Extraction of Important Characteristics for Data-Informed Guidance and Counseling from Daily Usage Log Data

Junya ATAKE^{a*}, Chia-Yu HSU^a, Izumi HORIKOSHI^b & Hiroaki OGATA^b

^aGraduate School of Informatics, Kyoto University, Japan

^bAcademic Center for Computing and Media Studies, Kyoto University, Japan

*atake.junya.86t@st.kyoto-u.ac.jp

Abstract: In many schools, teachers ensure that learners acquire academic knowledge and competences and promote their comprehensive psychological and social development through a process known as Guidance and Counseling (G&C). However, this process often involves dealing with difficult tasks and requires considerable effort from teachers. In recent years, as ICT tools have become more common, log data on daily use and learning log data have accumulated. These data have enabled teachers to understand learner processes. Utilizing data across contexts is expected to deepen learners understanding, which is the basis of G&C. However, despite these needs and potential, the use of data for G&C has not yet been fully explored. Therefore, this study examines how the log data accumulated in the Goal-Oriented Active Learner (GOAL) system can be used by teachers to understand learners. Specifically, we extracted the characteristics of each learner's situation for various contexts from the log data, which can capture learning and daily routines. We then interviewed two teachers to determine how the visualized characteristics might be useful for G&C. The results suggest that the visualized characteristics are valuable for understanding learners' conditions both inside and outside of school, and these characteristics could support teachers' G&C activities, transcending specific subjects and activities.

Keywords: Guidance and Counseling, Learning analytics, Across contexts

1. Introduction

In 21st-century schools, teachers ensure that learners acquire academic knowledge and competences but also help them develop character, morals, soft skills, and psychological wellbeing (Naraswari, 2024). In numerous countries, this educational approach is known as Guidance and Counseling (G&C), and aims to support individuals in developing useful skills, insights, and the ability to adapt to their environments, thereby enhancing their overall functionality (Dianovi, 2022; Naraswari, 2024; Lai-Yeung, 2014). Although specialized staff such as counselors or vocational teachers are sometimes responsible, homeroom teachers mainly conduct educational activities to promote learners' holistic psychological and social development. A deep understanding of students is fundamental for their functioning in the various phases of G&C (MEXT, 2022). However, even for experienced teachers, comprehensively understanding each learner's circumstances presents a significant challenge because it requires considerable effort and time (MEXT, 2022). In addition, G&C serves students and their parents, meaning such activities need to be conducted throughout the organization (Gysbers et al., 2012).

In recent years, as the use of ICT tools has become more common in education, data related to learners' learning history and log data on daily life through wearable devices have accumulated (Ogata et al., 2023; Dong et al., 2023). By using log data, attempts have been made to predict learner dropout and provide teacher interventions (Khor, 2024; Agus, 2018). In addition, extant studies have assessed learners' cross-curricular learning behaviors to better understand their complex online learning patterns and their diverse learning experiences

(Song et al., 2024). Other possibilities include the integration of log data with other data such as academic performance and attendance data to capture the overall condition and status of learners and examine the relevance of learner issues (MEXT, 2022). However, despite such needs and potential, the utilization of cross-contextual trace data for G&C has not been examined and how it can be used to support teacher interventions has not been explored.

This study thus investigates whether characteristics can be extracted from log data and utilized in G&C to help teachers gain a deeper understanding of their students. We assumed that the extracted characteristics would be utilized by homeroom teachers, who often perform G&C regardless of the specific subject matter. While communicating with teachers who conduct G&C, we examined the activities conducted under this process and the characteristics that can be utilized to understand learners.

The novelty of this study is that it examines the characteristics of deep student understanding by current teachers to realize more comprehensive data-informed G&C and determine learner variety from log data as a first step toward this goal. This is a method for teachers to gain a comprehensive picture of each learner's situation from log data and improve their G&C practices.

