This is the accepted manuscript of the article, which has been published in EASEAI 2019: Proceedings of the 1st ACM SIGSOFT International Workshop on Education through Advanced Software Engineering and Artificial Intelligence. 2019, 21-26. https://doi.org/10.1145/3340435.3342720

Application of Data Clustering for Automated Feedback Generation about Student Well-Being

Mikko Kylväjä
School Day
Helsinki, Finland
mikko.kylvaja@schoolday.fi

Pekka Kumpulainen
Datatohtori
Tampere, Finland
pekka.kumpulainen@datatohtori.fi

Anne Konu
Faculty of Social Sciences
Health Sciences
Tampere University, Finland
anne.konu@tuni.fi

ABSTRACT

Investment in the well-being of today's schoolchildren is an important investment in the future. We believe that learning does not happen in the absence of well-being. This data-oriented research studies how automation utilizing data analysis algorithms could help provide the students with feedback and guidance about their well-being related issues. We implemented a system that combines data processing methods and research-based knowledge to serve that purpose.

Our target was to develop an automated feedback system utilizing information from a large data set collected from well-being surveys from students, as well as research-based well-being knowledge. The system can be used to provide automated feedback for students who answer a well-being survey.

CCS CONCEPTS

• Applied computing~Education • Applied computing~Life and medical sciences • Computing methodologies~Artificial intelligence • Computing methodologies~Modeling and simulation

KEYWORDS

Student, Well-being, Feedback, Automation, Clustering, Hierarchical clustering, Application, Design Science.

1 Introduction

In this paper we present a novel process to generate well-being related feedback automatically and anonymously for students. Nothing can replace an adult presence in students' everyday life, but automated solutions can help to predict negative phenomena early, improve awareness of well-being in everyday life and help teachers to use their time efficiently. If well-being related problems are not solved early, the effect may be life-long for the student, with a corresponding cost for society.

Bullying in schools is an important and commonly researched topic (e.g. Olweus 1993, Salmivalli 2009, Rigby & Smith 2011). However, more research is needed about school well-being as a complex phenomenon. However, school well-being is a far more complex phenomenon. There exist some whole school approaches to school well-being, e.g. WHO Health promoting schools (Langford 2014) and Comprehensive School Health (http://www.jcsh-cces.ca/index.php/about/comprehensive-school-health/what-is-csh), MindMatters (Wyn et al. 2000) and School Well-being Model (Konu & Rimpelä, 2002). This paper takes the whole school approach and aims to find an intelligent and timely solution for helping to promote students' well-being in schools. The current framework was selected because it is theoretically grounded and focuses both on the well-being concept and the definition of health. In addition, as a notable difference to other approaches is the subcategory "means for self-fulfillment". (Konu & Rimpelä 2002.) The theoretical framework is thoroughly tested and published (e.g. Konu & Koivisto 2011; Konu et al. 2015) and broadly cited by other researchers.

The data set used in the research was collected from a total of 64,139 lower secondary school students, between the ages of 13 to 15 years, from Finland. The theoretical framework and data used in this research as well as related work is presented in section 2.

The automated feedback solution is explained in section 3. We explain the overall process, data preprocessing, clustering method and finally how the results can be used in an automated feedback system. In section 4 we present the result clusters and examples of feedback given to students. Finally, section 5 presents conclusions of the work and further research ideas.

In earlier research clustering has been utilized in providing automated personalized feedback separately for students (Kaleeswaran 2016), and for health (Rabbi 2015).

2 Theoretical framework

We use the School Well-being Model (Konu & Rimpelä, 2002) as the theory behind the current research on the use of data-based automation to help to promote students' well-being. In the model, learning, education and well-being are combined (Figure 1). Well-being consists of four categories: school conditions (e.g. physical conditions, enjoyable atmosphere, teaching organization, rules and punishments, services like school nurse and student counselor); social relationships (e.g. student-student, teacher-student and parents-school relationships); means for self-fulfillment (e.g. student appreciation, encouragement, guidance and help to students, student participation); health status (self-experienced health, psychosomatic symptoms like pains, mood, sleep, fear and common flus).

