more true for me), 1 was inputed and 0 otherwise.

The Mixed Graphical Model uses regularization, a procedure that has been shown to help recover the true network structure in data in case the data were simulated under a network model. Regularization has the goal to avoid estimating spurious relationships among items (i.e. false positive relations), and results in a parsimonious network structure. The regularization technique used here is the Least Absolute Shrinkage and Selection Operator (LASSO; [46]), which shrinks all edges and sets very small edges to exact zero. A paper that explains lasso regularization in network models in detail can be found elsewhere [39]. The model also displays the proportion of a given node’s variance, which can be predicted by its connected neighbours, providing an intuitive measure of how much of the value is determined by the other nodes [47].

# Findings

Table 1 shows the main demographic variables of the cohort by educational track. Most (83.1%) participants were born in Finland. While on average the sample consist of both boys and girls (43.5% vs. 56.5%), educational tracks were heavily divided by gender: Practical Nurse track had the highest amount of girls (82.3%) and IT track lowest (16%). Age ranged from 16 to 49, with the average age being 18.50. Altogether there were 190 students who reported being 20-year-olds or older.

Table 1

*Baseline demographics of educational tracks. Nur = Practical nurse, HRC = Hotel, restaurant and catering studies, BA = Business and administration, IT = Business information technology. Omitted are 24 participants, who reported “other” as their track, as well as 81 participants from whom the data is not available.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Nur | HRC | BA | IT | Full sample |
| n | 402 | 213 | 282 | 163 | 1165 |
| Mean study year (sd, median) | 1.7 (0.9, 1.0) | 1.9 (0.7, 2.0) | 1.7 (0.9, 1.0) | 1.7 (0.9, 1.0) | 1.7 (0.9, 1.0) |
| Mean age (range, median) | 18.8 (16.0-49.0, 17.0) | 18.5 (17.0-27.0, 18.0) | 18.0 (16.0-35.0, 17.0) | 18.5 (17.0-43.0, 17.0) | 18.5 (16.0-49.0, 18.0) |
| Born in Finland (%) | 80.0 | 88.2 | 87.1 | 87.9 | 83.1 |
| % girl | 82.3 | 60.6 | 39.0 | 16.0 | 56.5 |
| % allocated to intervention | 68.9 | 31.5 | 53.5 | 46.6 | 53.6 |

Table 2 shows summary statistics for primary outcome variables with their intra-class correlations (ICCs) for class and school (see supplementary website for ICCs for all variables).

At baseline, 79% students provided at least 4 days with a minimum of 10 hours per day of valid accelerometer data. On average, the youth reported engaging in at least 30 minutes of MVPA on 2.80 days a week.

Table 2

*Primary outcome variables with their class and school ICCs. Primary outcome variables highlighted with asterisks. Accelerometer results, including wear time, are only included from those participants who met the cutoff of at least 10 hours of measurement time for at least four days.*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Mean | CI95 | ICC class | ICC school | n |
| Daily moderate-to-vigorous PA time (accelerometer)\* | 1h 29min | 1h 19min - 1h 40min | .083 | .100 | 706 |
| Daily light PA time (accelerometer) | 1h 34min | 1h 26min - 1h 42min | .074 | .074 | 706 |
| Daily standing time (accelerometer) | 1h 32min | 1h 19min - 1h 45min | .140 | .098 | 706 |
| Daily time spent sitting or lying down (accelerometer)\* | 9h 34min | 8h 56min - 10h 12min | .086 | .145 | 706 |
| Daily number of times sitting was interrupted (accelerometer)\* | 28.0 | 24.7  -    31.4 | .058 | .085 | 706 |
| Number of days with >30 MVPA min previous week (self-report)\* | 2.8 | 2.6  -     3.0 | .047 | < .001 | 1082 |

## Theoretical mediators: Traditionally presented results

In table 3 below, we present the means for the primary outcome variables by gender and intervention group allocation. We do not present statistical tests for several reasons. First of all, following the logic of Neyman-Pearson hypothesis testing [48, 49], to keep error rate under the alpha level, one would have to correct for multiple testing and it is unclear how many tests one should correct for, when hypotheses are not pre-specified [18, 50]. Ignoring this – especially in our case, where it is unclear how to heed the recommendation to justify one’s alpha level [51] – error rates can become surprisingly high [20]. Besides the logic of hypothesis testing, and more importantly, we are not interested in whether the populations differ by any arbitrarily small amount on any on the variables. What matters is, whether this difference will affect the analysis and interpretation of outcomes for the main trial (i.e. a null hypothesis of nil difference is not sensible in our case), and the minimal meaningful effect size is of primary importance here. However, two caveats follow: Firstly, effect sizes not accounting for the multilevel structure of the data inflate the standard errors, possibly even making zero effects appear as medium-sized ones [52]. Secondly, it is not a trivial task to derive trustworthy effect sizes for nested data. Although some solutions exist (e.g. [52]), they have not yet been empirically validated for finite populations in the second or third levels [53], nor is there a straightforward software implementation at the time of writing. [**REIJO** maybe refurbish this?] Therefore, we opt to present the means with their corresponding confidence intervals, encouraging the readers to refrain from merely considering non-overlapping intervals between groups as hypothesis tests.

