Visualisation and network analysis of physical activity and its determinants: Demonstrating opportunities in analysing baseline associations in the Let's Move It trial

Matti T. J. Heino[1], Eiko Fried[2], Reijo Sund[3], Ari Haukkala[1], Keegan Knittle[1], Katja Borodulin[1], Antti Uutela[1], Vera Araujo-Soares[4], Tommi Vasankari[5], & Nelli Hankonen[1]

[1] Faculty of Social Sciences, University of Helsinki, PO Box 54, 00014 University of Helsinki, Finland

[2] Department of Clinical Psychology, Leiden University, Wassenaarseweg 52, 2333 AK, Leiden, The Netherlands

[3] Faculty of Health Sciences, University of Eastern Finland, PO Box 1627, 70211 Kuopio, Finland

[4] Institute of Health and Society, Medical Faculty, Newcastle University, Baddiley-Clarke Building, Richardson Road NE2 4AX, United Kindom

[5] UKK institute for Health Promotion Research, Kaupinpuistonkatu 1, 33500 Tampere, Finland

Trial Registration Number: ISRCTN10979479. Registered retrospectively: 31.12.2015. The authors declare that they have no competing interests.

Author Note

Correspondence concerning this article should be addressed to Matti T. J. Heino, Faculty of Social Sciences, University of Helsinki, PO Box 54, 00014 University of Helsinki, Finland. E-mail: matti.tj.heino@gmail.com

Abstract

Visualisation and network analysis of physical activity and its determinants: Demonstrating opportunities in analysing baseline associations in the Let's Move It trial

Background: The Let's Move It intervention aimed to increase physical activity (PA) and reduce sedentary behaviours (SB) among vocational school students. This study visually explores the baseline cohort of a cluster-randomised trial testing the intervention, making use of recent developments in software and network analyses to go beyond tabulated results.

Methods: At baseline, 1166 adolescents, distributed across 6 school clusters and four educational tracks, completed measures of PA and SB, theoretical predictors of these behaviors, and body composition. Data were tabulated and visualised, and network analyses explored relations between predictor variables and outcomes.

Results: Average daily moderate-to-vigorous PA was 65min (CI95: 57min-73min), and SB 8h44min (CI95: 8h04min-9h24min), with 25.8 (CI95: 23.5-28.0) interruptions to sitting. The four educational tracks were heterogeneous, and self-reported behaviour change technique use was generally low. Network analyses revealed direct relationships between PA and behavioural experiments, planning, and autonomous motivation.

Conclusions: Presenting complex data visually unearths dynamics involved in complex multi-causal systems beyond what is possible with tabulated results alone. Researchers should engage with these new methods and publish extended descriptive expositions as website supplements, which can increase the speed and quality of scientific communication.

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

Visualisation and network analysis of physical activity and its determinants: Demonstrating opportunities in analysing 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 among individuals with low socioeconomic status (SES) (Elgar et al., 2015, p. @dielemanTrendsFutureHealth2018). Patterns of low PA among adults begin earlier in the life course, with evidence that declines in PA and increases in SB begin during childhood and adolescence (Husu et al., 2016; Mäkelä et al., 2016). This highlights the need for further research into interventions to improve PA and SB among adolescents.

As adolescents spend a significant amount of their time in schools, the school setting provides valuable opportunities for PA and SB interventions (van Sluijs et al., 2008). The Let's Move It intervention aimed to reduce SB and increase PA among adolescents in vocational schools, and was developed using stakeholder input and co-creation with target group representatives, as well as theories and empirical evidence from behavioural science (Hankonen, Heino, Kujala, et al., 2017; Hynynen et al., 2016). Contrary to typical school-based interventions with relatively homogeneous participants, this trial was carried out in vocational schools with distinct and varied educational tracks (i.e. practical nurse, business information and communication technology, business administration, and hotel, restaurant and catering). Understanding the implications of these distinct tracks on the way participants engage in both PA and SB will support a better understanding of the individual and contextual determinants of behaviour and more informed interpretations of the results obtained in the trial.

The hypothesised programme theories (Moore et al., 2015; Rogers, 2008) for changing PA and SB differed from one another. In order to increase PA, one needs to make a conscious effort and implement self-regulatory skills (e.g. action planning and overcoming barriers to PA) to make optimal use of opportunities. The Let's Move it intervention places a particular emphasis on helping adolescents understand and use techniques to manage their motivation and behaviour (see also Hankonen, Heino, Hynynen, et al. (2017) and Hankonen (2018)). To date, there is little knowledge about how the use of these techniques links to each other, and it would be important to examine these links empirically. The theoretical model for changing SB, on the other hand, is more driven by environmental opportunities, such as having the option to stand up during class.

In order to increase moderate-to-vigorous-intensity PA, the Let's Move It intervention targeted several behavioural determinants, including 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/). To reduce total SB and introduce breaks in SB, the programme aimed to change the school environment by training teachers in the use of active teaching techniques and altering physical choice architecture in classrooms (Köykkä et al., 2018). The intervention also included poster campaigns in schools, a website, and materials to target community actors and parents (Köykkä et al., 2018). More information of the content of the intervention and the development of it is reported elsewhere (Hankonen, Absetz, & Araujo-Soares, 2019; Hankonen et al., 2016; Hankonen, Heino, Hynynen, et al., 2017).

It has long been a standard recommendation for quantitative analyses to investigate data visually as a core precursor of conducting statistical analyses (Cleveland, 1993; Tukey, 1977). However, in social and life sciences, such visualisations are rarely 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 (Trafimow, Wang, & Wang, 2018). Medians, modes, skewness and kurtosis provide helpful additional information, but human cognition places limits on evaluating these statistics simultaneously, especially when comparing groups of observations. For example, two distributions can have different means but the same mode, different modes but the same mean, or the same mean and standard deviation but a meaningful skew. Summary statistics conventionally calculated from the data leave important distributional properties uncovered, as illustrated in recent discussions on the inadequacy of bar plots (Saxon, 2015; Weissgerber, Garovic, Savic, Winham, & Milic, 2016; Weissgerber, Milic, Winham, & Garovic, 2015).

Data visualisations are crucial supplements to large numerical tables of descriptive statistics (Tay, Parrigon, Huang, & LeBreton, 2016). With visualisations, researchers can communicate large amounts of information – including the associated uncertainty – in an accessible format, without requiring extensive mathematical expertise from the reader. This is important for researchers who intend to build on previous results (Chalmers & Glasziou, 2009). Such practices may reduce problems that have led to the recent loss of confidence in the reproducibility and replicability of research findings (Gigerenzer, 2018; Kepes & McDaniel, 2013; Nosek, Ebersole, DeHaven, & Mellor, 2018; Nosek, Spies, & Motyl, 2012; Simmons, Nelson, & Simonsohn, 2011; Smaldino & McElreath, 2016). Fully open data sharing would be ideal, but this is not always possible due to privacy concerns (Expert Advisory Group on Data Access, 2015) and, at the time of writing, remains a lamentably rare practice (Vanpaemel, Vermorgen, Deriemaecker, & Storms, 2015). In addition, open data does not necessarily accommodate stakeholders with low technical expertise in data analysis and visualisation, such as clinicians, patients and policy makers; see Hallgren, McCabe, King, and Atkins (2018), p. 2.

Three recent developments give impetus to a new approach. First, many journals now allow publication of supplementary online materials, which circumvents both word and figure restrictions of traditional manuscripts. Second, statistical software such as R (R Core Team, 2015) has recently become increasingly mainstream among applied researchers, with many free tutorials available online, opening the door for a variety of data visualisation techniques.

Third, novel statistical methods in social and health psychology, such as psychological network analysis, may help to understand relationships between variables by making better use of visual representations of associations.

The aims of this paper are to describe central characteristics of the Let's Move It trial baseline cohort, focusing on co-primary outcomes and other activity measures (as measured by accelerometry) of the trial both arms, genders and educational tracks in both trial arms. A further aim is to describe psychological and social correlates, as well as hypothesised determinants of the intervention's effect on moderate-to-vigorous PA (MVPA), with detailed visualisations of the dataset provided in an extensive supplementary website. As a sub-aim, we also investigate the network of relationships between MVPA, quality of motivation and BCT use at baseline. We provide all code as open source scripts, so that other researchers can use those scripts as templates to visualise their own datasets in a format that requires no special skills or tools to view.