2. Literature Review

2.1 Theoretical Basis of Guidance and Counseling

Although the terms and concept of G&C vary slightly by country (e.g., Comprehensive School Guidance and Counseling Approach. Whole School Approach), the common point is that it is an activity that takes place in schools (Simões et al., 2021; Gysbers et al., 2012). A deep understanding of teachers' learners is fundamental for enhancing G&C activities (MEXT, 2022). Teachers have three main types of understanding of students: (1) understanding of complex psychological and interpersonal relationships; (2) professional, objective, and empathetic understanding; and (3) mutual understanding among students, parents, teachers, and staff (MEXT, 2022). In this process, it is necessary to grasp each student's needs and interests, as well as the home environment, upbringing, capabilities, aptitudes, and interests, which acknowledge the learner's presence and positive behaviors while preventing negative ones (MEXT, 2022). Teachers engage in this activity through learning activities, greetings, classroom activities, individual interviews, and communication with parents (Felder et al., 2005; MEXT, 2022). However, in G&C, the definitions and methods of understanding the learner vary, and criteria are not as clear as those of assessments in learning activities. Therefore, to understand learners based on G&C data, we first need to understand its structure and phases and define learner understanding in these phases.

In many countries, guidance and counseling practices are divided into several phases (Myrick, 1987). There are two main phases: the Developmental Phase, which focuses on promoting the holistic development of learners, and the Remedial Phase, which focuses on addressing learners' difficulties. In Japan, for example, the structure is divided into two axes, three classes, and four layers as part of the national educational policy. The two axes are Proactive and Reactive, and the three classes are "Support holistic development," "Problem preventive," and "Remedy difficult problems" as minor items. A fourth layer consists of these three classes plus "Problem preventive," which is divided into two parts, "Early detection of the problem" and "prevention education." In other words, Proactive is a developmental phase and Reactive is a prevention and remedial phase. Such a multi structural approach is used worldwide, and global trends have shifted from a remedial to a preventive approach in education in recent years (Myrick, 1987; Gysbers & Henderson, 2012; Yuen, 2002). This is because learners experience a sense of self-presence through inspiration, discovery, and creation in the classroom and during experiential activities, which builds mutual trust and prevents problems. In this sense, the developmental support phase is important.

Based on the above, we organized G&C into three phases as per Figure 1. In the Figure, the left-hand side is the developmental phase, which targets all learners and is less urgent. The right-hand side is the crisis and remedial phase, which is individualized and urgent. The intermediate phase prevents these problems. Although each phase has a different approach,

they share the same underlying first step: understanding the learner. However, the goal of understanding the learner differs in each phase, and for the homeroom teacher to understand the learner based on data, it is necessary to examine the characteristics of each phase. Therefore, this study focuses on two phases that are most likely to be captured in the data—the developmental and the preventive phases—, and examines the characteristics of these phases.

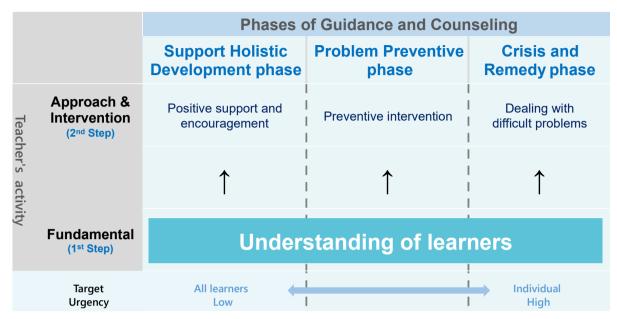


Figure 1 Multilayered Structure of Guidance and Counseling and Teacher's Activity.

2.2 Technology-enhanced Guidance and Counseling

Next, we will review related studies and examine the research gap in data-informed G&C based on the literature review of G&C and ICT, and learning analytics(LA) technology studies.

2.2.1 Guidance and Counseling with Information and Communication Technology

In Japan, the Ministry of Education, Culture, Sports, Science and Technology (MEXT) has proposed the "Comfortable, Customized and Optimized Locations of Learning (COCOLO) Plan," which uses ICT tools to monitor and understand various aspects of learners (MEXT, 2022). For example, the Awareness and Monitoring App is a web application developed for schools to facilitate early detection and prompt response to bullying (Nara Prefectural Board of Education, 2023). The system organizes signs of bullying into three levels based on 17 indicators of learners' behavior. By continuously recording observations in the app, teachers can share information about students' behaviors. This system not only aids in preventing bullying before it occurs but also supports quick intervention and the continuous monitoring of subtle changes in students' behavior. This proactive approach enhances learners' overall safety and well-being by enabling a nurturing environment, in which potential issues can be promptly addressed. Problems with this method include the need for periodic implementation, reliance on learners' subjective responses, and challenges in its application for children at a lower developmental stage.