Table 3

*Main mediating variables of PA and SB*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Variable | Girls | Boys | Intervention | Control | Total |
| PA action planning | 2.7 (2.6 - 2.8) | 2.8 (2.7 - 2.9) | 2.7 (2.6 - 2.8) | 2.8 (2.7 - 2.9) | 2.8 (2.7 - 2.8) |
| PA agreement-BCTs | 3.1 (2.9 - 3.2) | 3.1 (3.0 - 3.3) | 3.0 (2.9 - 3.2) | 3.2 (3.0 - 3.4) | 3.1 (3.0 - 3.2) |
| PA amotivation | 1.5 (1.4 - 1.5) | 1.6 (1.5 - 1.7) | 1.5 (1.4 - 1.6) | 1.5 (1.4 - 1.7) | 1.5 (1.5 - 1.6) |
| PA autonomous regulation | 3.3 (3.2 - 3.5) | 3.6 (3.4 - 3.7) | 3.3 (3.2 - 3.5) | 3.5 (3.3 - 3.6) | 3.4 (3.3 - 3.5) |
| PA controlled regulation | 1.9 (1.8 - 2.0) | 1.8 (1.7 - 1.8) | 1.8 (1.7 - 1.9) | 1.9 (1.8 - 1.9) | 1.8 (1.8 - 1.9) |
| PA coping planning | 2.4 (2.4 - 2.5) | 2.6 (2.5 - 2.7) | 2.5 (2.4 - 2.6) | 2.5 (2.4 - 2.6) | 2.5 (2.4 - 2.6) |
| PA descriptive norm | 4.3 (4.1 - 4.5) | 4.5 (4.4 - 4.7) | 4.3 (4.1 - 4.5) | 4.5 (4.3 - 4.7) | 4.4 (4.2 - 4.6) |
| PA frequency-BCTs | 2.5 (2.4 - 2.6) | 2.6 (2.5 - 2.7) | 2.5 (2.4 - 2.6) | 2.6 (2.4 - 2.7) | 2.5 (2.4 - 2.6) |
| PA injunctive norm | 4.6 (4.4 - 4.8) | 4.8 (4.5 - 5.0) | 4.5 (4.3 - 4.7) | 4.8 (4.6 - 5.0) | 4.7 (4.5 - 4.8) |
| PA intention | 5.3 (5.1 - 5.5) | 5.5 (5.2 - 5.7) | 5.4 (5.1 - 5.7) | 5.4 (5.1 - 5.7) | 5.4 (5.2 - 5.6) |
| PA opportunities | 5.1 (5.0 - 5.1) | 5.2 (5.1 - 5.3) | 5.1 (5.0 - 5.2) | 5.2 (5.1 - 5.3) | 5.1 (5.1 - 5.2) |
| PA outcome expectations | 4.7 (4.6 - 4.8) | 4.5 (4.4 - 4.6) | 4.6 (4.4 - 4.7) | 4.7 (4.5 - 4.8) | 4.6 (4.5 - 4.7) |
| PA perceived behavioural control | 5.2 (5.1 - 5.3) | 5.5 (5.4 - 5.6) | 5.3 (5.1 - 5.5) | 5.3 (5.1 - 5.5) | 5.3 (5.2 - 5.5) |
| PA self-efficacy | 5.1 (5.0 - 5.3) | 5.3 (5.2 - 5.5) | 5.2 (5.0 - 5.3) | 5.3 (5.1 - 5.4) | 5.2 (5.1 - 5.4) |
| SB descriptive norm | 3.2 (3.0 - 3.4) | 3.4 (3.1 - 3.6) | 3.2 (3.0 - 3.4) | 3.3 (3.1 - 3.5) | 3.2 (3.1 - 3.4) |
| SB injunctive norm | 4.0 (3.8 - 4.1) | 4.1 (3.9 - 4.3) | 3.9 (3.8 - 4.1) | 4.1 (4.0 - 4.2) | 4.0 (3.9 - 4.1) |
| SB intention | 3.8 (3.5 - 4.1) | 3.6 (3.3 - 3.9) | 3.7 (3.2 - 4.2) | 3.7 (3.3 - 4.2) | 3.7 (3.4 - 4.1) |
| SB outcome expectations | 4.5 (4.4 - 4.6) | 4.3 (4.2 - 4.4) | 4.4 (4.2 - 4.6) | 4.4 (4.3 - 4.6) | 4.4 (4.3 - 4.5) |

## Graphical presentation

Next, we present results graphically, to give the reader a richer perspective than from what can be gauged from considering summary statistics only.

### Activity during the day

From Figure 1, we can see that the patterns of activity within gender and intervention allocation groups are similar, but the IT track differs in their activity pattern to an extent.

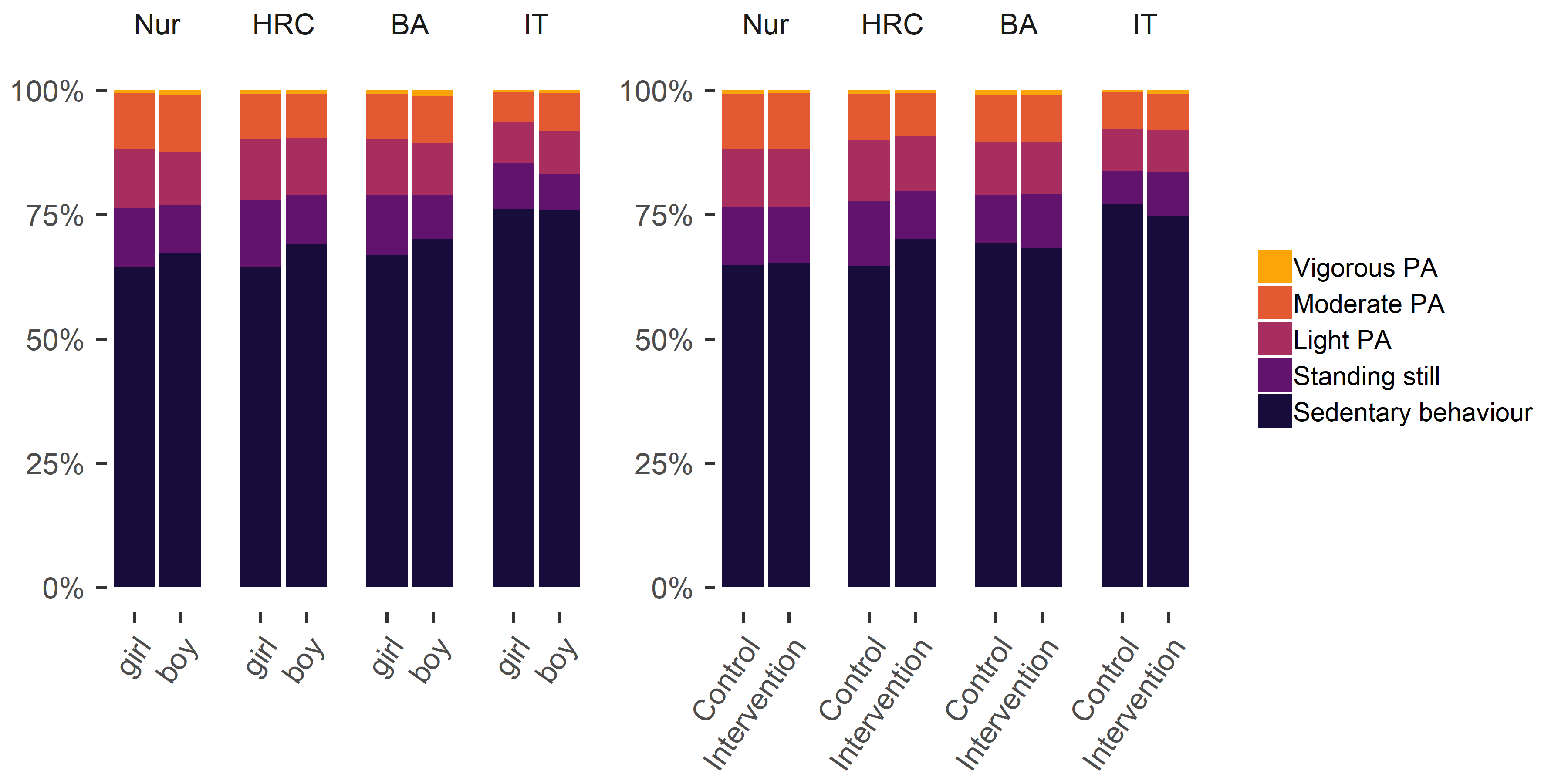


Figure 1 Stacked bar plot drawn with R package ggplot, showing averaged proportions of activity during an average day, separated by gender, arm and educational track. Nur = Practical nurse, HRC = Hotel, restaurant and catering, BA = Business and administration, IT = Information and communications technology.