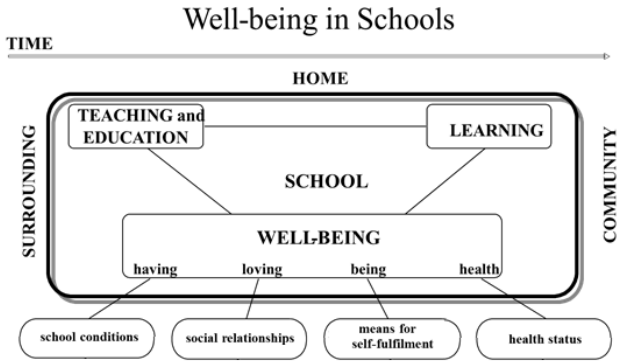

Figure 1. The School Well-being Model (modified) (Konu and Rimpelä, 2002)

A method to measure the well-being of primary, lower and upper secondary students and for school personnel was developed based on the model (Konu & Lintonen 2006a; Konu & Lintonen 2006b). The method is called School Well-being Profile and it is an internet-based service, which produces comparable data for schools on-line. The validity and reliability of the measurement method has been proven in earlier research. Reliability was good in each dataset (Konu & Koivisto 2011; Konu et al. 2015). The use of the Profile is free, and it announces in its front page that the data is anonymously used in research. The answering is voluntary, and it happens during the school hours. The data for this research was derived from this service.

The data was derived yearly from the internet-based School Well-being Profile during school years from 2008-2009 to 2017-2018 (N=90,310). Each year the data was secured anonymously in the Finnish Social Science Data Archive (https://www.fsd.uta.fi/en/; e.g. Konu, 2017-2018). The data consist of the answers of upper secondary school students (aged 13-15 years) from the schools that have used the School Well-being Profile service. Names of the students and schools were not saved in database (and they are not even asked). No sensitive questions were included in the questionnaire. The 81 statements deal with experiences about school, its activities and experienced health with common symptoms like neck-pain, tiredness, mood and flu. The answering options are five-point Likert scale, which are shown in Table 1.

Most of the questions (67) use the *General* answering options (Table 1). Eleven questions use the *Symptom* answering options. Two questions use the *Bullying* answer options. One question uses the *Experienced health* answer options.

Table 1. Answering options in different question categories

General questions	Bullying questions	Experienced health	Symptom questions
totally agree	not at all	very good	not at all
agree	few times during the term	quite good	rarely
nor agree or disagree	2-3 times during a month	nor good or bad	about once a month
disagree	once a week	bad	about once a week
totally disagree	several times a week	very bad	almost every day

3 Solution and methods

The solution can be divided into two processes. The first process trains the system with collected data and research-based knowledge to identify typical patterns in well-being data. The second process utilizes the trained system for providing automated feedback to individual student or school staff.

3.1 Training process

The training process prepares the system for actual use: providing automated feedback. The process consists of the six phases presented in Figure 2.

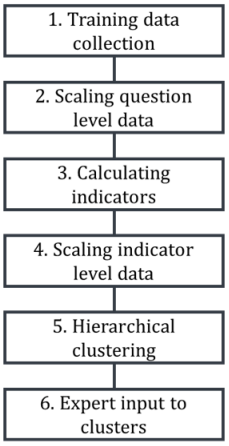

Figure 2. Training process has six phases.

Phase 1

In the first training phase the data set is collected from a large number of students by using the anonymous web forms of the School Well-being Profile as explained in Section 2. In this study we restricted the data to the students who had answered all 81 questions in the form. The incomplete answers were removed from the training data set, which resulted in a total of 64,139 answers.