Methods

This study analyses baseline data from a cluster-randomised controlled trial testing Let's Move It, a complex whole-school system multi-level intervention conducted in Finnish vocational schools. Details of the Let's Move It trial have been described in the study protocol (Hankonen et al., 2016). At baseline, consenting participants in both intervention and control groups answered an electronic survey, underwent bioimpedance measurements and were instructed to wear an accelerometer for seven consecutive days. The baseline data collection started in January 2015 and ended in April 2016.

Six school units were included in the study. There were four educational tracks in the schools from which students were recruited: 1. Practical Nurse (Nur), 2. Hotel, Restaurant and Catering (HRC), 3. Business and Administration (BA), and 4. Information and

Communications Technology (IT). Schools were paired so that there would be matching numbers of students from each educational track for both members of the pair. Blinded randomisation by a statistician was then conducted so that a random member of each pair was selected as intervention school, the other as control school (details reported in Hankonen et al. (2016)). Student participants were blind to allocation at baseline.

All conducted analyses and visualisations with accompanying code, can be found in the supplementary website at https://git.io/fNHuf (permalink at Heino and Sund (2019)), previously piloted in (Heino et al., 2018a). Source code to reproduce this manuscript (written with the R package `papaja` (Aust, 2014/2019)), and all its figures can be found at https://git.io/fptcC.

Measures

The measures are presented briefly, as they have been previously described in Hankonen et al. (2016), and all individual items of the scales are available in the supplementary website (see section https://git.io/fjfLw).

Primary outcome variables of the trial. In the LMI trial, there were multiple primary outcomes. The primary outcome for PA was moderate to vigorous PA (MVPA), measured by accelerometry and self-reports. Primary outcomes for sedentary behaviour (SB) were measured by accelerometry; they included time spent sitting or lying down, and the number of times sitting was interrupted during the day.

Self-reported MVPA. Self-reported MVPA was measured with two questions in accordance with the NordPAQ measurement (Fagt et al., 2012). The first question asked participants about the number of days during the last week in which they did more than 30 minutes of MVPA, the other probed the overall amount of MVPA (in hours) during the past seven days.

Accelerometer-measured MVPA and SB. No more than seven days after responding to the questionnaire, students were given an accelerometer to be worn on 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. Definitions of the parameters are described in detail in the supplementary website (section https://git.io/fjJNi).

Theoretical predictors of PA. The determinants 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. In other words, answering a single item of a specific scale sufficed. All items, response options, and descriptive statistics of scales are available in the supplementary website (section https://git.io/fAj0e); made using R package `codebook` (Arslan, forthcoming) for automatic dataset documentation.

Statistical analysis. We used RStudio (RStudio Team, 2015) 1.1.456 running R (Version 3.5.3; R Core Team, 2018) 3.5.3 for all our analyses and figures.

In our case (no confirmatory hypotheses), confidence intervals are more appropriate to report than p-values, as they provide readily interpretable values on the same scale as the original variable, accommodating inferences of practical relevance (Gardner & Altman, 1986; Nosek et al., 2018; Sterne, 2001; Wasserstein & Lazar, 2016). Hence, we omit explicit statistical testing from the tables.

Activity data was explored by utilising 100% stacked bar charts, which are useful when comparing proportions which add to 100%. MVPA data was, in addition, examined with augmented raincloud ridge plots to unveil distributional properties. Psychological and social determinants were examined with diamond plots (Peters, 2018), and heuristic effect sizes between means of intervention arms and genders transformed from Cohen's d to Pearson's r. Distributions of BCT use were examined with histograms.

Psychological network analysis was used to estimate and visualise relations among BCT use, motivation and MVPA. Such networks contain nodes (variables) and edges (statistical relationships between variables). Unlike in social network analysis, the connections are not directly observed, but are estimated. We used network models that estimate conditional dependence relations among a set of variables, which can be interpreted similarly to partial correlations. An edge between two variables implies that they are related after controlling for all other variables; the absence of an edge implies that the two variables are (conditionally) independent.

The Mixed Graphical Model uses regularisation, 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 (Haslbeck & Waldorp, 2015). Regularisation has the goal to avoid estimating spurious relationships among items (i.e. false positive relations), and results in a parsimonious network structure. The regularisation technique used here is the Least Absolute Shrinkage and Selection Operator (LASSO; Tibshirani (1996)), which shrinks all edges and sets very small edges to exact zero. A paper that explains lasso regularisation in network models in detail can be found elsewhere (Epskamp & Fried, 2018).

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 (Fried & Cramer, 2017). Identifying these determinants of importance can thus supplement traditional structural equation modeling (SEM) approaches. Network analysis has recently been taken up in many fields such as social psychology (Dalege et al., 2017a, 2016), personality (Mõttus & Allerhand, 2017), intelligence (Van Der Maas, Kan, Marsman, & Stevenson, 2017), psychopathology (Fried et al., 2017), and empathy research (Briganti, Kempenaers, Braun, Fried, & Linkowski, 2018), and is beginning to be applied for health behaviours on a broader scale. Several helpful tutorial papers aimed at empirical researchers are available (Costantini et al., 2015, 2017; Dalege et al., 2017b; Epskamp et al., 2018; Epskamp & Fried, 2018), and also exist for health psychology context in particular (Hevey, 2018).

To ease interpretation of the network analysis, we dichotomised the heavily skewed controlled motivation variable in such a way that 1 represents answers 3 (‘partly true for me’) or higher, and 0 the rest. In addition, BCT use variables were dichotomised by giving 0 if a person reports completely disagreeing with their statements, or never having used the technique, and 1 otherwise. A correlation matrix of the variables can be found in the supplement (https://git.io/fhAgk).

Findings

In this section, we first present data in traditional numeric tables, and follow up by augmenting them with graphical illustrations. Table 1 shows the main demographic variables of the cohort by educational track. Among 638 intervention arm participants, 80.5% (429/533) reported having been born in Finland. Among the 528 control arm participants, the percentage was 88.7% (423/477).

While on average the sample was relatively balanced on 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 (16.0%). Age ranged from 16 to 49, with the average age being 18.50. Altogether there were 190 (16%) students who reported being at least 20 years old.

Table 2 shows summary statistics for primary outcome variables with their intra-class correlations (ICCs) for class and school (see supplementary website, section https://git.io/fjIcc, for ICCs of all variables). The ICC can be interpreted as the proportion of the variable's variance accounted for by group membership.

At baseline, 63.6% students provided at least 4 days with a minimum of 10 hours per day of valid accelerometer data. On average, the participants reported engaging in at least 30 minutes of MVPA on 2.80 days a week. Accelerometer data indicated, that girls were as active as boys (mean 65 vs. 67 minutes). Given that boys are generally more active than girls (Husu et al., 2016), this result will be elaborated on below.

To give the reader a richer perspective than from what can be gauged from considering these summary statistics only, we present the results graphically in figure 1. We can see that the patterns of average baseline activity, as measured by the accelerometer, are similar within gender and intervention allocation groups. However, the charts reveal that the IT track is more sedentary compared to other tracks and that girls are actually *less* active in each educational track.

The plot shows the average activity types relative to measurement time, but hides variability around the averages. The graph does not depict, for example, that while the average portion of time spent in sedentary behaviour for the IT track was 72.0%, almost half (42.0%) of that track's participants were sedentary more than 75% of the time.

Zooming in on accelerometer-measured MVPA, table 3 gives us statistics – some of which more commonly reported, others less so – on the variable.

Figure 2 displays an augmented density plot, representing and elaborating on information from table 3. The density curves can be read like a histogram, but the shape is not dependent on the bar width. They also help illustrate differences across groups, revealing potential differences in variability and distribution shape.

As the diamonds in figure 2 illustrate, participants who study practical nursing are the most active, followed by HRC students and BA students, with the IT track being the least active. There is considerable variation within tracks though. This explains the gender difference in MVPA: the practical nurse track is the largest, and its students, mostly girls, are the most active. The IT students, mostly boys, are the least active.

In sum, boys did more MVPA in every educational track (mean differences in minutes: 12.80 for Practical nurse, 5.40 for Hotel, restaurant and catering, 11.30 for Business and administration, and 19.80 for IT). In spite of this, girls appear more active in the aggregate. This is also known as the Simpson's paradox, and is best investigated by visualising data (see Kievit, Frankenhuis, Waldorp, and Borsboom (2013) for an introduction). Examining the left side of figure 2 reveals the difference between boys and girls in MVPA, the difference between Practical nurse and IT tracks, the differences in gender composition, and differences in the amount of participants in each track. These, when taken together, contribute to a comprehensive understanding of the data.