2.2.2 Guidance and Counseling Using Learning Analytics Technology

Another method that can be considered is the use of learning analytics technology, a technique for understanding the learning process of learners and profiling users using trace data accumulated on computers and tablet devices. LA provides Monitoring, Assessment, and Analysis to offer personalized learning (Khor, 2024; Agus, 2018). This makes it possible to identify the learning progress, level of understanding, and areas of weakness for each subject,

and to provide learning support optimized to individual learners. Studies have also been conducted to prevent dropouts through at-risk detection (Oliveira, 2021). In this sense, learning analytics have contributed to understanding learners' situations from data.

There are several other examples of applying learning analytics technology in G&C. One describes the development of learning analytics solutions, including a predictive model and personal tutor dashboard, to optimize learner support and academic outcomes in Singapore's polytechnics (Hwee et al.,2023). Another study proposes an Al-powered system that predicts learners' chance of university admission and recommends similar institutions based on their academic history and preferences, using a model for high accuracy (Majjate 2023). However, the use of data in each of the multiple G&C phases, with different approaches in different phases, has not been fully examined.

2.3 Research Gap and Research Question

The above literature review indicates that the use of data in each G&C phase (developmental, preventive, and therapeutic) has not been fully explored. Therefore, this study creates indicators from log data, identifies the G&C characteristics that can be captured from these indicators, and explores their applicability to each phase. Teacher feedback and expectations regarding the extracted characteristics are then collected to identify specific challenges and areas for improvement regarding the application of technology in actual educational settings. To accomplish this, we pose the following research questions:

RQ1: What Guidance and Counseling characteristics can be captured from log data?

RQ2: What feedback and expectations do teachers have for the extracted characteristics?

3. Methodology

First, based on the literature review, we analyzed log data and created indicators and characteristics (RQ1). We then interviewed two teachers on their feedback and expectations of the extracted indicators and characteristics (RQ2). The sequence of the analytical flow is shown in Figure 2.

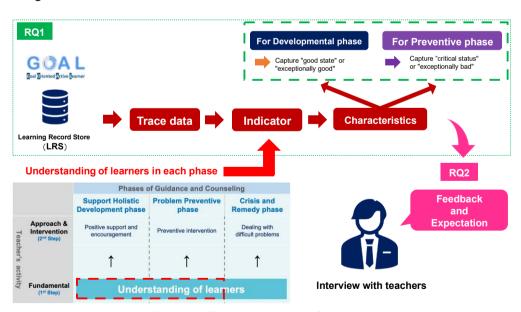


Figure 2 The Procedure of Analysis

3.1 Data Collection and Analysis

This study used log data of daily activities in the Goal-Oriented Active Learner (GOAL) system to answer RQ1. This system supports the execution of data-driven Self-Directed Learning (SDL) and the acquisition of self-directed skills through the DAPER (data collection—analyze—

–plan–execution and monitoring–reflect), process model. This approach synthesizes personal health and learning data logs from multiple sources and presents them to learners (Majumdar et al., 2019), allowing learners to see the goals they have set for themselves, activity indicators, and scores. Specifically, the SDL that learners perform includes learning activity and physical activity measured by wearing a GARMIN device; we used five datasets: math activity time, English activity time, math score, stress level, and number of steps taken. The data gathered from 200 learners cover the period from June 2020, when the students enrolled (the start was two months later than usual due to COVID-19), to March 2023, when the students graduated.

3.2 Data Processing

We attempted to integrate and scale various log data with the different contexts in the GOAL system because the meanings of these numbers vary by context. English and math activity times represent the time used for independent study, which is an indicator of engagement, such as learners' high or low interest in learning. Math scores are the learning scores on the tasks given by the teacher. For example, the activity time values were recorded in minutes per activity, while scores were recorded against the maximum score. In addition, the number of steps taken per day was recorded. For stress levels, values were recorded from 0 to 100, with values closer to 0 indicating lower stress levels and values closer to 100 indicating higher stress levels. The GARMIN device classifies stress into four levels: 0–25 shows rest, 25–50 represents low, 51–75 means middle, and 76–100 high.