The plot shows the averages aggregated across individuals and, while being clear, hides variability in the activity types. For example, while the average portion of measurement time spent in sedentary behaviour was 68%, almost every fourth (23%) participant was sedentary more than 75% of the time.

Figure 2 displays a density plot. It can be read like a histogram, but the shape is not dependent on the bar width, which is often set by software defaults and may not reflect the research needs at hand. The density curve also helps illustrate differences across groups. This graph, in comparison to the previous one, reveals variability and potential distributional differences, but does not show individual observations which determine the height of the curve.

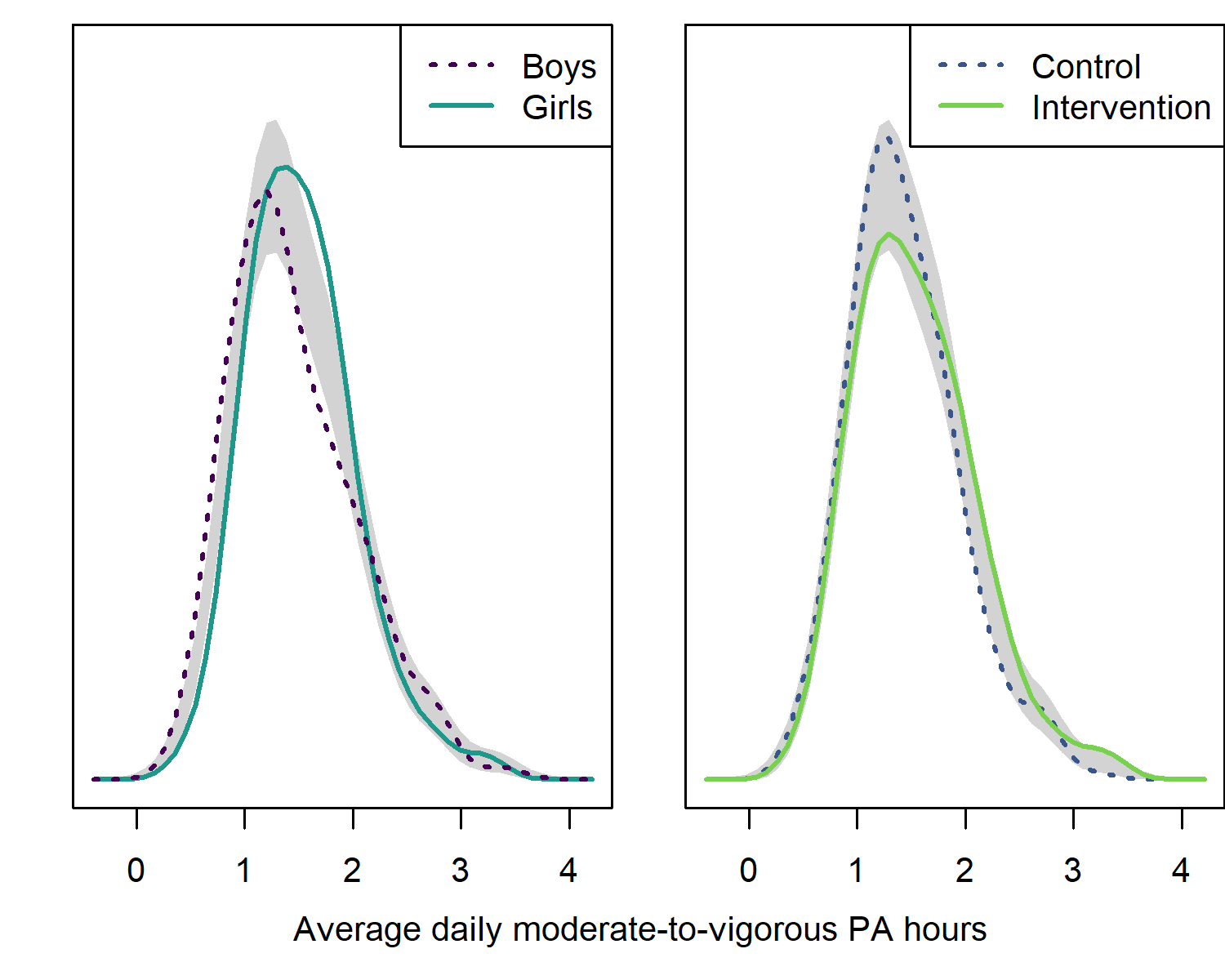


Figure 2 Kernel density plot drawn with R package sm, showing accelerometer-measured moderate-to-vigorous PA minutes, separated by gender and arm. Grey region depicts bootstrapped area of equivalence, given independent observations.

We can see, that girls seem to be slightly more active than boys (mean 92 vs. 87 minutes), or more specifically, there are more girls who reached an average of 1.5-2 hours of MVPA, and less of them who were lower. Although the grey band does not take clustering in classes, schools and educational tracks into account, it provides a heuristic for determining divergences between groups. Given that boys are generally more active than girls [3], this warrants a closer inspection, as presented by the density plots in figure 3.