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 self-reported MVPA than the accelerometer-measured one. Accelerometer measurement also indicated, that boys engaged in more sedentary time and interrupted sitting less often than girls (see supplementary website, sections https://git.io/fjvWv and https://git.io/fjvCj).

Theoretical determinants

In table 4 below, we present the means for the primary outcome variables by gender and trial arm.

In 14 of the 18 variables presented here, the mean of the control group is more favourable than that of the intervention group (average unadjusted advantage 1.91%). In figure 3, the results are visualised in a concise manner.

We can observe, for example, that SB descriptive norms are bimodal and thus the means are not representative of typical participants. In addition, several of the variables are skewed, which has implications on analytical choices as well as interpretations of the mean values.

Behaviour change technique usage. There were no clear differences in frequency-dependent BCT use between genders or arms (Figure 4).

Figure 4 tells 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 aforementioned forms can also be observed in the distributions of agreement-dependent BCTs, as presented in Figure 5.

Demonstration of network analysis

Figure 6 shows a LASSO regularised mixed graphical model of BCT use, motivation and the two MVPA measures. We can observe, that after taking into account all the other nodes in the network and regularising small connections to zero, autonomous motivation appears to serve as a link between many BCTs and MVPA. In fact, only having a plan made by someone else, and having tried out new ways to be physically active (during the past three weeks), are directly connected to either of the MVPA nodes. In addition, use of certain BCTs are coupled particularly closely: Comparatively strong links exist between goal setting and having an own PA plan, between identifying barriers and planning to overcome them (i.e. problem solving/coping planning), and between goal setting and an own PA plan (i.e. action planning). We can also see a triad, where reflecting positive consequences is connected to goal review, through having thought of personal reasons to do PA, as well as less strongly coupled social support and having made changes to home environment. Such connections can be understood as variables influencing each other, but can also be indicative of underlying latent variables (i.e., the three variables are causal consequences of a shared origin) (Molenaar, 2010).

Conclusions

This study investigated the baseline characteristics of the Let's Move It trial cohort, making use of modern tools to visualise key results and exhaustively report the analyses, findings and analytical choices made. We found high levels of sedentary behaviour in the sample, with heterogeneity across educational tracks. MVPA, motivation and BCT use were modeled as a network, which highlighted the relevance of autonomous motivation in associations between PA and BCT use.

In contrast to earlier international and Finnish data collected in the general population (e.g. Husu, Suni, et al. (2016)), girls performed slightly more PA than boys in this sample. This is due to the practical nurse track being most active and mostly female; in other words, after accounting for track, no meaningful gender differences in accelerometer-measured MVPA could be seen. Further, boys reported doing more MVPA than girls, and the accelerometer-measurement implied boys 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 pairing educational tracks in randomisation, such that all tracks were represented in both arms. The practical nurse track was simultaneously the largest, the most active and had the highest percentage of girls, which means that potential gender differences in eventual intervention results should be interpreted with caution.

To our knowledge, this is one of the first studies to measure the use of potential BCTs 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, many participants report not having 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 (Michie, Abraham, Whittington, McAteer, & Gupta, 2009). To our knowledge, this is also the first trial to measure the use of a range of BCTs among both control and intervention arm participants.

Comprehensive, transparent reporting of results leads to a vast amount of information to be presented: visual exposition is thus vital. Visualising distributions makes the variability among study participants more salient, which informs us about the distributional assumptions that underlie many common statistical techniques. Modern and traditional approaches to data visualisation also allow us to go further than just comparing means (Rousselet, Pernet, & Wilcox, 2017), and provide opportunities to avoid drawing false conclusions (e.g. in the case of Simpson's paradox) based on summary statistics alone.

The results of the network analysis highlight, how most naturally used BCTs – exceptions including having a plan made by someone else, and trying out new forms of PA – possibly require autonomous motivation to affect MVPA. This finding, if corroborated in longitudinal data, would support the theoretical framework of the intervention, which held autonomy support and behavioural experiments at the forefront. So far, network models have been largely used as a tool for exploring empirical relationships among variables, often with little existing theory (Fried et al., 2017; Mõttus & Allerhand, 2017). One could understand this as the first generation of network papers in psychology, and there have been recent calls for a second generation that is confirmatory in nature, and based on existing theories of relationships among biological, psychological and social variables (Fried & Cramer, 2017).

The study also has limitations. It should be noted that while we consider 7-day accelerometry (with inclusion criterion of accumulating more than 4 days of over 10 hours wear time) an approximation of a participant's true habitual PA and SB in their daily life, it is not an errorless measure and it does not capture all forms of activity. Additionally, the questionnaire to measure the BCTs requires future validation (Bringmann & Eronen, 2016; Flake & Fried, 2019; Hankonen, 2018).

In the network model used, regularisation techniques are applied to remove spurious relations and control for multiple testing (for an in-depth tutorial on such regularised network models, see Epskamp and Fried (2018), and for a health psychology specific use case, see Hevey (2018)). At the same time, these networks estimate relations that are akin to partial correlations to derive the conditional dependence structure among variables. Potential pitfalls of these models and their application have been discussed elsewhere in detail (Fried & Cramer, 2017; Guloksuz, Pries, & Van Os, 2017). Most importantly, while in social networks one can include all relevant nodes (e.g., all people in a classroom or company), this is not so in biopsychosocial networks, where the question of what items to include as nodes remains a challenging question. Relations among items are often interpreted as putative causal pathways (although many other interpretations exist, Epskamp and Fried (2018)), which means one should not include two variables that are simply two indicators of the same construct (e.g. the items 'I often feel sad' and 'I often feel blue'). Another important challenge is that one should avoid statistically controlling for common effects, also known as colliders: If in the true model both A and B independently cause C, C is a collider.

If one controls for C in the model, a negative relation between A and B will emerge where no relation exists in the true model. This applies to all regression models and network models that are based on regressions, and it can be challenging to determine if a given variable is a collider. Rohrer Rohrer (2018) provides an approachable introduction to causal inference in observational data.

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 these 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 dynamic processes. Few individual participants are described by the group-level summary statistics: In fact, using Daniels' (Daniels, 1952) definition of an 'approximately average individual' as falling in the middle 30% of the range of values, only 1.50% of participants can be considered 'average' on all of the primary outcome measures (see supplementary website, section https://git.io/fpOy1). 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. Still, and especially when processes are considered, group-level data does not inform the individual-level mechanisms of action in the case of non-ergodic systems, and hence the agreement between features of these two levels should be investigated (Fisher, Medaglia, & Jeronimus, 2018).

In conclusion, this analysis of baseline data from the Let's Move It intervention trial indicates that randomisation did not result in highly disproportionate groups, i.e. the differences between arms were small – although, in the case of complex systems, even minimal differences may proliferate and lead to group imbalances (Rickles, 2009). It also highlights that vocational school students 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. Transparent sharing of analyses and analytical choices is imperative for increasing confidence in research findings.

List of abbreviations.

PA = Physical activity

MVPA = Moderate-to-vigorous physical activity

SB = Sedentary behaviour

BCT = Behaviour change technique

Nur = Practical nurse

HRC = Hotel, restaurant and catering studies

BA = Business and administration

IT = Business information technology

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 https://osf.io/jn9ax/ after the anonymisation process has been completed in August 2019. All analyses and code are available at https://git.io/fNHuf (permalink at Heino and Sund (2019), GitHub repository at https://git.io/fjIQ6). The electronic questionnaire form is available at https://git.io/fjIP5.

Competing interests. The authors declare that they have no competing interests.

Authors' contributions. MH wrote the analysis code, including the full online supplement, formulated the initial draft of the manuscript and revised it in collaboration

with all co-authors. TV was responsible for planning and analysing the PA and SB measured from data collected with accelerometer. RS and EIF provided expertise regarding the statistical analyses. KB, AH, AU, VA-S, TV, RS and NH contributed to planning of the trial design and data collection including the measures used. NH, with the study co-applicants, conceived of the study. NH acted as principal investigator of the research project. All authors read and approved the final manuscript.

Funding. MH was supported by Academy of Finland (grant number 295765) and Ministry for Education and Culture, Sports Science projects (grant number OKM/81/626/2014). 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. We are also grateful to Frederik Aust for technical support in creating a reproducible manuscript, as well as Ruben Arslan in creating the codebook.

