DESIGNING FOR MOTIVATION, WHEN THERE IS NONE: LESSONS OF DESIGNING A PYTHON-BASED DATA LITERACY PROGRAM MANDATORY FOR ALL STUDENTS

Johannes van Deest and Salomey A. Addo
Leuphana University of Lueneburg, Germany
deest@leuphana.de

With data exponentially growing lately, data literacy has become essential for individuals and organizations to prosper. Nonetheless, it is rare and seems an unattainable goal to the majority of people. Researchers, policymakers, and educators are exploring ways of getting learners excited about data, and the Leuphana University of Luneburg is not left behind in this quest. Starting in the winter semester of 2021/2022, all 1,450 first-year students will undergo basic data literacy education at Leuphana, which includes the teaching of related knowledge and skills as well as relevant ethical issues. In this paper, we theoretically explore how to design teaching-learning settings that, given heterogeneous learning groups, can transform initially unmotivated students into intrinsically motivated lifelong learners of data competencies. Furthermore, in light of our empirical experience, we derive and discuss design principles that will enable readers to find a context-adapted solution.

INTRODUCTION

In recent times, data has been the starting point for knowledge generation, value creation, and the basis for making better decisions for individuals, organizations, and economies. Our world has changed through digital transformation and so have the capabilities that are required to succeed. Successfully navigating and prospering in a digital world requires us to bring a wide range of digital competencies to the table (Carretero et al., 2017). Being able to manage, make use of, and critically assess data - known as data literacy (Ridsdale et al., 2015) - is one of the most desired skills amongst many in our society today. By all means, there is a plethora of literature that answers the question of why data literacy is a desired qualification (Ridsdale et al., 2015).

Through the development of frameworks, experts (e.g., from the EU Science Hub) have made these digital and future-relevant competencies tangible, for both broader and more specific proficiency areas, such as data literacy (Carretero et al., 2017; Ridsdale et al., 2015). However, these same experts also agree that the field of digital and technology-based know-how is so broad and subject to constant evolution that lifelong learning is becoming an imperative (European Commission, n.d.) and yet no individual can become an omniscient expert (Schüller et al. 2019).

Educational institutions and stakeholders have then begun to incorporate parts of the aforementioned frameworks into their curricula and are experimenting with approaches to deliver them.

In this paper, we focus on a didactic difficulty many educational settings face: intrinsic motivation for lifelong learning of digital competencies is meant to be triggered, but the mandatory nature of graded courses creates only extrinsic motivation. Therefore, we will address the following guiding question in this paper: *How do we design teaching-learning settings that, given heterogeneous learning groups, can turn students without initial motivation into intrinsically motivated lifelong learners of data competencies?*

Literature distinguishes extrinsic and intrinsic motivations, or lack thereof: amotivation (Ryan & Deci, 2000a). Intrinsic motivation is understood to mean that an activity is performed for its own sake and there are no external incentives (Ryan & Deci, 2000a). It is believed to last longer than extrinsic motivation, which is triggered by external rewards, constraints, or punishments (Tohidi & Jabbari, 2012; Ryan & Deci, 2000a). Both types of motivation are needed for a successful learning process (Ozcelik et al., 2013). Extrinsic motivation can provide an initial push to initiate learning, while intrinsic motivation provides self-motivation for ongoing learning (Li & Lynch, 2016).

To answer our guiding question, we explore promising approaches to intrinsic motivation and effective learning and, in light of our empirical experience, derive and discuss design principles that will enable readers to find a context-adapted answer to our guiding question.

APPROACH

In October 2018, third-party funded project “Data Driven X (DATAx)” started with the proclaimed goal of reaching all approximately 1,450 freshmen of mostly non-technical degree programs, to make them data literates and life-long learners of data literacy. We iteratively developed and established an interdisciplinary data literacy teaching program, tackling all dimensions of data literacy - knowledge, skills, and values (Ridsdale et al., 2015) - in different formats: A lecture for knowledge transfer, a tutorial for skill transfer, and a term paper for critical reflection and development of data ethics. The content of the formats is thereby not canonically oriented toward frameworks, but builds on the students' prior knowledge and includes other digital competencies deemed relevant. In this context, we consider especially the understanding of computers and their functionalities, as well as their handling and associated ways of thinking ("algorithmic/computational thinking") as desirable. Also, we believe that a better understanding of computers improves data analysis skills, for example through the improved understanding of how digital data is created. In view of this, our data literacy program is designed to teach practical data analysis skills using the Python programming language, even though this creates perceived additional hurdles. Python was thereby chosen because it gives students access to a wider variety of application contexts and potential learning paths. In addition, Python is very popular for predictive data analysis (machine learning), the application of which is the subject of courses to follow in higher semesters.