Therefore, the meanings of these values differ by context. Hence, we represent these different values on a 10-point scale using the following method: values closer to 10 represent a good state of the learner, and values closer to 0 represent a critical state (Figure 3).

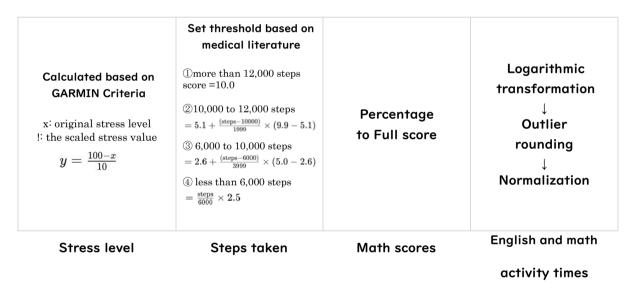


Figure 3 Scaling Logic and Scores

For example, because activity time depends on the overall distribution, we performed logarithmic transformation rounding for outliers, followed by normalization. Because normalization takes a distribution from 0 to 1, the value multiplied by 10 was used as the value after scaling. The score is expressed as a percentage of the full score on a 10-point scale. Next, because the math score values had maximum values, they were converted so that the score percentage of the full score was expressed on a scale of 10. The number of steps was then scaled based on previous studies that examined the associations between steps and adolescent health (Tudor-Locke et al., 2011; Weres et al., 2022), that is, 10 for more than 12,000 steps per day, from 5.1 to 9.9 for between 10,000 and 12,000 steps per day, from 2.5 to 5 for 6,000 to 10,000 steps, and below 2.5 for less than 6,000 steps. Finally, the stress levels were calculated. The original value was determined using the GARMIN device, which uses heart rate variability. The original scale uses 0 for a low-stress level and 100 for a high one,

where multiplying by 0.1 would have the opposite meaning. Therefore, we subtracted 100 from the recorded value, took the absolute value, and multiplied it by 0.1. In addition, the four predefined stress levels were used to give meaning to the scaled values: 10 to 7.6 = very good, 7.5 to 5.1 = good, 5.0 to 2.5 = medium, and 2.5 to 0 = critical. The scaling of these datasets resulted in the following distribution (Figure 4).

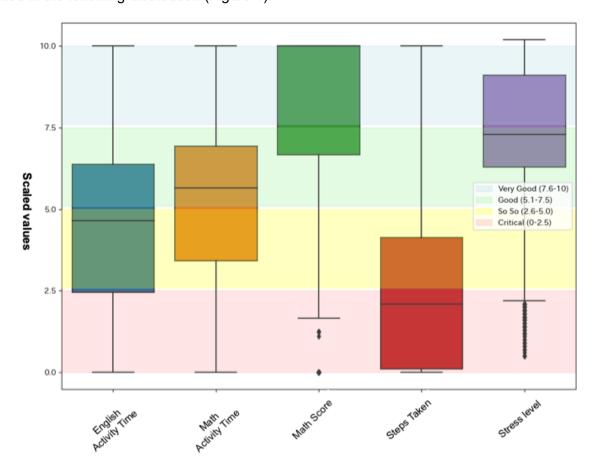


Figure 4 Box plot for scaled scores of each context

3.3 Interview Design

Semi-structured interviews were conducted with school teachers using the GOAL system to obtain their expectations and feedback on the potential of the indicators and characteristics in RQ1. We conducted the experiment twice, owing to time limitations. The first interview was conducted with the director of G&C at the school, who was also a classroom teacher. In the first interview, we asked about the actual practice of G&C at their schools, including how teachers understood students, such as their behaviors and signs, as well as the interventions they focused on during each phase. In addition, to determine the characteristics that could be captured from the log data, we asked about the status of ICT use in understanding students and their expectations of using log data. The second interview was conducted with the director of G&C and another classroom teacher. In this interview, we mainly asked for feedback and expectations regarding the visualized characteristics. The results of the semi-structured interviews informed the analysis of RQ2.