References

Arslan, R. C. (forthcoming). How to automatically document data with the codebook package to facilitate data re-use. *Advances in Methods and Practices in Psychological Science*. https://doi.org/10.31234/osf.io/5qc6h

Aust, F. (2019). *Papaja (Preparing APA Journal Articles) is an R package that provides document formats to produce complete APA manscripts from RMarkdown-files (PDF and Word documents) and helper functions that fac.*. Retrieved from https://github.com/crsh/papaja (Original work published 2014)

Briganti, G., Kempenaers, C., Braun, S., Fried, E. I., & Linkowski, P. (2018). Network analysis of empathy items from the interpersonal reactivity index in 1973 young adults. *Psychiatry Research*, *265*, 87–92. https://doi.org/10.1016/j.psychres.2018.03.082

Bringmann, L. F., & Eronen, M. I. (2016). Heating up the measurement debate: What psychologists can learn from the history of physics. *Theory & Psychology*, *26*(1), 27–43. https://doi.org/10.1177/0959354315617253

Chalmers, I., & Glasziou, P. (2009). Avoidable waste in the production and reporting of research evidence. *Lancet*, *374*(9683), 86–89. https://doi.org/10.1016/S0140-6736(09)60329-9

Cleveland, W. S. (1993). *Visualizing data*. Hobart Press.

Costantini, G., Epskamp, S., Borsboom, D., Perugini, M., Mõttus, R., Waldorp, L. J., & Cramer, A. O. (2015). State of the aRt personality research: A tutorial on network analysis of personality data in R. *Journal of Research in Personality*, *54*, 13–29. https://doi.org/10.1016/j.jrp.2014.07.003

Costantini, G., Richetin, J., Preti, E., Casini, E., Epskamp, S., & Perugini, M. (2017). Stability and variability of personality networks. A tutorial on recent developments in network psychometrics. *Personality and Individual Differences*. https://doi.org/10.1016/j.paid.2017.06.011

Dalege, J., Borsboom, D., Harreveld, F., Waldorp, L. J., & Maas, H. L. (2017a). Network structure explains the impact of attitudes on voting decisions. *Scientific Reports*, *7*(1), 4909.

Dalege, J., Borsboom, D., van Harreveld, F., van den Berg, H., Conner, M., & van der Maas, H. L. J. (2016). Toward a formalized account of attitudes: The Causal Attitude Network (CAN) model. *Psychological Review*, *123*(1), 2–22. https://doi.org/10.1037/a0039802

Dalege, J., Borsboom, D., van Harreveld, F., & van der Maas, H. L. (2017b). Network analysis on attitudes: A brief tutorial. *Social Psychological and Personality Science*, *8*(5), 528–537.

Daniels, G. S. (1952). *The" Average Man"?* AIR FORCE AEROSPACE MEDICAL RESEARCH LAB WRIGHT-PATTERSON AFB OH.

Dieleman, J. L., Sadat, N., Chang, A. Y., Fullman, N., Abbafati, C., Acharya, P., … Alizadeh-Navaei, R. (2018). Trends in future health financing and coverage: Future health spending and universal health coverage in 188 countries, 2016–40. *The Lancet*, *391*(10132), 1783–1798.

Elgar, F. J., Pförtner, T.-K., Moor, I., De Clercq, B., Stevens, G. W. J. M., & Currie, C. (2015). 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*, *385*(9982), 2088–2095.

https://doi.org/10.1016/S0140-6736(14)61460-4

Epskamp, S., Borsboom, D., & Fried, E. I. (2018). Estimating psychological networks and their accuracy: A tutorial paper. *Behavior Research Methods*, *50*(1), 195–212. https://doi.org/10.3758/s13428-017-0862-1

Epskamp, S., Costantini, G., Haslbeck, J., Cramer, A. O. J., Waldorp, L. J., Schmittmann, V. D., & Borsboom, D. (2019). Qgraph: Graph Plotting Methods, Psychometric Data Visualization and Graphical Model Estimation (Version 1.6.1). Retrieved from https://CRAN.R-project.org/package=qgraph

Epskamp, S., & Fried, E. I. (2018). A Tutorial on Regularized Partial Correlation Networks. *Psychological Methods*. https://doi.org/10.1037/met0000167

Expert Advisory Group on Data Access. (2015). *EAGDA Report: Governance of Data Access*. Retrieved from https://web.archive.org/web/20180914101149/https://wellcome.ac.uk/sites/default/files/governance-of-data-access-eagda-jun15.pdf

Fagt, S., Andersen, L. F., Anderssen, S. A., Becker, W., Borodulin, K., Fogelholm, M., … Trolle, E. (2012). *Nordic Monitoring of diet, physical activity and overweight : Validation of indicators*. Retrieved from http://urn.kb.se/resolve?urn=urn:nbn:se:norden:org:diva-1639

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences*, 201711978. https://doi.org/10.1073/pnas.1711978115

Flake, J. K., & Fried, E. I. (2019). *Measurement Schmeasurement: Questionable Measurement Practices and How to Avoid Them*. https://doi.org/10.31234/osf.io/hs7wm

Fried, E. I., Borkulo, C. D. van, Cramer, A. O. J., Boschloo, L., Schoevers, R. A., & Borsboom, D. (2017). Mental disorders as networks of problems: A review of recent insights. *Social Psychiatry and Psychiatric Epidemiology*, *52*(1), 1–10. https://doi.org/10.1007/s00127-016-1319-z

Fried, E. I., & Cramer, A. O. (2017). Moving forward: Challenges and directions for psychopathological network theory and methodology. *Perspectives on Psychological Science*, *12*(6), 999–1020.

Gardner, M. J., & Altman, D. G. (1986). Confidence intervals rather than P values: Estimation rather than hypothesis testing. *BMJ*, *292*(6522), 746–750. https://doi.org/10.1136/bmj.292.6522.746

Gigerenzer, G. (2018). Statistical Rituals: The Replication Delusion and How We Got There. *Advances in Methods and Practices in Psychological Science*, 2515245918771329. https://doi.org/10.1177/2515245918771329

Guloksuz, S., Pries, L. K., & Van Os, J. (2017). Application of network methods for understanding mental disorders: Pitfalls and promise. *Psychological Medicine*, *47*(16), 2743–2752.

Hallgren, K. A., McCabe, C. J., King, K. M., & Atkins, D. C. (2018). Beyond path diagrams: Enhancing applied structural equation modeling research through data visualization. *Addictive Behaviors*. https://doi.org/10.1016/j.addbeh.2018.08.030

Hankonen, N. (2018). *Participants' enactment of behavior change techniques: A call for increased focus on what people do to manage their motivation and behaviour*. https://doi.org/10.31234/osf.io/pa4wg

Hankonen, N., Absetz, P., & Araujo-Soares, V. (2019). *Changing activity behaviours in vocational school students: The stepwise development and optimised content of the*

'Let's Move it' intervention. https://doi.org/10.31234/osf.io/ak68f

Hankonen, N., Heino, M. T. J., Araujo-Soares, V., Sniehotta, F. F., Sund, R., Vasankari, T., … Haukkala, A. (2016). '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*, *16*, 451–466. https://doi.org/10.1186/s12889-016-3094-x

Hankonen, N., Heino, M. T. J., Hynynen, S.-T., Laine, H., Araújo-Soares, V., Sniehotta, F. F., … Haukkala, A. (2017). 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*, *14*(1). https://doi.org/10.1186/s12966-017-0484-0

Hankonen, N., Heino, M. T. J., Kujala, E., Hynynen, S.-T., Absetz, P., Araújo-Soares, V., … Haukkala, A. (2017). What explains the socioeconomic status gap in activity? Educational differences in determinants of physical activity and screentime. *BMC Public Health*, *17*(1), 144. https://doi.org/10.1186/s12889-016-3880-5

Haslbeck, J. (2019). Mgm: Estimating Time-Varying k-Order Mixed Graphical Models (Version 1.2-6). Retrieved from https://CRAN.R-project.org/package=mgm

Haslbeck, J., & Waldorp, L. J. (2015). Structure estimation for mixed graphical models in high-dimensional data. *arXiv Preprint arXiv:1510.05677.*

Heino, M. T. J., Knittle, K., Haukkala, A., Vasankari, T., & Hankonen, N. (2018a). Simple and rationale-providing SMS reminders to promote accelerometer use: A within-trial randomised trial comparing persuasive messages. *BMC Public Health*, *18*(1), 1352. https://doi.org/10.1186/s12889-018-6121-2

Heino, M. T. J., & Sund, R. (2019). *Source code: Visualisation and network analysis of physical activity and its determinants*. https://doi.org/10.5281/zenodo.2628764

Heino, M. T. J., Vuorre, M., & Hankonen, N. (2018b). Bayesian evaluation of behavior change interventions: A brief introduction and a practical example. *Health Psychology and Behavioral Medicine*, *6*(1), 49–78. https://doi.org/10.1080/21642850.2018.1428102

Hevey, D. (2018). Network analysis: A brief overview and tutorial. *Health Psychology and Behavioral Medicine*, *6*(1), 301–328. https://doi.org/10.1080/21642850.2018.1521283

Husu, P., Suni, J., Vähä-Ypyä, H., Sievänen, H., Tokola, K., Valkeinen, H., … Vasankari, T. (2016). Objectively measured sedentary behavior and physical activity in a sample of Finnish adults: A cross-sectional study. *BMC Public Health*, *16*(1). https://doi.org/10.1186/s12889-016-3591-y

Husu, P., Vähä-Ypyä, H., & Vasankari, T. (2016). 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*, *16*(1), 338.