Development Phases and Feedback Instances	Fall 19/20: Phase I	Fall 20/21: Phase II	Fall 21/22: Phase III
Student Numbers	250	600	[1.400]
Course Numbers	10	35	[42]
Mode	Face-to-Face	Online	Online
Formats	Co-Learning Tutorial	Tutorial & Co-Learning Office Hours	[Tutorial & Co-Learning Office Hours]

Figure 1. Development stages and key facts of the DATAx projects’ data literacy tutorials

Our mandatory tutorials, in which students learn the basics of programming with Python for data analysis, have students enrolling with no or limited programming background. They thereby provide a self-contained setting with heterogeneous learners and diverse levels of prior knowledge for investigating our guiding question. Figure 1 schematically depicts their development path and shows the central surveys (S) and feedback workshops (F) that are part of our empirical lessons. Underpinned by the theory in the next chapter, they form the basis of our designing of future tutorials.

How to design teaching settings that inspire life-long learning around data?

In our data analysis tutorials, two things make inspiring students' intrinsic motivation particularly difficult. First, students are often emotionally distant and have misconceptions about programming due to previous negative experiences (Bain et al., 2014). Second, the exercise itself can be very frustrating, as one inevitably stumbles over error codes (Murphy & Thomas, 2008).

In our search for an appropriate theoretical approach to elicit intrinsic learning motivation in students, the heterogeneity of learners limits us to look for theories with general validity. We, therefore, focus on one of the sub-theories of the *self-determination theory* (SDT) from Ryan and Deci (2000b), the *Basic Psychological Need Theory* (BPNT). The BPNT theory, whose magnitude is expressed by a rapidly increasing number of empirical studies (Vansteenkiste et al., 2020) has been attested valid in a wide variety of contexts in meta-analyses (Van den Broeck et al., 2016). BPNT postulates that humans have three psychological needs - the need for *autonomy*, the need for *competence*, and the need for social *relatedness* - and states that their degree of fulfillment during any action has a high impact on an individual's psychological well-being, motivation, and engagement for that given action. Here, needs can sometimes conflict or mutually influence each other and the strength of needs might vary between individuals (Ryan & Deci, 2000b). We discuss this in the following chapter on "Design".

Given the theory's wide applicability and ability to make the individual’s self-regulated motivation a controllable construct, it is useful for the designing of teaching and learning settings

(Golapan et al., 2017). It helps make implementation decisions and identify areas for improvement by sharpening our awareness of situations where needs are not being met.

However, these decisions are also influenced by how one understands the learning process. On the one hand, we learn by interacting with artifacts, and literature reveals the more attractive (Keller, 2008) and reality-based they are (Ridsdale et al., 2015), the more successful is the learning outcome and sustaining of motivation. On the other hand, we share the perspective of recent learning theories that individuals sharpen their understanding through interaction with others and attach meaning to knowledge (Golopan et al., 2017; Wenger, 2010). What we know and learn defines who we are and want to be and vice versa (Wenger, 2010). Hence, social learning environments should always inspire a desire to let the learned become part of one's identity (Wenger, 2010). Research shows that usually, the teacher plays a crucial role in this regard (Hattie, 2012).

In our work, we reflect on these theoretical approaches against the background of our own empirical experience and translate them into design imperatives to answer our guiding question. The structure of the analysis process as well as the outline of the following chapter were inspired by a perspective from the organizational theory article on organizational design archetypes by Greenwood and Hinings (1988). They suggest that an organizational design consists of only three elements - interpretive schemas (values), processes, and structures.

IMPLICATIONS FOR PRACTICE

As mentioned above, we adopt an organizational lens that differentiates into values, processes, and structures to capture the totality of DATAx tutorials and their design to galvanize intrinsic motivation. However, because structures and processes are often not sharply distinguishable, there will be two subsections: one on the values and the other on the processes and structures behind our project and the tutorials. In both sections, we refrain from reproducing every detail of our course designs throughout the different phases, especially within structures and processes, as they have changed a lot. Rather, we aggregate our experiences into five central design aspects and principles that offer an enhanced transferability to the reader's context. This synthesis will be discussed against the background of BPNT and our empirical experiences with individual design choices in the two pilot phases 2019/2020 and 2020/2021. Furthermore, we highlight key tensions and challenges that must be considered when making design decisions alongside these principles and illustrate the result of negotiating these tensions with our plans for the upcoming winter semester.