4. Results and Discussion

4.1 RQ1 Guidance and Counseling Characteristics Extracted from Data

Polar charts were used to visualize the status of learners in different contexts, representing all scaled learner values (Figure 5). This also reflects learner status. An example of this visualization is shown in Figure 6.

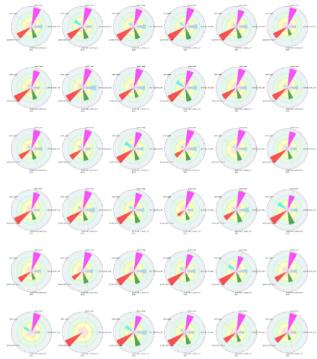


Figure 5 Example list of learner state visualizations across contexts

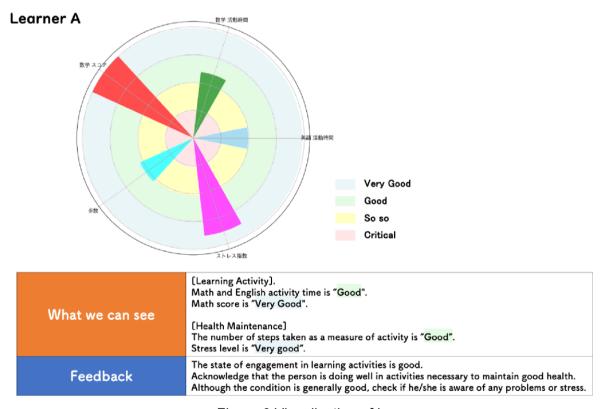


Figure 6 Visualization of learners

To answer the first research question, we analyzed GOAL system log data to identify key indicators that reflect the learners' current state. The following key indicators and characteristics emerged from the data visualization.

The first is the time and score of the learning activities. These indicators reflect learners' engagement within and achievement of educational tasks. Using these indicators, high engagement and good scores indicate positive engagement and learning attitudes, while low engagement and low scores may indicate potential problems, such as a lack of interest or understanding, which are important for student understanding.

The second is the number of steps taken. This indicator serves as a measure of physical activity, and is important for maintaining health. Using this indicator, the maintenance of a

constant number of steps is indicative of health, physical condition, and mental health, whereas a low number may indicate potential health problems or low motivation, which are characteristics of student understanding.

Finally, stress levels were considered. Stress levels are directly related to the psychological state of learners. High stress negatively affects learning and overall health. Stress level indicators can thus be important characteristics of a student's mental health and can identify students who need additional support or intervention.

Consequently, these indicators can be used to extract important cross-contextual learner characteristics from log data that are important for G&C.

4.2 RQ2 Teachers' Feedback and Expectations for the Characteristics

Semi-structured interviews were conducted with the teachers while showing them the visualization in Figure 6 and other learners' visualizations. The two teachers were asked about their expectations and feedback on their characteristics. The main questions related to RQ2 were as follows:

- Q1. How and to what extent do you think this visualization is likely to reveal the status of your students? (four-point scale)
- Q2. What changes are you likely to detect in your learners when the values of these five indicators change (increase or decrease)?
- Q3. As a classroom teacher, how do you think you can use these characteristics for G&C? Table 1 shows the results.

Table 1 Interview responses from two teachers

Question	Common answers
Q1	Of the four-point scale, responses of 3 and 4 · Time
	and effort in learning activities.
	Stress level of students.
	Both teachers responded that they could understand the learners' condition, including outside of school.
Q2	Possible changes in interest in school and learning. Possibility
	of change in mental and physical health.
Q3	To understand learners' efforts and engagement.
	To praise the student for good behavior and correct undesirable behavior.
	· As material for discussion in individual meetings with each learner.
	As material for discussion in meetings with their parents.

From the interviews, teachers considered the characteristics valuable for understanding learners' conditions, both in and out of school, and believed it could help detect changes in their interest in school and learning, as well as their mental and physical health. Additionally, these characteristics can be used in various aspects of G&C such as recognizing and encouraging positive behavior, addressing undesirable behavior, and serving as discussion points in meetings with learners and their parents. One teacher suggested that knowing the stress status can help them better understand how students are doing. He responded that the value of the indicator is not everything, but that even children whose numbers are not very high need to be looked at, providing an opportunity to communicate with the student, as he/she may be trying too hard. This feedback indicates that teachers have expectations for the potential of these characteristics to enhance the understanding of learners and provide effective G&C tailored to each learner's needs.