Hynynen, S.-T., van Stralen, M. M., Sniehotta, F. F., Araújo-Soares, V., Hardeman, W., Chinapaw, M. J. M., … Hankonen, N. (2016). A systematic review of school-based interventions targeting physical activity and sedentary behaviour among older adolescents. *International Review of Sport and Exercise Psychology*, *9*(1), 22–44. https://doi.org/10.1080/1750984X.2015.1081706

Kepes, S., & McDaniel, M. A. (2013). How Trustworthy Is the Scientific Literature in Industrial and Organizational Psychology? *Industrial and Organizational Psychology*, *6*(3), 252–268. https://doi.org/10.1111/iops.12045

Kievit, R. A., Frankenhuis, W. E., Waldorp, L. J., & Borsboom, D. (2013). Simpson's paradox in psychological science: A practical guide. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00513

Köykkä, K., Absetz, P., Araújo-Soares, V., Knittle, K., Sniehotta, F. F., & Hankonen, N. (2018). Combining the reasoned action approach and habit formation to reduce sitting time in classrooms: Outcome and process evaluation of the Let's Move It teacher intervention. *Journal of Experimental Social Psychology*. https://doi.org/10.1016/j.jesp.2018.08.004

Mäkelä, K., Kokko, S., Kannas, L., Villberg, J., Vasankari, T., Heinonen, J. O., … Selänne, H. (2016). 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*, *6*.

Michie, S., Abraham, C., Whittington, C., McAteer, J., & Gupta, S. (2009). Effective techniques in healthy eating and physical activity interventions: A meta-regression. *Health Psychology*, *28*(6), 690–701. https://doi.org/10.1037/a0016136

Molenaar, P. C. M. (2010). Latent variable models are network models. *Behavioral and Brain Sciences*, *33*(2-3), 166–166. https://doi.org/10.1017/S0140525X10000798

Moore, G. F., Audrey, S., Barker, M., Bond, L., Bonell, C., Hardeman, W., … Baird, J. (2015). Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*, *350*, h1258. https://doi.org/10.1136/bmj.h1258

Mõttus, R., & Allerhand, M. (2017). Why do traits come together? The underlying trait and network approaches. *SAGE Handbook of Personality and Individual Differences*, *1*, 1–22. Retrieved from https://www.researchgate.net/profile/Rene_Mttus/publication/312341252_Why_do_traits_come_together_The_underlying_trait_and_network_approaches/links/587de06808ae4445c06e16fb.pdf

Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, 201708274. https://doi.org/10.1073/pnas.1708274114

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific Utopia II. Restructuring Incentives and Practices to Promote Truth Over Publishability. *Perspectives on Psychological Science*, *7*(6), 615–631. https://doi.org/10.1177/1745691612459058

Peters, G.-J. (2018). *Diamond Plots: A tutorial to introduce a visualisation tool that facilitates interpretation and comparison of multiple sample estimates while respecting their inaccuracy*. https://doi.org/10.31234/osf.io/fzh6c

Peters, G.-J. (2019). Ufs: Quantitative Analysis Made Accessible (Version 0.2.0). Retrieved from https://CRAN.R-project.org/package=ufs

Peters, G.-J., Verboon, P., & Green, J. (2018). Userfriendlyscience: Quantitative Analysis Made Accessible (Version 0.7.2). Retrieved from https://CRAN.R-project.org/package=userfriendlyscience

R Core Team. (2015). R: A Language and Environment for Statistical Computing (Version 3.1.1). Vienna, Austria: R Foundation for Statistical Computing.

R Core Team. (2018). *R: A language and environment for statistical computing*. Retrieved from https://www.R-project.org/

Rickles, D. (2009). Causality in complex interventions. *Medicine, Health Care, and Philosophy*, *12*(1), 77–90. https://doi.org/10.1007/s11019-008-9140-4

Rogers, P. J. (2008). Using Programme Theory to Evaluate Complicated and Complex Aspects of Interventions. *Evaluation*, *14*(1), 29–48. https://doi.org/10.1177/1356389007084674

Rohrer, J. M. (2018). Thinking Clearly About Correlations and Causation: Graphical Causal Models for Observational Data. *Advances in Methods and Practices in Psychological Science*, *1*(1), 27–42. https://doi.org/10.1177/2515245917745629

Rousselet, G. A., Pernet, C. R., & Wilcox, R. R. (2017). Beyond differences in means: Robust graphical methods to compare two groups in neuroscience. *European Journal of Neuroscience*, *46*(2), 1738–1748. https://doi.org/10.1111/ejn.13610

RStudio Team. (2015). RStudio: Integrated Development Environment for R (Version 0.99.491). Retrieved from http://www.rstudio.com/

Saxon, E. (2015). Beyond bar charts. *BMC Biology*, *13*(1), 60. https://doi.org/10.1186/s12915-015-0169-6

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science*, *22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Smaldino, P. E., & McElreath, R. (2016). The natural selection of bad science. *Open Science*, *3*(9), 160384. https://doi.org/10.1098/rsos.160384

Sterne, J. A. C. (2001). Sifting the evidence—what's wrong with significance tests? Another comment on the role of statistical methods. *BMJ*, *322*(7280), 226–231. https://doi.org/10.1136/bmj.322.7280.226

Tay, L., Parrigon, S., Huang, Q., & LeBreton, J. M. (2016). Graphical Descriptives A Way to Improve Data Transparency and Methodological Rigor in Psychology. *Perspectives on Psychological Science*, *11*(5), 692–701. Retrieved from http://pps.sagepub.com/content/11/5/692.abstract

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso: A retrospective. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *73*(3), 273–282. https://doi.org/10.1111/j.1467-9868.2011.00771.x

Trafimow, D., Wang, T., & Wang, C. (2018). Means and standard deviations, or locations and scales? That is the question! *New Ideas in Psychology*, *50*, 34–37. https://doi.org/10.1016/j.newideapsych.2018.03.001

Tukey, J. W. (1977). *Exploratory data analysis* (Vol. 2). Reading, Mass.

Van Der Maas, H. L. J., Kan, K.-J., Marsman, M., & Stevenson, C. E. (2017). Network Models for Cognitive Development and Intelligence. *Journal of Intelligence*, *5*(2), 16. https://doi.org/10.3390/jintelligence5020016

Vanpaemel, W., Vermorgen, M., Deriemaecker, L., & Storms, G. (2015). Are We Wasting a Good Crisis? The Availability of Psychological Research Data after the Storm. *Collabra: Psychology*, *1*(1). https://doi.org/10.1525/collabra.13

van Sluijs, E. M., Skidmore, P. M., Mwanza, K., Jones, A. P., Callaghan, A. M., Ekelund, U., . . . Wareham, N. J. (2008). 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*, *8*(1), 388.