Phase 2

This phase performs scaling for individual answer data when necessary. All questions are answered with same five-point Likert scale, but questions have differences in what point each answer is alarming. This phase is needed to emphasize the differences of questions that are used in same indicator. The selection of scaling function for each question should be based on the research of students' well-being. In our study, only 2 questions ("*classmates intervene in bullying*" and "*teachers are interested in my situation*") were scaled individually in order to equalize their importance, and to prevent them from over emphasizing the indicators they were included in. For the other questions the indicator level scaling was satisfactory. For the 2 individually scaled questions, a mapping from the linear scale [1 2 3 4 5] to the scale [1 3 4.2 4.7 5] was performed. The scaling procedure is discussed in more detail in training phase 4.

Table 2. Indicators used in research organized in 4 main categories. Reliability is calculated as Cronbach's alpha.

Indicator	Questions	Scaling function	Reliability, alpha
i. School conditions			
1 physical conditions	11	A	0,89
2 coziness	5	A	0,83
3 schoolwork	4	B	0,76
4 rules	2	A	0,82
5 health and social services	4	A	0,83
ii. Social relationships			
1 class community	4	B	0,90
2 bullying	3	C	0,51
3 schoolmates	3	B	0,83
4 student-teacher relationships	4	B	0,88
5 parents' participation	5	C	0,90
iii. Means for self-fulfillment			
1 students' appreciation	4	B	0,85
2 students' participation in development	1	A	-
3 teachers' expectations	1	A	-
4 students' own capabilities	9	B	0,93
5 learning support	7	B	0,92
6 optional subjects and clubs	2	A	0,44
iv. Health status			
1 perceived health	1	B	-
2 pains	4	B	0,81
3 moods	3	B	0,84
4 sleep	2	B	0,73
5 fear	1	C	-
6 common colds	1	A	-

Phase 3

The scaled answer data was combined into 22 groups, which are shown in Table 2. The smallest group had only one question and the largest one had 11 questions. The groups were formed based on topic association. The basis of the grouping was both the School Well-being Model and the 20 years accumulated research knowledge of the students' answers.

Secondly, the reliability (Cronbach's alphas) of the groups were tested. Although two of the groups did not have high alpha, they were kept as they were. Those were the group "*bullying*" (to be a victim or to be a perpetrator and if classmates intervene if a student is being bullied) and the group "*optional subjects and clubs*" (which means do the students have choices for their studies and organized free time in schools).

Indicators were formed from each group by calculating the average of all the questions' data in the group, using the scaled data for the two questions as described above. The reliabilities of the indicators were good except for bullying and optional subjects and clubs (see Table 2).

Phase 4

The indicators are scaled to have equal importance and sensitivity to make their values comparable. This type of scaling also equalizes the meaning of the indicators in the distance metrics used in the next phase. The scaling is achieved by applying piecewise linear functions (Kumpulainen, 2009), which also normalizes the data from range [1, 5] to [0, 1]. In this study we used three scaling functions. Table 2 shows which one was applied to each indicator. The shapes of the scaling functions are presented in Figure 3.

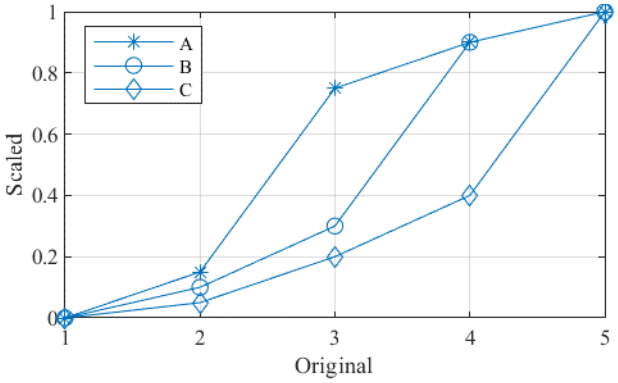

Figure 3. Scaling functions used in the research. Letters refer to Table 2.