Values

Values are indispensable. They shape organizations over time and often outlast processes or structures (Greenwood & Hinings, 1988). Moreover, they are essential for individuals' alignment with the organization, for the notion of being part of the organization, and ultimately for engagement in the organization (Wenger, 2010). Hence, they must be included in the design imperatives for a motivating learning environment. The underlying values of our course remained largely unchanged across all stages of development.

Design principle 1 - Cultivate a supportive atmosphere: As it is at the core of DATAx projects' values to reach all students, *inclusivity* is one of the central values. As a necessary condition for inclusivity and the creation of social bonds, all forms of discrimination must be avoided. Therefore, we adopt a Code of Conduct (CoC) identical to and adapted from the Berlin Code of Conduct (https://berlincodeofconduct.org) to normalize social behavior and exclude all forms of discrimination, thus creating fertile ground for the emergence of social relatedness. This value informs all program decisions, for example, from recruiting teaching staff to deciding which datasets students should work with.

In addition, we actively promote peer support, but at the same time introduce a so-called "honor code". This is intended to discourage students from plagiarizing by making it clear that good help does not consist of passing on solutions and that cheating destroys potential competence opportunities. The honor code was also introduced as a consequence of the increasing incidence of plagiarism in the digital-only mode of 2020/2021 and will be emphasized in the courses of 2021/2022.

Design principle 2 - Cultivate a digital growth mindset: We believe that *everything can be learned* and success is the result of consistent efforts but not solely talent. By focusing on effort, we offer an autonomous choice: it is up to the students to become good at something. Not only does this mindset grant autonomy, by making it an active choice to become good at something, but it also limits the ubiquitous focus on performance in higher education and the perceived importance of grades. Moreover, we believe that *failures are learning opportunities*. Encountering errors is imminent in programming activities. Hence, it is important to build resilience by reframing and embracing them. By teaching a growth-oriented mindset (Yaeger & Dweck, 2012), we transmit those two values to our students. The goal is to change students' learning perceptions ("*I can't do it yet, but soon I will*" instead of "I can't do it"). This leads to an attribution of the mistake made to a lack of effort, rather than a lack of talent (Murphy & Thomas, 2008). Through this intervention, for which studies have shown that it can have an immediate significant positive effect on academic achievement (Yaeger & Dweck, 2012) that is not only limited to the subject at hand but other skill areas as well (Murphy & Thomas, 2008), we create a safety net for negative competence experiences. In our first phase, faculty members were briefed to intentionally target this mindset change. Unfortunately, this aspect has been lost in the wake of Corona-related changes and has not been systematically implemented. For 21/22, the intervention is a permanent part of instructor preparation and a "Growth Mindset" manifest is planned for students as digital delivery.

Structures & Processes

Structures and processes have changed, sometimes dramatically, in recent years due to evolving contextual conditions, such as changes in curricula integration approaches, our scaling (see Figure 1), and the pandemic situation. Our first phase was not affected by the latter, and the course was designed primarily as a synchronous face-to-face class with few digital support structures. However, the second phase had to be in an "online-only" mode. The planned third phase now includes many digital elements and some flexible ones that could optionally take place in presence. From the combination of BPNT and our empirical experience in Phase 1 and 2, we derive three additional design principles that have proven particularly helpful in developing processes and structures.

Design principle 3 - Independent social learning: This principle emphasizes the relevance of learning in social contexts. First, because this is how we construct our knowledge and negotiate meaning (Gopalan, 2017), and second, because it is one of the basic psychological needs to be met. At the same time, it also illustrates that we all have different learning styles, backgrounds, personalities, and interests that need to be considered. This balancing act of giving individual space to think through and solve problems and group cooperation needs to be resolved.

In 2019/2020, we attempted to strike this balance by designing our course as a co-learning format led by an instructor (90 minutes weekly). Students worked on self-study materials, but were not alone in doing so, an environment was created to allow cooperation among the students. This approach was very well received by them, but could not be maintained due to the pandemic situation. In 2020/2021, we tried to create social cohesion within courses by establishing fixed learning groups at the beginning of the semester. Of course, this only works to a limited extent under purely digital conditions. In addition, we reduced the amount of face-to-face teaching with instructors (90 minutes bi-weekly) and introduced a new format with teaching assistants (students of higher semesters), in which they supported the small groups with advice and assistance in their learning process: weekly office hours (20 minutes weekly). Although some of these peer groups did not function ideally in terms of group coherence, the post-course survey indicated that 56% of students perceived great support from their peers and 89% perceived support from their instructors. Students also indicated that office hours were by far the most popular element of the course. Based on the good feedback, we will maintain this in 2021/2022 and intensify it in a way that office hours will be extended (30 minutes weekly) and processes to create harmony will be scheduled. The office hours can and should also be conducted in presence on campus if the situation allows. Teaching with faculty, which will continue to be scheduled online, has also been changed to weekly 45-minute sessions, and additional interactive elements have been provided, such as small assignments to be worked on together in breakout sessions in Zoom or bigger ones centered around a data analysis process, where groups of 9 students work as a team, but are assigned individual roles. Overall, we try to create as many avenues for meeting the need for social belongingness while creating the learning path mainly around self-learning materials and individual assignments.