Wasserstein, R. L., & Lazar, N. A. (2016). The ASA's Statement on p-Values: Context, Process, and Purpose. *The American Statistician*, *70*(2), 129–133. https://doi.org/10.1080/00031305.2016.1154108

Weissgerber, T. L., Garovic, V. D., Savic, M., Winham, S. J., & Milic, N. M. (2016). From Static to Interactive: Transforming Data Visualization to Improve Transparency. *PLOS Biology*, *14*(6), e1002484. https://doi.org/10.1371/journal.pbio.1002484

Weissgerber, T. L., Milic, N. M., Winham, S. J., & Garovic, V. D. (2015). Beyond Bar and Line Graphs: Time for a New Data Presentation Paradigm. *PLOS Biology*, *13*(4), e1002128. https://doi.org/10.1371/journal.pbio.1002128

Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., … RStudio. (2018). Ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics (Version 3.1.0). Retrieved from https://CRAN.R-project.org/package=ggplot2

Wilke, C. O., & ggridges), R. (. for ggplot2 code copied to. (2018). Ggridges: Ridgeline Plots in 'ggplot2' (Version 0.5.1). Retrieved from https://CRAN.R-project.org/package=ggridges

Table 1

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

Variable	Nur	HRC	BA	IT	Full samp
n	402	213	282	163	1166
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
Born in Finland (%)	80.1	88.3	89.7	86.7	84.4
% girl	82.3	60.6	39.0	16.0	56.5
% allocated to intervention	68.9	31.5	53.5	46.6	54.7

Table 2

Key variables with their class and school intra-class correlations (ICCs). Let's Move It trial's primary outcome variables marked with asterisks. Accelerometry data is missing from 435 participants, of whom 169 due to not meeting the cutoff of at least 10 hours of measurement time for at least four days. Survey data missing from 84 participants.

Variable	Mean	CI95	ICC class	ICC school	n
Daily moderate-to-vigorous PA time (accelerometer)*	1h 5min	0h 57min - 1h 13min	.089	.062	731
Daily light PA time (accelerometer)	2h 51min	2h 32min - 3h 9min	.111	.110	731
Daily standing time (accelerometer)	1h 24min	1h 15min - 1h 34min	.122	.041	731
Daily time spent sitting or lying down (accelerometer)*	8h 44min	8h 4min - 9h 24min	.115	.138	731
Daily number of times sitting was interrupted (accelerometer)*	25.8	23.5 - 28.0	.047	.080	731
Number of days with >30 MVPA min previous week (self-report)*	2.8	2.6 - 3.0	.047	< .001	1082

Table 3

Statistics describing accelerometer-measured moderate-to-vigorous physical activity in different educational tracks. Values not corrected for eff

Gender	Arm	Nur	HRC	
girl	control	M=73.0; SD=29.5; skewness=0.9; kurtosis= 0.6; n=104	M=57.5; SD=22.3; skewness=0.8; kurtosis= 1.0; n=90	M
girl	intervention	M=71.8; SD=28.4; skewness=1.0; kurtosis= 1.6; n=227	M=52.0; SD=23.6; skewness=0.8; kurtosis=-0.2; n=39	M
boy	control	M=72.7; SD=28.9; skewness=0.3; kurtosis=-1.1; n=21	M=56.1; SD=27.4; skewness=1.5; kurtosis= 2.0; n=56	M
boy	intervention	M=89.6; SD=39.1; skewness=0.7; kurtosis= 0.2; n=50	M=71.2; SD=43.9; skewness=0.9; kurtosis= 0.5; n=28	M

Table 4

Main theoretical determinants of physical activity (PA) and sedentary behaviour (SB). Mean (CI95, taking into account school and class mem
and coping planning are evaluated on a scale from 1 to 4, autonomous / controlled regulation, amotivation and behaviour change technique (BC
from 1 to 6 – all other variables from 1 to 7.

Variable	Girls (n = 603-611)	Boys (n = 459-467)	Intervention (n = 570-579)	Control (n = 492-499)	Total (
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
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
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
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
PA descriptive norm	4.3 (4.1 - 4.5)	4.6 (4.4 - 4.7)	4.3 (4.1 - 4.5)	4.5 (4.3 - 4.7)	4.4 (4.2
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
PA outcome expectations	5.4 (5.2 - 5.5)	5.1 (5.0 - 5.3)	5.2 (5.0 - 5.5)	5.3 (5.1 - 5.5)	5.3 (5.1
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
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
PA autonomous regulation	3.3 (3.2 - 3.5)	3.6 (3.4 - 3.7)	3.4 (3.2 - 3.5)	3.5 (3.3 - 3.6)	3.4 (3.3
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
PA amotivation	1.5 (1.4 - 1.5)	1.6 (1.5 - 1.7)	1.5 (1.4 - 1.6)	1.5 (1.4 - 1.6)	1.5 (1.5
PA agreement-BCTs	3.1 (2.9 - 3.2)	3.2 (3.0 - 3.3)	3.0 (2.9 - 3.2)	3.2 (3.0 - 3.4)	3.1 (3.0
PA frequency-BCTs	2.5 (2.4 - 2.6)	2.6 (2.4 - 2.7)	2.5 (2.4 - 2.6)	2.6 (2.4 - 2.7)	2.5 (2.4

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
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
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
SB outcome expectations	4.9 (4.8 - 5.0)	4.5 (4.4 - 4.7)	4.8 (4.5 - 5.0)	4.8 (4.6 - 5.0)	4.8 (4.6

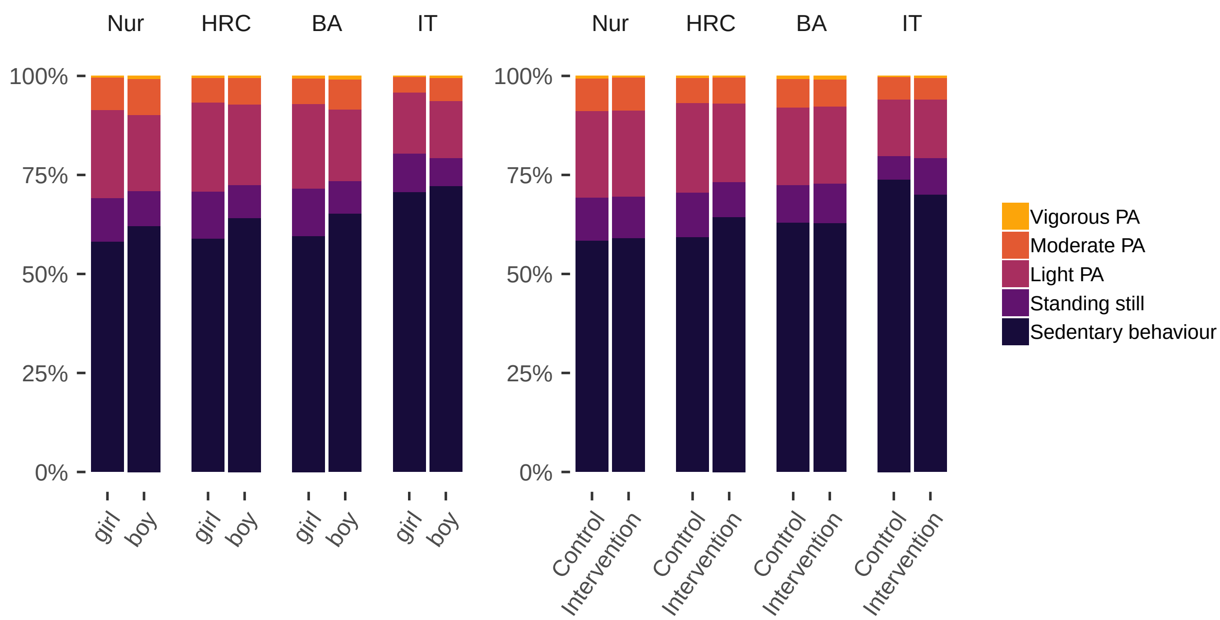

Figure 1. Stacked bar plot drawn with R package `ggplot` (Wickham et al. (2018), code available at https://git.io/fptlp), showing proportions of accelerometer-measured physical activity (PA) in relation to measurement time, averaged over genders, arms and educational tracks. Nur = Practical nurse, HRC = Hotel, restaurant and catering, BA = Business and administration, IT = Information and communications technology.