Scaling function A is the most forgiving one. Even answer 3/5 is reasonably good in the scaled space [0, 1]. Function B scales answer 4 to a reasonably good value, but 3 is already alarmingly low. Function C is the most sensitive: immediately when answers start to deviate down from 5, the scaled value drops quickly.

Phase 5

Clustering is an unsupervised method for identifying groups in data (Everitt, 2001). Similar observations are grouped together based on the data, without any prior knowledge about the groups. In this study, clustering was used to identify the typical patterns

in the well-being profiles. The data set was clustered by hierarchical clustering with Ward linkage (Ward, 1963) and Euclidean distance metric. Clustering is an unsupervised method where the ground truth may exist but is not known. Therefore, there are no objective and absolute criteria for optimal clustering, including the number of clusters. The assessment of the results depends on the task at hand and the goals of the process. The goals are based on what information the clustering is expected to provide (Hennig, 2015). Clustering, as well as the validation of the results, should not be treated as an application-independent mathematical problem, but should always be studied in the context of its end-use (Luxburg 2012). The aim of clustering in this case was as described by Hennig page 55: "*information reduction and structuring of sets of entities from any subject area for simplification, effective communication, or effective access/action such as complexity reduction for further data analysis, or classification systems*".

Phase 6

In the final training phase, an expert associates feedback messages for each result cluster, which will be later used in the trained system as feedback to students. Feedback should be based on research, but on the other hand formulated in a motivating and constructive manner. At this point the feedback is created by a research expert in the area of school well-being. In the future, the feedback will be created, tested and discussed by a professional team, which will include researchers, psychologist, student counsellor, teachers, parents and students.

3.2 System application process

The system application process is straightforward. With current computational speeds, automated feedback from the system application process can be given instantaneously.

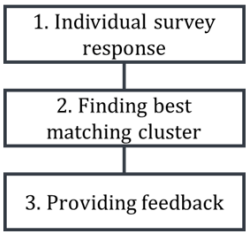

Figure 4. The process to use the trained system has three phases.

In the first phase of using the system, an individual student answers the survey questions anonymously. The survey is the same that has been used to collect the data for the system.

Secondly, the response data is processed as in phases 2, 3 and 4 of the training process. The individual questions are scaled, the indicators are calculated and scaled as specified in the process. The best matching cluster to student's answer is detected by finding the minimum distances between the students processed answer and each cluster's mean vector.

Finally, the feedback made based on well-being research in the final training phase can be given to the student anonymously already during the same session he or she has answered the survey questions.

4 Results

In the clustering phase of the system we tested several clustering criteria for determining the number of clusters. The results were not useful, suggesting either very few (2 to 3 clusters), or as many as possible. Therefore, we selected 20 clusters, which provided sufficiently informative and meaningful clusters for the purpose of this work. It is important to note that each cluster has its own characteristics. Sometimes the clusters vary in a similar manner, but then the levels of the indicators were different. The process works similarly independently of number of the result clusters as long as the clusters are meaningful for the system functionality. What is meaningful depends on the specific aim of clustering in the application of interest (Hennig, 2015). Clustering validation indexes can be used to find preferred number of clusters, but it has been argued that experts are more consistent in assessing clusterings (Lewis, 2012).

We selected five result clusters to show examples about expert feedback to students or possibly to school staff. Positive feedback can be always given straight to the student. Feedback about possible issues with well-being should be given carefully with an encouraging tone. In some cases, the student needs help from adults. The system should provide information for the student on how to get help from teacher(s), the school nurse or counselor.