Some suggestions were made based on questions not directly related to the RQs. First, these characteristics are consistent with teachers' understanding of learners. Specifically, when we asked "What do we expect to learn about learners from log data?," teachers

responded that (1) it would reveal activities that are not limited to specific subjects, such as learning activities, exercise, and sleep; (2) it would be coherent, such as a diary; and (3) it would provide a glimpse into learners' inner aspects. Currently, teachers ask learners to record their learning times and reflections on a form (paper-based) to understand their situation. However, they wanted the recorded data to be visualized in a manner that was not selfreported. These findings suggest that these characteristics may be useful for teachers' understanding of their students. In addition, when asked whether they had ever used learning analytics tools for G&C, the answer was no for both. This was because they had not used these tools to understand learners from the perspective of teaching each subject area. From this perspective, teachers desire to develop a multifaceted understanding of their learners using log data and other digital tools.

In summary, the extracted characteristics are useful for understanding learning activities and stress states, learners' effort and physical activity, and possible changes in interest and health status. Based on these results, teachers can better understand their students' situations through the extracted characteristics and provide appropriate G&C to each student.

5. Conclusions and Future Work Directions

This study explored the potential use of data in supporting G&C by using log data stored in the GOAL system to characterize and visualize learner situations through semi-structured interviews with two teachers. The results indicated that the characteristics extracted from the log data are useful for teachers to understand the time and effort status of learners' learning activities, stress states, and health statuses.

One of the key findings is that the visualized characteristics are valuable for understanding learners' conditions both inside and outside school. Teachers can use these characteristics to recognize positive behaviors, address undesirable ones, and serve as discussion points during meetings with learners and their parents. In addition, log data can be used to understand a wide range of activities and habits. Based on these findings, it is expected that teachers will better understand learners' situations through the extracted characteristics and provide appropriate G&C in each phase.

There are two main areas for future research. The first is the further expansion and integration of log data. In this study, we used log data from the GOAL system, which integrates data from daily use. However, to use these data for G&C, it is necessary to consider data such as learners' learning logs for many subjects and life habits. Furthermore, by integrating diverse data, such as socio-emotional aspects and family environment that influence student behavior and outcomes, it will be possible to analyze learners' situations to understand them from a more multifaceted perspective.

Second, only two teachers were interviewed in this study. The sample size should be expanded. G&C are based on teachers' understanding of their students, but simply visualizing data and providing feedback are not sufficient to encourage teachers to become more aware of their students. In addition, developing a system that detects abnormal stress levels or sudden changes in behavior and feedback based on teachers' diverse views is important to support teachers' decision-making for G&C intervention.

These directions will improve the effectiveness of data-informed G&C and contribute to a more adequate educational environment.

Acknowledgments

This work was partially supported by the following research grants: JSPS KAKENHI (A) JP23H00505 and(B) JP22H03902 and the National Institute for Educational Policy Research: Educational Data Analysis and Research Promotion Project FY2023-2025.

References

Naraswari, I. A. M. D., Dantes, N., Suarni, N. K., Gading, I. K., & Suranata, K. (2024). Solutionfocused brief counseling to improve student's social-emotional skills and psychological wellbeing. Jurnal EDUCATIO: Jurnal Pendidikan Indonesia, 10(1), 106-113.