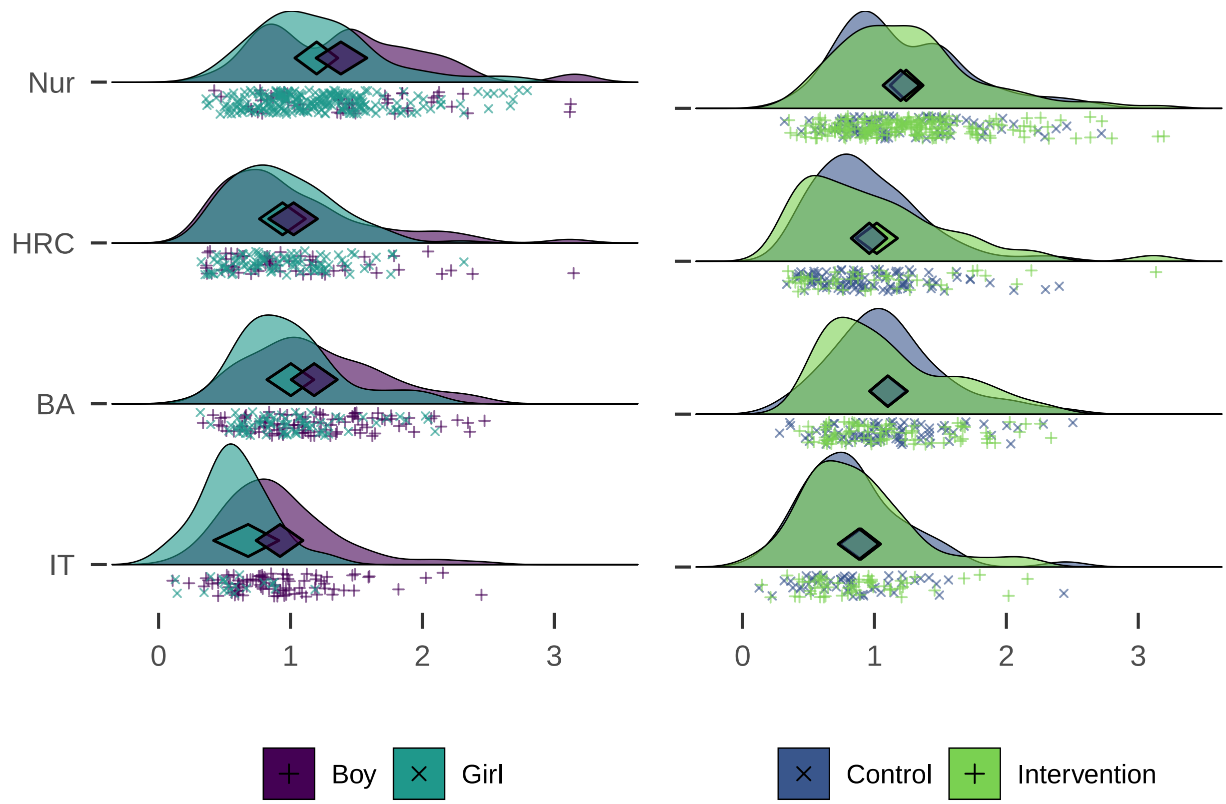

Figure 2. Raincloud ridge plot combined with a diamond plot, drawn with R packages `ggridges` (Wilke & ggridges), 2018) and `userfriendlyscience` (Peters et al. (2018), code available at https://git.io/fjLBG), showing hours of accelerometer-measured moderate-to-vigorous physical activity for different educational tracks. Midpoints of diamonds indicate means, endpoints 95% credible intervals (see (Heino et al., 2018b) for interpretation). 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.

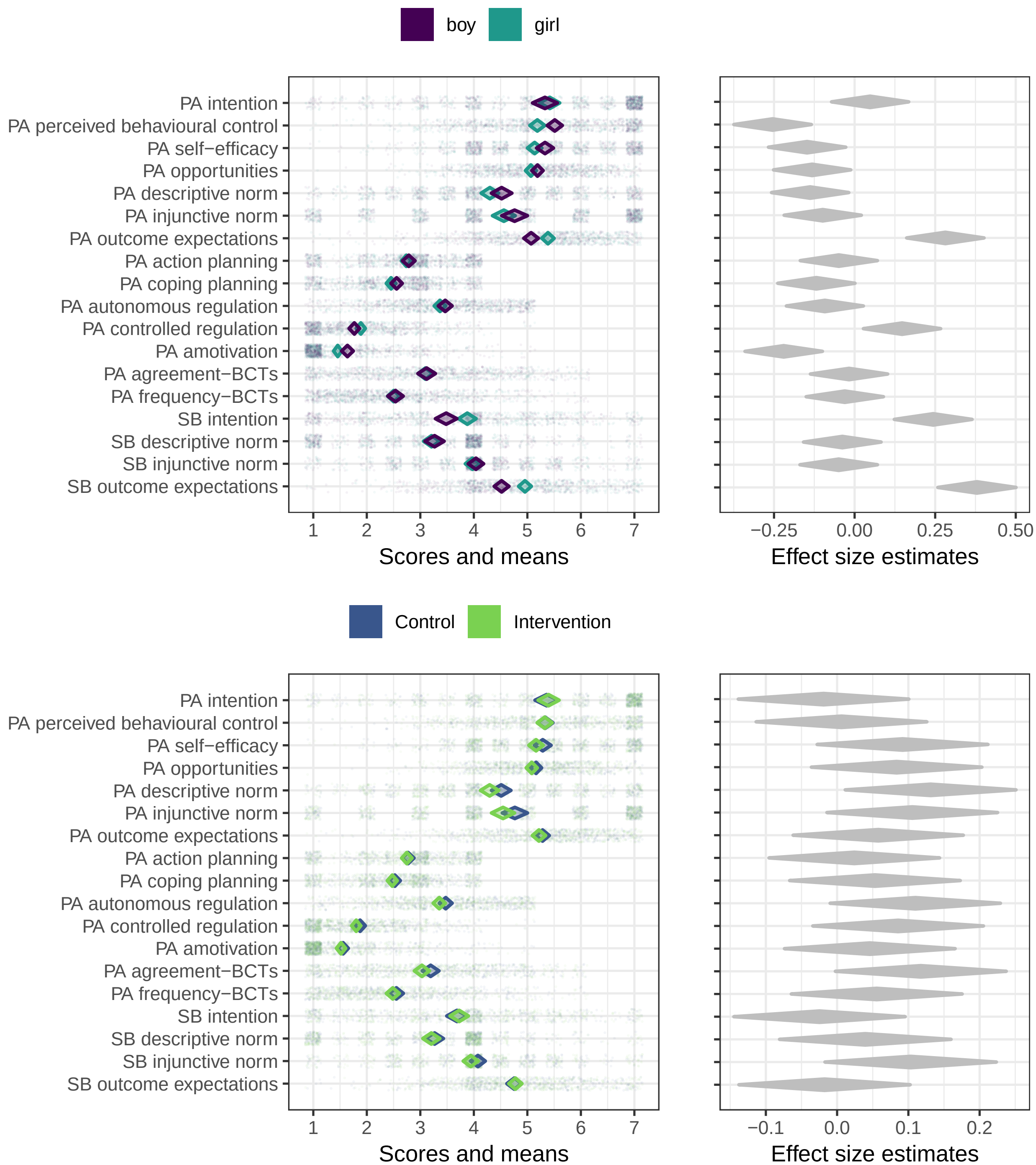

Figure 3. Diamond comparison plot drawn with R package `ufs` (Peters (2019), code available at https://git.io/fjLBB), showing means (middle of diamonds), 99% confidence intervals (endpoints of diamonds) and individual answers (dots) separated by gender and arm. Rightmost plots show heuristic effect sizes for differences in means (transformed to Pearson's r). ICC is not accounted for in any plot.