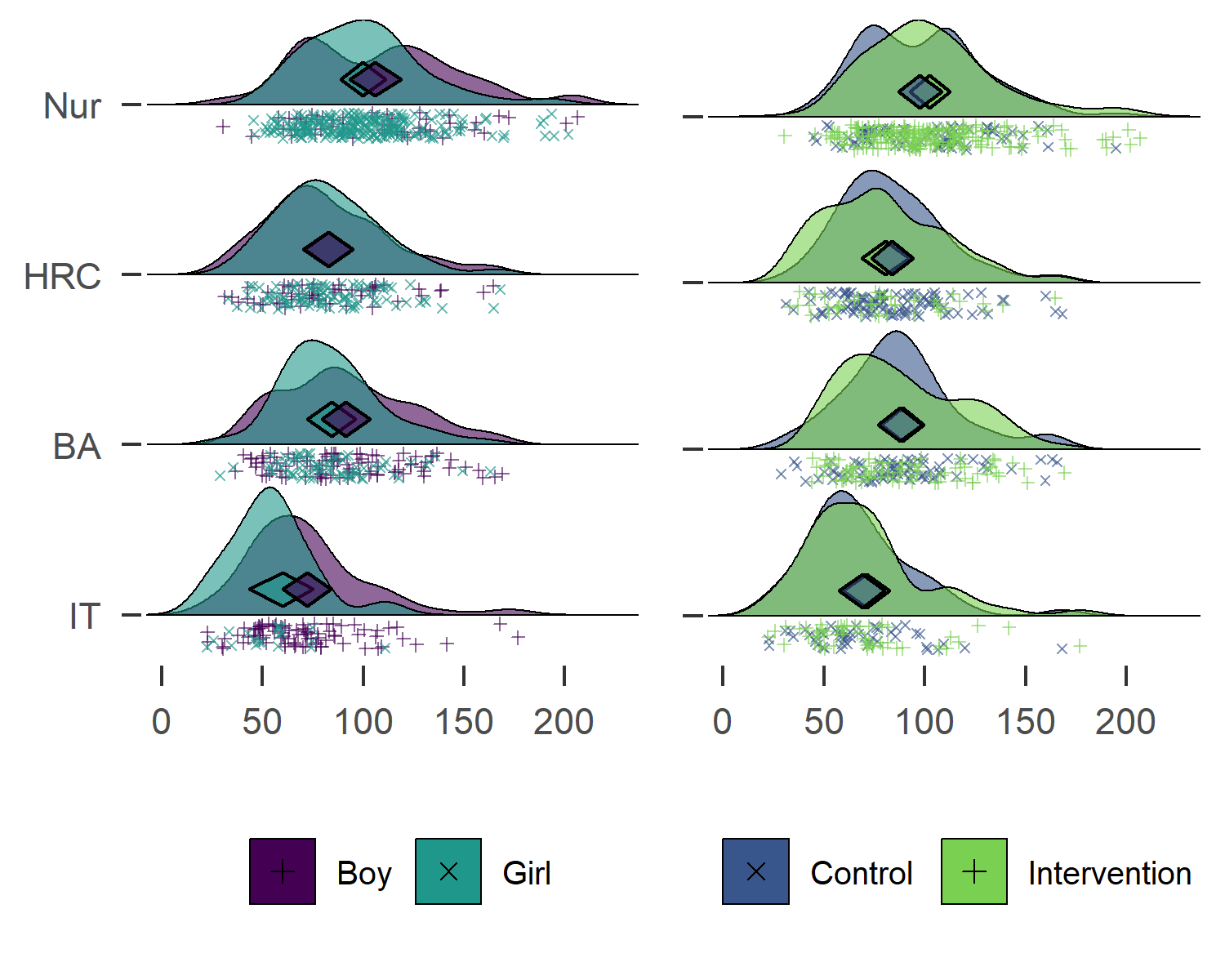


Figure 3 Raincloud ridge plot combined with a diamond plot, drawn with R packages ggridges and userfriendlyscience, showing hours of moderate-to-vigorous PA for different educational tracks. Midpoints of diamonds indicate means, endpoints 95% credible intervals. Individual observations are presented under the density curves, with random scatter on the y-axis to ease inspection. Nur = Practical nurse, HRC = Hotel, restaurant and catering, BA = Business and administration, IT = Information and communications technology.

As can be seen from the x-axis placement of diamonds in figure 3, participants who study to be practical nurses are the most active, followed by HRC students and BA students, with the IT track being the least active, though there is considerable variation within tracks. This explains the difference in MVPA among girls and boys: the practical nurse track is the largest, and its students are most active, as well as mostly girls. The information technology students are the least active, but mostly boys.

In sum, boys were very actually slightly more active in most of the tracks, although mean differences in minutes (7.40 for Practical nurse, -0.40 for Hotel, restaurant and catering, 7.50 for Business and administration, and 17.10 for Information and communications technology) are small. In spite of this, girls appear more active in the aggregate. This is also known as the Simpson’s paradox, and is best approached by visualising data (see [54] for an introduction).

Similar plots for all primary outcome variables can be found in the supplement. In brief, regardless of track, boys reported more days with at least 30 minutes of MVPA, while reporting more e.g. gym training, which was more strongly connected to the survey-measured MVPA than the accelerometer-measured one. Accelerometer measurement also indicated, that boys were more sedentary and interrupted sitting less often than girls. These results are also described in the supplement.

### Behaviour change technique usage

There were no clear differences between frequency-dependent BCTs between genders and intervention allocation, as shown in Figure 4.

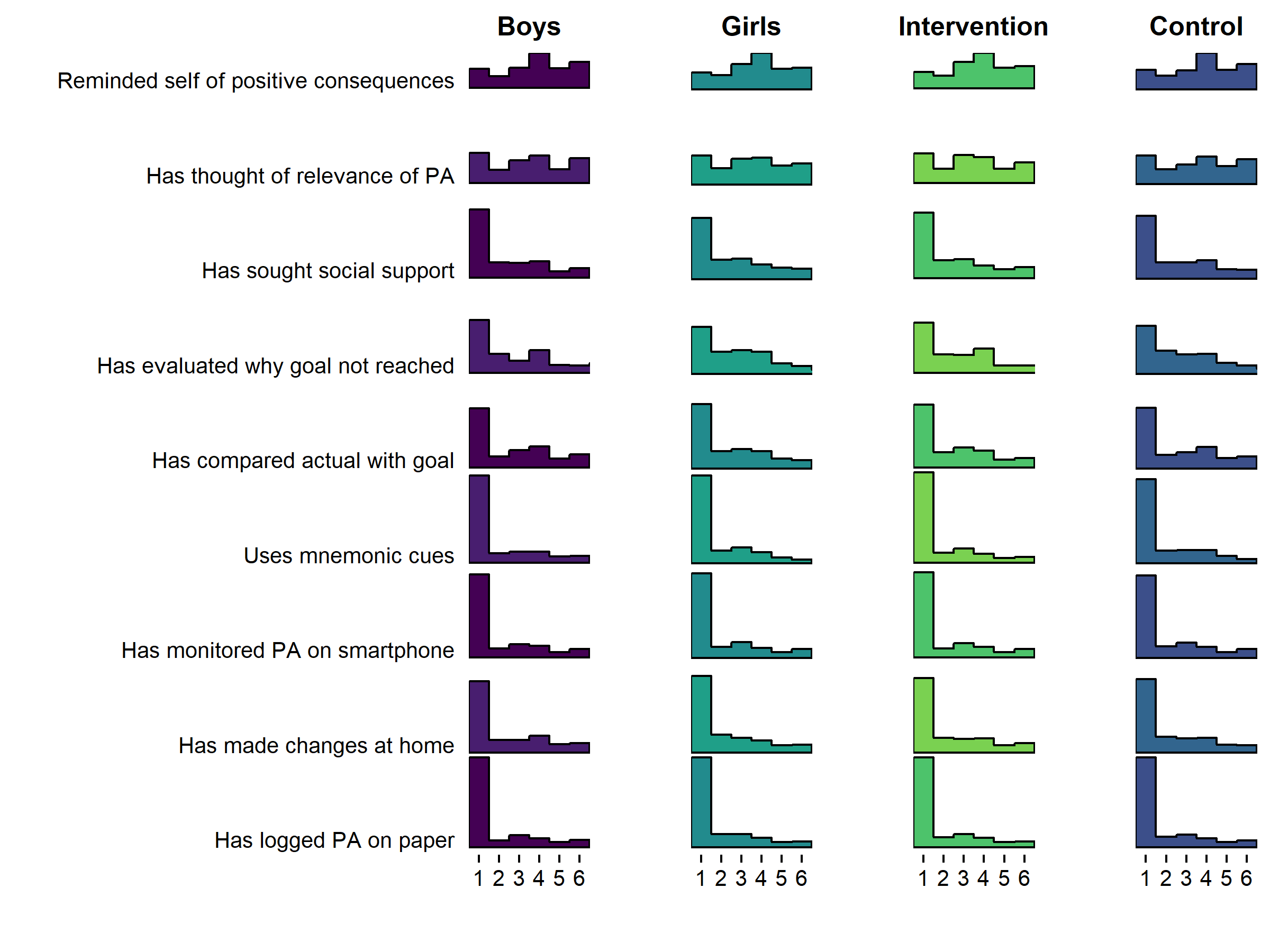


Figure 4 Histogram drawn with R package ggridges, showing self-reported use of frequency-dependent BCTs (1 = Not once … 6 = Daily).