Cluster 3: Everything seems to be well. The main concern is sleep or tiredness. It might have some effect also to the minor decline in the schoolwork indicator. The feedback given to student: "*Great, it seems that you are doing really well. Could you sleep a bit more? Are you tired during school hours?*"

Cluster 9: Both social relationships and studies seem to be fine. There is a minor decline in parents' support and schoolwork, but the main concern here are the symptoms (pains, mood, sleep/tiredness and fear). The feedback given to student: "*You seem to do really well with schoolmates and teachers. Also, your studies go fine, but would you like to speak to some adult (teacher, counsellor, school nurse, parents) about your pains, mood and tiredness.*"

Cluster 14: Health status is good. However, there seem to be some problems with schoolwork, class community, teachers, parents, student's appreciation, studies and student's own capabilities. This might indicate the existence of some bullying. Are teachers' expectations too low? The feedback given to student: "*You seem to be very healthy student. Do you need more help with your studies from your teachers and parents?*"

Cluster 16: Relations with schoolmates seem to be fine, less so with teachers and parents. A small concern about bullying exists. Experiences about schoolwork, student's appreciation and possibilities to influence school development are low. The main concern is health. There seem to be many symptoms, like pains, low mood, sleep/tiredness and fear. The feedback given to student: "*You are doing well with your schoolmates, but a bit less so with teachers. Do your parents support you enough with schooling? Your experiences about studies and your capabilities are quite fine. Are there some organizational factors in schoolwork that disturbs you? Do you feel that you are not appreciated, and you are not heard when developing school? You could speak about your pains, mood, sleep/tiredness and fear to some adult at school and at home.*"

Cluster 18: The main relief is that student experiences him/herself healthy, although there exist many concerns. Problems with rules and teachers as well as with studies dominate the picture. Some light is seen in the relations with schoolmates and support from parents. Also, feelings about own capabilities are not very low. The feedback given to student: "*You seem to have problems with rules and teachers. Do you have feelings that you are not appreciated at school and you have no influence in school development? It seems that teachers expect too much from you and you do not get enough support. Great that you are doing well with your schoolmates and parents. Great thing is that you seem to be very healthy.*"

5 Discussion

The starting point for this work was the research done for School Well-being Model and the data collected for that framework. We developed a methodology to automate the well-being feedback process.

During the work we learned, that scaling and normalizing the data as well as calculating meaningful indicators from the data is necessary. Defining scaling functions and grouping the questions to indicators was based on expert opinions, but it was necessary in order to preprocess the data. With this manual input the resulting clusters were meaningful and easier to interpret than those achieved using raw data without proper preprocessing.

Based on this research, we can conclude that hierarchical clustering is a viable method for finding student well-being related patterns. When the patterns were visualized to well-being research specialist, she was able to explain the differences among clusters as well as give a feedback that could be given to student. However, the feedback should be ethical and carefully planned. Each time when feedback is given, there should be a plan of who will see the feedback, in addition to student him/herself, and what are the actions that should be taken to help the student.

Several phases of the training process rely on experts' input. For instance, defining the scaling functions, choosing the meaningful number of clusters and composing the feedback messages to students. Obviously, human input provides a risk for the validity of the feedback, but the aim of the method is to automate a process that otherwise is purely manual and, as such, even more prone to human errors.

The scaling functions were defined by experts, who were familiar with the data and well-being as a phenomenon. The method presented could be improved, if the scaling function definition could be partially or completely automated. However, it raises a dilemma, that is how to capture the importance of a question asked of a student in such an automated process. For instance, safety is commonly considered a more important well-being factor than school yard coziness.

Validation of the presented method can be done by repeating the process using a different data set. To test the applicability to other areas the data set could also be from some other domain than school well-being.