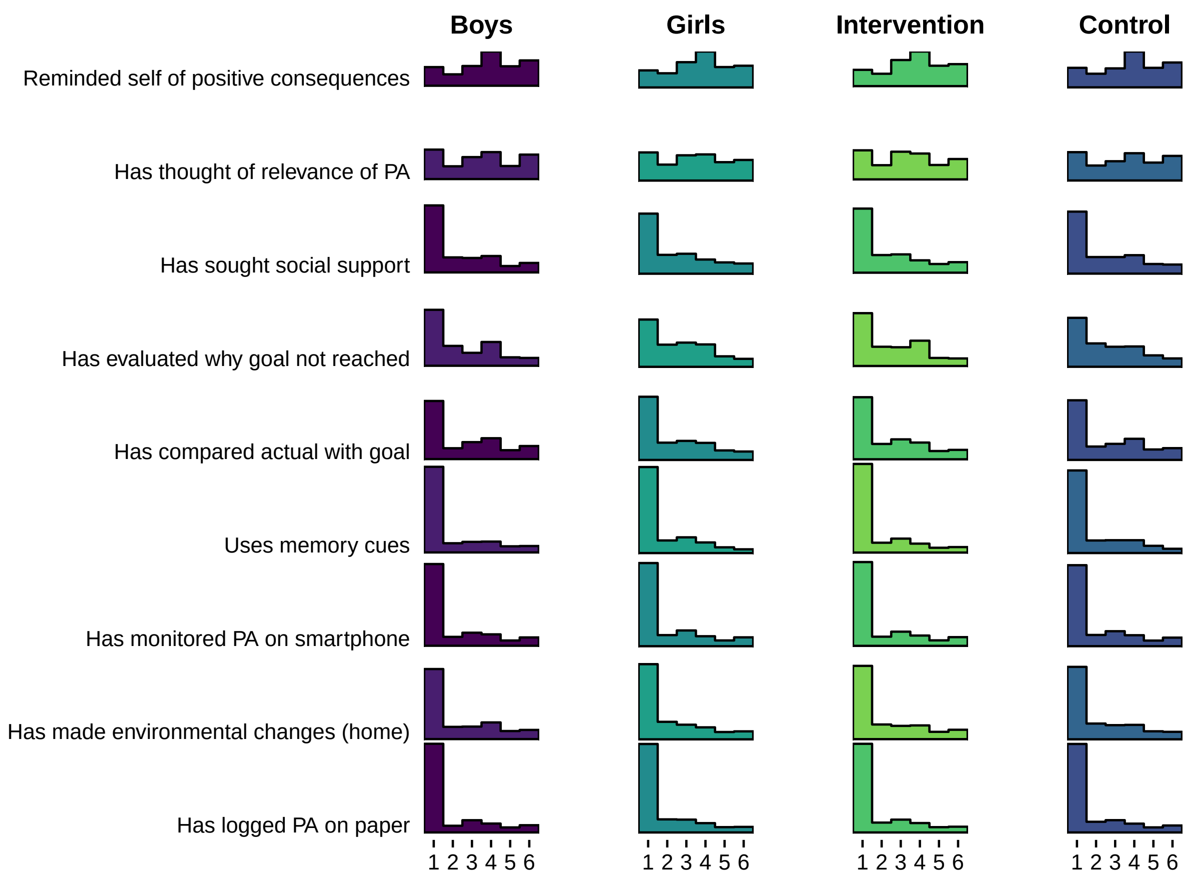

Figure 4. Histogram drawn with R package `ggridges` (Wilke and ggridges) (2018), code available at https://git.io/fpOLj), showing self-reported use of frequency-dependent BCTs (1 = Not once … 6 = Daily).

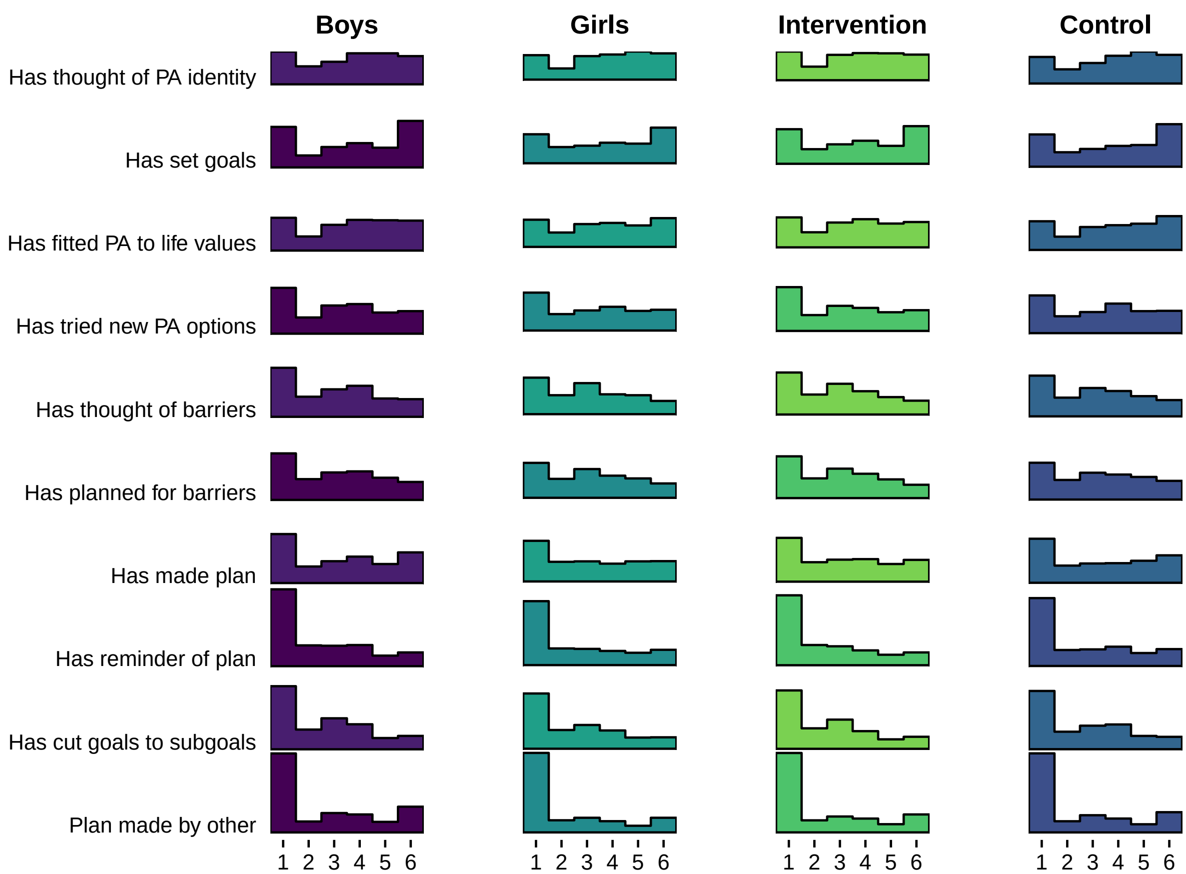

Figure 5. Histogram drawn with R package `ggridges` (Wilke and ggridges) (2018), code available at https://git.io/fjLBE), showing self-reported use of agreement-dependent BCTs (1 = Not at all true . . . 6 = Completely true).

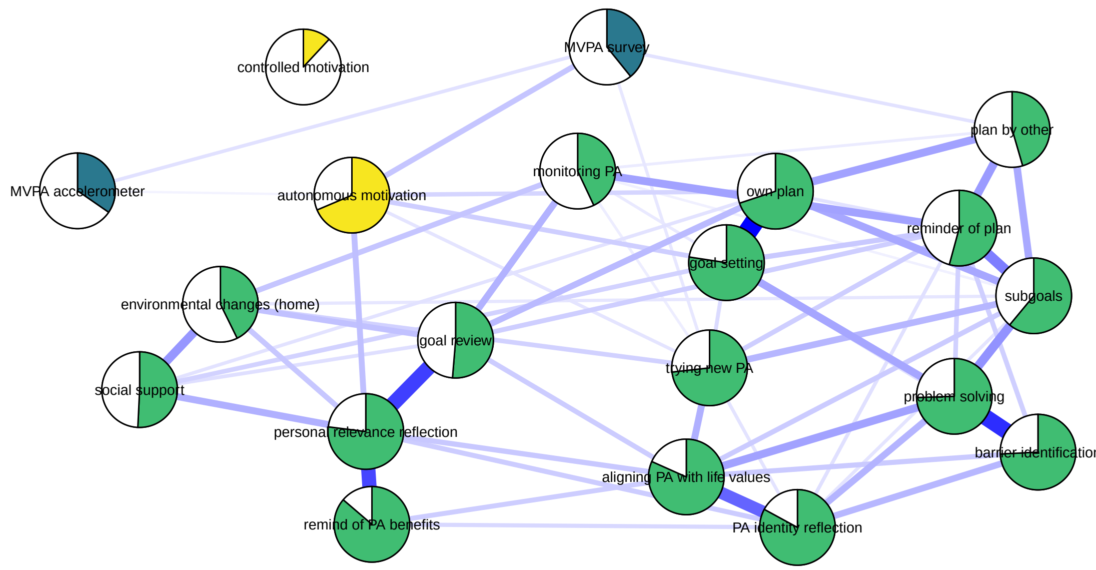

Figure 6. Mixed graphical model with LASSO regularisation and model selection by EBIC. Network models estimated and drawn with packages `mgm` (Haslbeck, 2019) and `qgraph` (Epskamp et al. (2019), code available at https://git.io/fpOXV). Blue lines indicate positive relationships. Plot shows the conditional dependence relationships between the variables of interest (edges which connect nodes), which can be interpreted akin to partial correlations. Pies depict means as proportion of theoretical maximum (in the case of accelerometer-measured moderate-to-vigorous physical activity (MVPA), mean as proportion of highest observed value); behaviour change technique (BCT) use and controlled motivation are dichotomised (see Methods). Colours distinguish the three types of nodes; MPVA, motivation, and BCT use.