Inspecting Figure 4 allows us to realise that the most frequent response is 1, indicating non-use of that BCT. In fact, a large number of BCTs seem to indicate a composite distribution, where one population reports never using the BCT, and another is seems normally distributed around the middle of the scale.

The same can be seen when observing agreement-dependent BCTs, as presented in Figure 5.

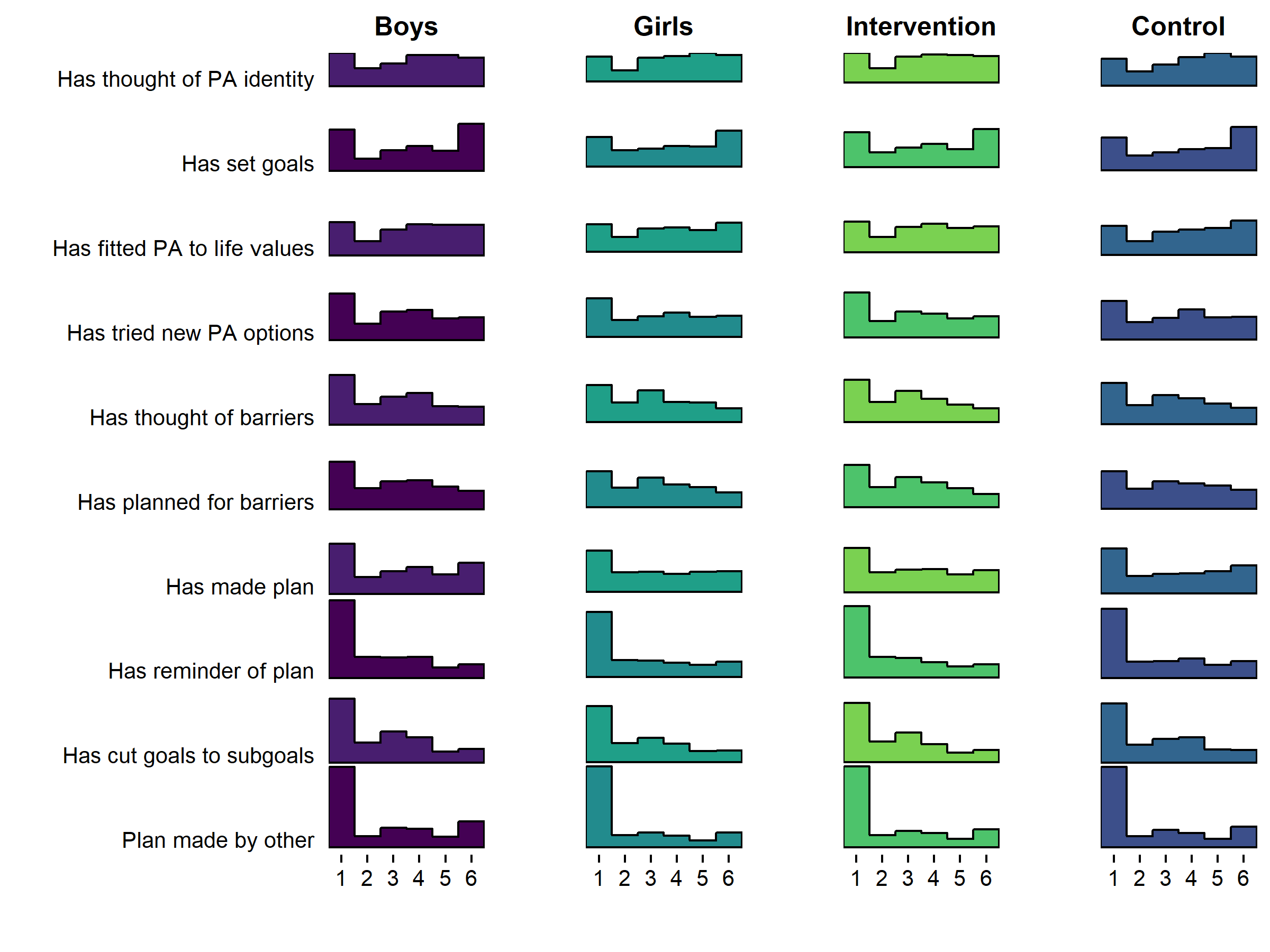


Figure 5 Histogram drawn with R package ggridges, showing self-reported use of agreement-dependent BCTs (1 = Not at all true … 6 = Completely true).

## Demonstration of network analysis

We can observe from figure 6, that after conditioning on BCT use, motivation and the two MVPA measures, a similar structure to the bivariate correlations is produced. One difference is, that the link from BCT usage to MVPA accelerometer disappears. Bootstrap analysis of stability indicated, that in repeated sampling, the connection is estimated as non-zero less than half of the time (see supplement, section Networks and correlations). **EIKO**