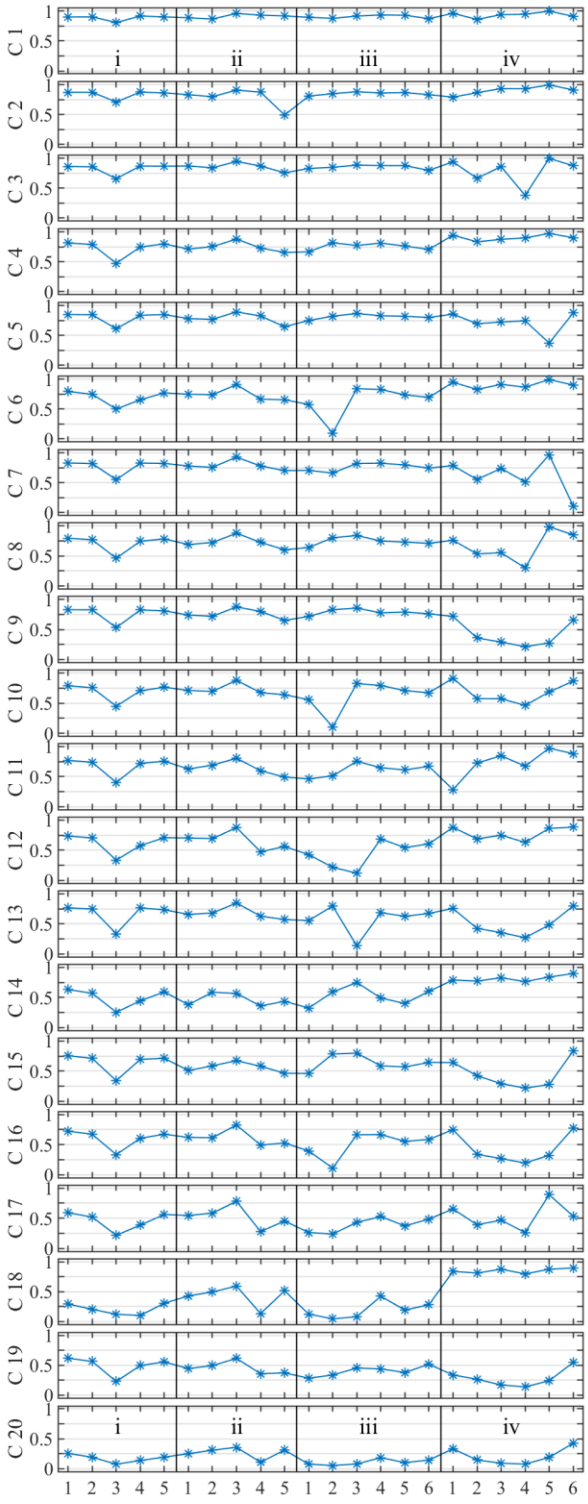

Figure 5. The mean vector of each cluster. Clusters are ordered descending based on average answer. Numbers i-iv refer to the indicator groups in Table 2, and the numbers on x axis refer to the indicators in each group.

Because this research focuses in the data processing method, user evaluation would be good for validating the feedback content correctness as well as how useful receivers perceive it.

For further research, we found several interesting topics. Firstly, the clustering method could be further developed so that also possible anomalies were found. For instance, the samples, that do not belong well to any of the result clusters could be analyzed by specialists. They are potential well-being patterns that are not common but are problematic for smaller subset of students.

Secondly, repeating the same research with data from several countries. It would be interesting to see what kind of differences there were in the result clusters.

Thirdly, we discussed in Chapter 4 about the number of clusters. We ended up with 20, since well-being researchers were assessing that as the best result. The vast number of existing clustering criteria could be tested to find if they would provide useful results for this purpose. Another option is to develop novel clustering criteria that could detect a useful number of clusters purely based on data, without expert supervision.

Finally, we believe, that the system can be significantly improved, if a feedback system from the users, and possibly a self-optimization mechanism, would be included in it.

ACKNOWLEDGMENTS

Big thanks to Mr. Mika Kasanen from School Day Helsinki for making this research possible. We thank Mr. Gordon Alford for the help with language.

REFERENCES

[1] Kylväjä Mikko, Kumpulainen Pekka, Hätönen Kimmo, 2005, Information summarization for network performance management. 10th IMEKO TC10 International Conference on Technical Diagnostics

[2] Kumpulainen Pekka, Kylväjä Mikko, Hätönen Kimmo, 2009, Importance of scaling in unsupervised distance-based anomaly detection. 19th IMEKO World Congress 2009

[3] B. Everitt, S. Landau, M. Leese, Cluster analysis. Edition: 4, Arnold, 2001.