- Dianovi, A., Siregar, D. ., Mawaddah, I. ., & Suryaningsih, S. (2022). Guidance and Counselling in Education. World Psychology, 1(2), 99–107. https://doi.org/10.55849/wp.v1i2.95
- Lai-Yeung, S. W. C. (2014). The Need for Guidance and Counselling Training for Teachers. Procedia Social and Behavioral Sciences, 113, 36-43. https://doi.org/10.1016/j.sbspro.2014.01.008
- Hiroaki OGATA, Rwitajit MAJUMDAR, Brendan FLANAGAN, Learning in the Digital Age: Power of Shared Learning Logs to Support Sustainable Educational Practices, IEICE Transactions on Information and Systems, 2023.
- Qian Dong and Rong Miao. 2023. Wearable Devices for Smart Education Based on Sensing Data: Methods and Applications. In Learning Technologies and Systems: 21st International Conference on Web-Based Learning, ICWL 2022, and 7th International Symposium on Emerging Technologies for Education, SETE 2022, Tenerife, Spain, November 21–23, 2022, Revised Selected Papers. Springer-Verlag, Berlin, Heidelberg, 270–281. https://doi.org/10.1007/978-3031-33023-0 24
- Gysbers, N. C., & Henderson, P. (2012). Developing & managing your school guidance & counseling program (5th ed.). American Counseling Association.
- Yuen, M. (2002). Exploring Hong Kong Chinese guidance teachers' positive beliefs: A focus group study. International Journal for the advancement of Counselling, 24(3), 169-182.
- Yige Song, Eduardo Oliveira 0001, Michael Kirley, Pauline Thompson. A Case Study on University Student Online Learning Patterns Across Multidisciplinary Subjects. In Proceedings of the 14th Learning Analytics and Knowledge Conference, LAK 2024, Kyoto, Japan, March 18-22, 2024. pages 936-942, ACM, 2024.
- Khor ET, K M. A Systematic Review of the Role of Learning Analytics in Supporting Personalized Learning. Education Sciences. 2024; 14(1):51. https://doi.org/10.3390/educsci14010051
- Agus, R. (2018). Learning Analytics Contribution in Education and Child Development: A Review on Learning Analytics. Asian Journal of Assessment in Teaching and Learning.
- Simões, C., Caravita, S., & Cefai, C. (2021). A Systemic, Whole-School Approach to Mental Health and Well-Being in Schools in the EU. Analytical Report. European Union.
- Felder, R. M., & Brent, R. (2005). Understanding student differences. Journal of engineering education, 94(1), 57-72.
- Myrick, R. D. (1987). Developmental guidance and counseling: A practical approach. Educational Media Corporation, PO Box 21311, Minneapolis, MN 55421.
- Tudor-Locke, C., Craig, C. L., Beets, M. W., Belton, S., Cardon, G. M., Duncan, S., ... & Blair, S. N. (2011). How many steps/day are enough? for children and adolescents. International Journal of Behavioral Nutrition and Physical Activity, 8, 1-14.
- Weres, A., Baran, J., Czenczek-Lewandowska, E., Leszczak, J., & Mazur, A. (2022). The association between steps per day and blood pressure in children. Scientific Reports, 12(1), 1422.
- de Oliveira, C. F., Sobral, S. R., Ferreira, M. J., & Moreira, F. (2021). How does learning analytics contribute to prevent students' dropout in higher education: a systematic literature review. Big Data and Cognitive Computing, 5(4), 64.
- Hwee Theng, L., Patrick, P., & Tien Chern, C. (2023). Enhancing pastoral care and academic monitoring efforts using the AiE-PFP learning analytics solutions. Paper presented at the 16th International Symposium on Advances in Technology Education, Matsue, Japan. Retrieved from
- Majjate, H., Bellarhmouch, Y., Jeghal, A., Yahyaouy, A., Tairi, H., & Zidani, K. A. (2023). Al-Powered Academic Guidance and Counseling System Based on Student Profile and Interests. Applied System Innovation, 7(1), 6.
- Majumdar R., Li H., Yang Y.Y., Akçapınar G., Flanagan B. and Ogata H. (2019). Adaptive Support for Acquisition of Self-direction Skills using Learning and Health Data, in Proceedings of the 19th IEEE ICALT pp.54-56, Maceió, Brazil July 2019.
- Ministry of Education, Culture, Sports, Science and Technology. (2022) Outline of Guidance and Counseling [in Japanese] https://www.mext.go.jp/content/20230220-mxt_jidou01-000024699201-1 pdf
- Nara Prefecture Board of Education. (2023, June 1). Nara Kodomo Mimamori System Operational Guidelines. Nara Prefecture Board of Education. https://www.pref.nara.jp/secure/296324/3-5.pdf