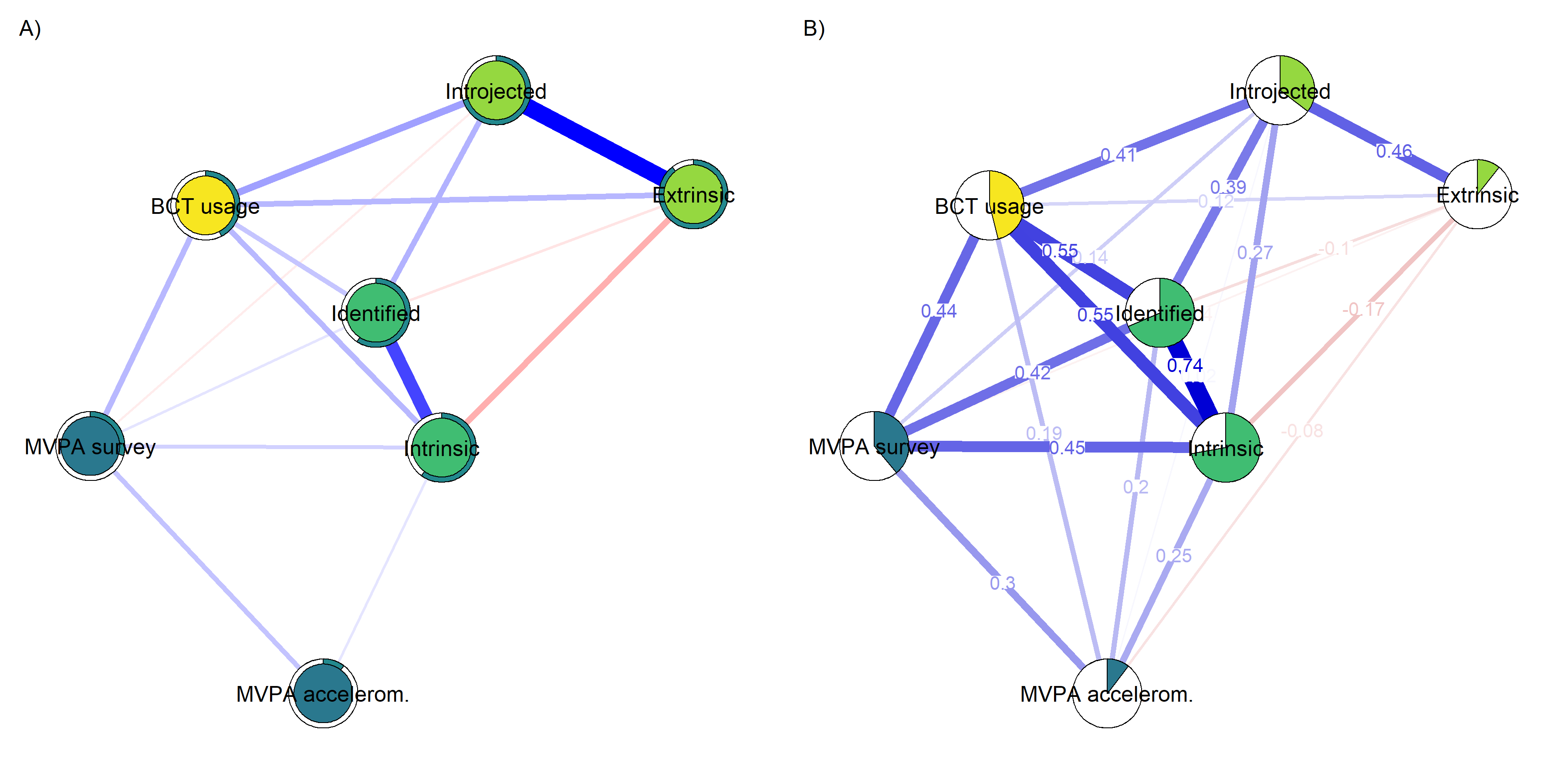


Figure 6 Network models drawn with packages mgm and qgraph; blue lines are for positive relationships, red for negative. A) Mixed graphical model with LASSO regularisation and model selection by cross-validation. Plot shows the conditional dependence relationships between the variables of interest (edges between nodes) and predictability (coloured proportion of the node’s outer rim). B) Bivariate correlation network showing the raw correlation structure. The proportion coloured with the pie indicates the mean of that variable, as % of theoretical maximum (in case of accelerometer-measured MVPA: proportion of mean daily accelerometer wear time).

The outer rim of the nodes in the plot on the left-hand side shows that extrinsic regulation is largely predictable (89.5%) by the connected nodes, whereas the accelerometer-measured MVPA is barely predictable at all (10.2%).

# Conclusions

This study investigated the baseline characteristics of the Let’s Move It trial cohort, making use of modern tools to exhaustively report the analyses performed and the analytical choices made.

Contrary with both international and Finnish evidence, girls performed more PA than boys in this sample. This could be explained by the practical nurse track being most active and mostly female; i.e. after accounting for track, no gender differences in accelerometer-measured MVPA could be seen. Further, boys reported doing more MVPA than girls, and the accelerometer-measurement implied they were also more sedentary and interrupted sitting less often. Intervention and control groups were similar in their accelerometer-measured MVPA. This observation supports the decision of randomising the educational tracks, such that all tracks were represented in both groups.

To our knowledge, this is one of the first studies to measure the use of potential behaviour change techniques comprehensively already at the trial baseline. As can be expected, many people indeed do use BCTs even before the intervention takes place. The results reveal that in the past three weeks, most participants had not used self-regulation related BCTs such as planning, problem solving or goal setting, which on the other hand have been indicated to be useful techniques for PA self-management [55].

Comprehensive, transparent reporting of results leads to a vast amount of information to be presented; visual exposition is thus vital. Visualising distributions makes it salient, how much variability there is among study participants, as well as how we should be wary of applying statistical techniques which rely on distributional assumptions.

The results of the network analysis .xxxxx It should be noted that while network analysis may be a useful tool for exploring and clarifying empirical relationships in novel areas and variables with little existing theory, the usefulness of this analytical method in testing and elaborating on existing theories of behaviour change has not yet systematically been considered. This may be a fruitful avenue. **EIKO**

Although some variables such as intention showed relatively high baseline levels in a large number of participants, we do not expect obvious ceiling effects to emerge in the PA psychosocial determinants. Due to the phenomenon of the practical nurse track being most active and mostly girls, potential gender differences in intervention results should be interpreted with caution.

The study and the LMI trial has a number of strengths, and will provide important information regarding ways to influence adolescent activity behaviour. Firstly, the recruitment of the baseline sample has succeeded well, with only x% **NELLI** of participants refusing to participate - this number has been considerably higher in other studies in this target group. Secondly, the study used also objective measures of PA and SB, shedding additional light to activity patterns, which has been commonly studied with self-reports only. Third, we have demonstrated state-of-the-art opportunities to visualise data and make it more understandable, and provided code as well as references to R packages, which we hope the reader will be able to make use of.

Current practice of evaluating determinant importance often relies on bivariate correlations [56]. This approach is limited in that when estimating a large number of correlations, one inevitably ends up with a potentially large amount of spurious relationships – in addition, it is unclear how many tests one should correct multiple testing for when doing exploratory analyses [18]. Network analysis tools provide easy-to-use regularisation techniques, and can be applied to the question of determinant relevance, demonstrating relationships between constructs in a readily graspable fashion. As the method controls for all the variables in the model, [watch out for colliders, consider if including cognitive/motivational nodes actually makes sense, be aware that interpretation may become difficult, …?] **EIKO**

The study also has limitations. It should be noted that the 7-day accelerometry, even with more than 4 days of over 10 hours wear time, is still an approximation of a participant’s true habitual physical activity and sedentary behaviour in their daily life – not an errorless measure of it, and does not capture all forms of activity.