[4] Ward Jr., J.H. (1963). Hierarchical grouping to optimize an objective function. Journal of the American Statistical Association 58 (301), pp. 236-244.

[5] Christian Hennig, What are the true clusters?, Pattern Recognition Letters, Volume 64, 2015, Pages 53-62, ISSN 0167-8655, https://doi.org/10.1016/j.patrec.2015.04.009. (http://www.sciencedirect.com/science/article/pii/S0167865515001269)

[6] Konu Anne and (Tampereen yliopisto): Koulun hyvinvointiprofiili 2017-2018: yläluokat 7-9. Versio 1.0 (2018-06-18). Yhteiskuntatieteellinen tietoarkisto. http://urn.fi

[7] Konu Anne and Rimpelä Matti, 2002, Well-being in Schools – a conceptual model. Health Promotion International, 17(1), 79-87.

[8] Konu Anne and Koivisto Anna-Maija, 2011, The School Well-being Profile – a valid instrument for evaluation. In LG Chova, M Belenguer, AL Martinez: EDULEARN11 Publications. Barcelona 4th-6th of July 2011, pp. 1842-1850, IATED.

[9] Konu Anne and Lintonen Tomi, 2006, Theory-based survey analysis of well-being in secondary schools in Finland. Health Promotion International, 21(1), 27-36.

[10] Konu Anne and Lintonen Tomi, 2006, School well-being in Grades 4-12. Health Education Research, 21(5), 633-642.

[11] Konu Anne, Joronen Katja and Lintonen Tomi, 2015, Seasonality in School Well-being: The Case of Finland. Child Indicators Research, 8(2), 265-277

[12] Langford R, Bonell CP, Jones HE, Pouliou T, Murphy SM, Waters E, Komro KA, Gibbs LF, Magnus D and Campbell R, 2014, The WHO Health Promoting School framework for improving the health and well-being of students and their academic achievement. The Cochrane database of systematic reviews, 4 (4). CD008958. ISSN 1469-493X DOI: https://doi.org/10.1002/14651858.CD008958.pub2

[13] Wyn Johanna, Cahill Helen, Holdsworth Roger, Rowling Louise and Carson Shirley, 2000, MindMatters, a whole-school approach promoting mental health and wellbeing. Australian and New Zealand Journal of Psychiatry, 34(4), 594-601.

[14] J. Lewis, M. Ackerman, V. de Sa, 2012, Human cluster evaluation and formal quality measures, in: Proceedings of the 34th Annual Conference of the Cognitive Science Society, Austin, Texas, 2012, pp. 1870–1875.

[15] Olweus, D., 1993. Bullying at school: What we know and what we can do. Oxford, UK: Blackwell Publishers.

[16] Rigby, K. & Smith, P.K., 2011. Is school bullying really on the rise? Social Psychology of Education, 14, 441-455. https://doi.org/10.1007/s11218-011-9158-y

[17] Salmivalli, C, 2009. Bullying and the peer group: A review. Aggression and Violent Behavior, 15, 112-120. https://doi.org/10.1016/j.avb.2009.08.007

[18] Ulrike von Luxburg, Robert C. Williamson, Isabelle Guyon: Clustering: Science or Art? Proceedings of ICML Workshop on Unsupervised and Transfer Learning, PMLR 27:65-79, 2012

[19] Shalini Kaleeswaran, Anirudh Santhiar, Aditya Kanade, Sumit Gulwani, Semi-Supervised Verified Feedback Generation, Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering, Pages 739-750.

[20] Mashfiqui Rabbi, Min Hane Aung, Mi Zhang, Tanzeem Choudhury, MyBehavior: Automatic Personalized Health Feedback from User Behaviors and Preferences using Smartphones, UbiComp '15 Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Pages 707-718.