The type of supplement used for this manuscript allows for presenting a lot, but not all, information due to resource considerations. One of the reader groups not fully considered are researchers and educators, who wish to use this data to guide intervention design. We would like to point out that the results, like most of the research in the area, only provide a group-level snapshot of a wide variety of constantly unfolding dynamical processes. Few individual participants are described by the group-level summary statistics: In fact, using Daniels’ [57] definition of an “approximately average individual” as falling in the middle 30% of the range of values, only 2% of participants can be considered “average” on all of the primary outcome measures (see supplement plots, section “Informativeness of averages”). Intervention designers looking at this cohort to choose to-be-targeted determinants for their study, may want to consider applying clustering techniques on the data once it becomes publicly available.

In conclusion, this analysis of baseline data from the Let’s Move It intervention trial indicates that randomization did not result in highly disproportionate groups, i.e. the differences between control and intervention groups were minimal – although, in the case of complex systems, even small differences may proliferate and lead to group imbalances [58]. It also has shown, that the vocational school students in Finland are hardly a homogenous group, but differ in many regards by their chosen educational track. Finally, graphical methods of presenting descriptive data are an important addition to traditional tables displaying means and standard deviations, and transparent sharing of analyses and analytical choices is imperative for increasing confidence in research findings.

### List of abbreviations

MVPA

### Declarations

### Ethics approval and consent to participate

The research proposal was reviewed by the Ethics Committee for Gynaecology and Obstetrics, Pediatrics and Psychiatry of the Hospital District of Helsinki and Uusimaa (decision number 367/13/03/03/2014).

### Availability of data and materials

The analysis data will be available at [OSF storage] on 1st January 2019. All analyses and code are available at <https://git.io/fNHuf> (permalink at [REF], repository at [REF]).

### Competing interests

The authors declare that they have no competing interests.

### Authors’ contributions

MH analyzed and interpreted the patient data regarding the hematological disease and the transplant. RH performed the histological examination of the kidney, and was a major contributor in writing the manuscript. TV was responsible for planning and analysing the PA and SB measured from data collected with accelerometer. All authors read and approved the final manuscript.

### Funding

MH was supported by Academy of Finland (grant number 295765). NH was supported by an Academy of Finland Research Fellowship (grant number 285283). The data were collected in a project funded by Ministry for Education and Culture, Sports Science projects (grant number OKM/81/626/2014).

### Acknowledgements

We would like to thank participating schools, their staff and students, as well as the numerous people who have helped in study design and data collection.

# References

1. Elgar FJ, Pförtner T-K, Moor I, De Clercq B, Stevens GWJM, Currie C. Socioeconomic inequalities in adolescent health 2002–2010: A time-series analysis of 34 countries participating in the Health Behaviour in School-aged Children study. The Lancet. 2015;385:2088–95. doi:[10.1016/S0140-6736(14)61460-4](https://doi.org/10.1016/S0140-6736(14)61460-4).

2. Dieleman JL, Sadat N, Chang AY, Fullman N, Abbafati C, Acharya P, et al. Trends in future health financing and coverage: Future health spending and universal health coverage in 188 countries, 2016–40. The Lancet. 2018;391:1783–98.

3. Husu P, Vähä-Ypyä H, Vasankari T. Objectively measured sedentary behavior and physical activity of Finnish 7-to 14-year-old children–associations with perceived health status: A cross-sectional study. BMC public health. 2016;16:338.

4. Mäkelä K, Kokko S, Kannas L, Villberg J, Vasankari T, Heinonen JO, et al. Physical Activity, Screen Time and Sleep among Youth Participating and Non-Participating in Organized Sports: The Finnish Health Promoting Sports Club (FHPSC) Study. Advances in Physical Education. 2016;6.

5. van Sluijs EM, Skidmore PM, Mwanza K, Jones AP, Callaghan AM, Ekelund U, et al. Physical activity and dietary behaviour in a population-based sample of British 10-year old children: The SPEEDY study (Sport, Physical activity and Eating behaviour: Environmental Determinants in Young people). BMC public health. 2008;8:388.

6. Hankonen N, Heino MT, Kujala E, Hynynen S-T, Absetz P, Ara’ujo-Soares V, et al. What explains the socioeconomic status gap in activity? Educational differences in determinants of physical activity and screentime. BMC public health. 2017;17:144. <https://bmcpublichealth.biomedcentral.com/articles/10.1186/s12889-016-3880-5>. Accessed 27 May 2017.

7. Hynynen S-T, van Stralen MM, Sniehotta FF, Ara’ujo-Soares V, Hardeman W, Chinapaw MJM, et al. A systematic review of school-based interventions targeting physical activity and sedentary behaviour among older adolescents. Int Rev Sport Exerc Psychol. 2016;9:22–44.

8. Rogers PJ. Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions. Evaluation. 2008;14:29–48.

9. Moore GF, Audrey S, Barker M, Bond L, Bonell C, Hardeman W, et al. Process evaluation of complex interventions: Medical Research Council guidance. BMJ. 2015;350:h1258.

10. Hankonen N, Heino MTJ, Hynynen S-T, Laine H, Ara’ujo-Soares V, Sniehotta FF, et al. Randomised controlled feasibility study of a school-based multi-level intervention to increase physical activity and decrease sedentary behaviour among vocational school students. International Journal of Behavioral Nutrition and Physical Activity. 2017;14.

11. Hankonen N, Heino MTJ, Araujo-Soares V, Sniehotta FF, Sund R, Vasankari T, et al. “Let’s Move It” a school-based multilevel intervention to increase physical activity and reduce sedentary behaviour among older adolescents in vocational secondary schools: A study protocol for a cluster-randomised trial. BMC Public Health. 2016;16:451–66.

12. Cleveland WS. Visualizing data. Hobart Press; 1993.

13. Tukey JW. Exploratory data analysis. Reading, Mass.; 1977.

14. Board of Governors of the Federal Reserve System, Chang AC, Li P. Is Economics Research Replicable? Sixty Published Papers from Thirteen Journals Say "Usually Not". Finance and Economics Discussion Series. 2015;2015:1–26.

15. Bond TN, Lang K. The Sad Truth About Happiness Scales: Empirical Results. Working Paper. National Bureau of Economic Research; 2018.

16. Gigerenzer G. Statistical Rituals: The Replication Delusion and How We Got There. Advances in Methods and Practices in Psychological Science. 2018;2515245918771329.

17. Kepes S, McDaniel MA. How Trustworthy Is the Scientific Literature in Industrial and Organizational Psychology? Industrial and Organizational Psychology. 2013;6:252–68.

18. Nosek BA, Ebersole CR, DeHaven AC, Mellor DT. The preregistration revolution. Proceedings of the National Academy of Sciences. 2018;201708274.

19. Nosek BA, Spies JR, Motyl M. Scientific Utopia II. Restructuring Incentives and Practices to Promote Truth Over Publishability. Perspectives on Psychological Science. 2012;7:615–31.

20. Simmons JP, Nelson LD, Simonsohn U. False-Positive Psychology Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. Psychological Science. 2011;22:1359–66.

21. Smaldino PE, McElreath R. The natural selection of bad science. Open Science. 2016;3:160384.

22. Stodden V, Seiler J, Ma Z. An empirical analysis of journal policy effectiveness for computational reproducibility. Proceedings of the National Academy of Sciences. 2018;115:2584–9.

23. Nosek BA, Errington TM. Reproducibility in cancer biology: Making sense of replications. Elife. 2017;6:e23383.

24. R Core Team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2017.

25. Fagt S, Andersen LF, Anderssen SA, Becker W, Borodulin K, Fogelholm M, et al. Nordic Monitoring of diet, physical activity and overweight : Validation of indicators. Nordic Council of Ministers; 2012.

26. Vähä-Ypyä H, Vasankari T, Husu P, Suni J, Sievänen H. A universal, accurate intensity-based classification of different physical activities using raw data of accelerometer. Clinical physiology and functional imaging. 2015;35:64–70.

27. Vähä-Ypyä H, Vasankari T, Husu P, Mänttäri A, Vuorimaa T, Suni J, et al. Validation of cut-points for evaluating the intensity of physical activity with accelerometry-based mean amplitude deviation (MAD). PLoS One. 2015;10:e0134813.

28. Tremblay MS, Aubert S, Barnes JD, Saunders TJ, Carson V, Latimer-Cheung AE, et al. Sedentary Behavior Research Network (SBRN) process and outcome. International Journal of Behavioral Nutrition and Physical Activity. 2017;14:75.

29. Vähä-Ypyä H, Husu P, Suni J, Vasankari T, Sievänen H. Reliable recognition of lying, sitting, and standing with a hip-worn accelerometer. Scandinavian Journal of Medicine & Science in Sports. 2018;28:1092–102.

30. RStudio Team. RStudio: Integrated Development Environment for R. 2016.

31. R Core Team. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2018. <https://www.R-project.org/>.

32. Dalege J, Borsboom D, van Harreveld F, van den Berg H, Conner M, van der Maas HLJ. Toward a formalized account of attitudes: The Causal Attitude Network (CAN) model. Psychological Review. 2016;123:2–22.

33. Dalege J, Borsboom D, Harreveld F, Waldorp LJ, Maas HL. Network structure explains the impact of attitudes on voting decisions. Scientific reports. 2017;7:4909.

34. Mõttus R, Allerhand M. Why do traits come together? The underlying trait and network approaches. SAGE handbook of personality and individual differences. 2017;1:1–22.

35. Van Der Maas H, Kan K-J, Marsman M, Stevenson CE. Network Models for Cognitive Development and Intelligence. 2017.

36. Fried EI, van Borkulo CD, Cramer AO, Boschloo L, Schoevers RA, Borsboom D. Mental disorders as networks of problems: A review of recent insights. Social Psychiatry and Psychiatric Epidemiology. 2017;52:1–10.

37. Briganti G, Kempenaers C, Braun S, Fried EI, Linkowski P. Network analysis of empathy items from the interpersonal reactivity index in 1973 young adults. Psychiatry Research. 2018;265:87–92.

38. Dalege J, Borsboom D, van Harreveld F, van der Maas HL. Network analysis on attitudes: A brief tutorial. Social psychological and personality science. 2017;8:528–37.

39. Epskamp S, Fried EI. A Tutorial on Regularized Partial Correlation Networks. Psychological Methods. 2018. doi:[10.1037/met0000167](https://doi.org/10.1037/met0000167).

40. Epskamp S, Borsboom D, Fried EI. Estimating Psychological Networks and their Stability: A Tutorial Paper. arXiv preprint arXiv:160408462. 2016.

41. Costantini G, Epskamp S, Borsboom D, Perugini M, Mõttus R, Waldorp LJ, et al. State of the aRt personality research: A tutorial on network analysis of personality data in R. Journal of Research in Personality. 2015;54:13–29.

42. Costantini G, Richetin J, Preti E, Casini E, Epskamp S, Perugini M. Stability and variability of personality networks. A tutorial on recent developments in network psychometrics. Personality and Individual Differences. 2017.

43. Fried EI, Cramer AO. Moving forward: Challenges and directions for psychopathological network theory and methodology. Perspectives on Psychological Science. 2017;12:999–1020.

44. Borsboom D, Mellenbergh GJ, van Heerden J. The theoretical status of latent variables. Psychological Review. 2003;110:203–19.

45. Bringmann LF, Eronen MI. Don’t blame the model: Reconsidering the network approach to psychopathology. Psychological review. 2018;125:606–33.

46. Tibshirani R. Regression shrinkage and selection via the lasso: A retrospective. Journal of the Royal Statistical Society: Series B (Statistical Methodology). 1996;73:273–82.

47. Haslbeck JMB, Waldorp LJ. How well do network models predict observations? On the importance of predictability in network models. Behavior Research Methods. 2018;50:853–61.

48. Haig BD. Tests of Statistical Significance Made Sound. Educational and Psychological Measurement. 2016.

49. Nickerson RS. Null hypothesis significance testing: A review of an old and continuing controversy. Psychological Methods. 2000;5:241–301.

50. de Groot AD. The meaning of “significance” for different types of research [translated and annotated by Eric-Jan Wagenmakers, Denny Borsboom, Josine Verhagen, Rogier Kievit, Marjan Bakker, Angelique Cramer, Dora Matzke, Don Mellenbergh, and Han L. J. van der Maas]. Acta Psychologica. 2014;148:188–94.

51. Lakens D, Adolfi FG, Albers CJ, Anvari F, Apps MAJ, Argamon SE, et al. Justify your alpha. Nature Human Behaviour. 2018;2:168–71.

52. Lai MHC, Kwok O-m. Estimating Standardized Effect Sizes for Two- and Three-Level Partially Nested Data. Multivariate Behavioral Research. 2016;51:740–56.

53. Lai MHC, Kwok O-m, Hsiao Y-Y, Cao Q. Finite population correction for two-level hierarchical linear models. Psychological methods. 2018;23:94.

54. Kievit RA, Frankenhuis WE, Waldorp LJ, Borsboom D. Simpson’s paradox in psychological science: A practical guide. Frontiers in Psychology. 2013;4.

55. Michie S, Abraham C, Whittington C, McAteer J, Gupta S. Effective techniques in healthy eating and physical activity interventions: A meta-regression. Health Psychology. 2009;28:690–701.

56. Crutzen R, Peters G-JY, Noijen J. Using Confidence Interval-Based Estimation of Relevance to Select Social-Cognitive Determinants for Behavior Change Interventions. Frontiers in Public Health. 2017;5.

57. Daniels GS. The" Average Man"? AIR FORCE AEROSPACE MEDICAL RESEARCH LAB WRIGHT-PATTERSON AFB OH; 1952.

58. Rickles D. Causality in complex interventions. Medicine, Health Care, and Philosophy. 2009;12:77–90.