Design principle 4 - Enable controlled autonomy: This principle emphasizes the individual's need for autonomous decision-making, but also raises awareness of a conflict that exists between the need for competence and the need for autonomy (Ryan & Deci, 2000b). Namely, when there is no prior knowledge at all in a subject area, autonomy can be overwhelming, leading to a negative experience of competence that may trump the benefits of autonomy for psychological well-being. Thus, we must initially curtail autonomy and provide a lot of structure, and then, as students become more competent and their learning progresses, successively allow for more autonomous decisions: controlled autonomy. For example, after students learn the basics of programming, we allow them to autonomously select a data set of interest, while still providing enough structure to the analysis process so that they are not overwhelmed. The post-course survey in 2020/2021 revealed that only 17% of students wished they had more involvement in the selection of the course content. This affirms our argument that creating room for autonomy with some elements of control helps inspire motivation for novices. In light of this, we create as much structure as possible, through clear learning objectives, a clear temporal organization of the semester (using the learning management system "Moodle"), recurring weekly procedures and tasks, fixed social structures, predefined choices for the tools to be used (Jupyter notebooks with Anaconda or CoLab) and a consistent set of different types of teaching materials. We encourage students to set and structure their own learning goals in learning diaries. They are also allowed to work on transfer and creative assignments to express their individual ideas. Moreover, with the blended learning approach introduced for 2020/2021, they are self-regulated in the completion of their assignments.

Design principle 5 - Design for adequate success: This principle emphasizes the importance of generating competence experiences in designing a motivating and psychologically satisfying learning experience. In practice, this is certainly the design principle we had the most challenges with implementing. Both courses 2019/2020 and 2020/2021 have shown that we regularly over-challenge students who start with no prior knowledge and sometimes under-challenge those with prior knowledge. For instance, one student in the post-course survey when asked how we could have supported his/ her learning stated that "I would have just liked smaller steps to comprehend what we were doing completely". In contrast, another student stated, "It would have helped me to have more positive competence experiences or feel less often incompetent if I had a lesson on object-oriented programming". Clearly, the first student sounded over-challenged while the second student sounded under-challenged. However, with the number of students we deal with, this may not be completely avoided.

Nevertheless, certain elements have proven to be beneficial, such as a formative assessment process consisting of several individual submissions that allow for regular feedback. The operational tension between available resources and the effort it takes to provide feedback must be taken into account. Therefore, for example, in contrast to the manual correction in 2019/2020, we rely on an automatic assessment of student problem sets (in Jupyter notebooks) thereafter. In addition, we also try to integrate non-graded submissions. The grading system that is established in formal education systems rarely produces positive experiences of competence. We, therefore, try to reduce the idea of performance embodied in grades by incorporating assignments that value effort and not peak performance. Either you fulfil minimum requirements and earn a certain number of points or you get no points at all. With this in mind, a data-centered group project is planned for the upcoming winter semester in which students can contribute to the success of the project in a variety of ways within different roles, providing a sense of accomplishment that does not necessarily have to do solely with performing data analysis. However, due to the necessity in the university system, points are accumulated and converted into a grade at the end of the semester. In addition, there are significantly more elements that allow for students' successes besides exam-related assessments. For instance, the self-study notebooks contain coding tasks which enhance the number of students' moments of competence. Also, gamification/problem solving tasks included in the instructors' sessions create shared moments of competence experience.

CONCLUSION

After having focused mainly on operational implementation and optimization in the design of the first two pilot phases of our interdisciplinary programming course, we are now bringing together our practical experiences and theory in the preparation of our third phase.

Instead of focusing on improving learning materials, our approach aims at reprogramming learners' formerly negative emotions towards programming by promoting psychological and emotional well-being. To do so, we transfer basic psychological needs theory (BPNT) from self-determination theory (SDT) to the context of data literacy education and operationalize it into five guiding principles for creating motivating contexts, especially with heterogeneous learning groups, namely: (i) cultivating a supportive atmosphere, (ii) cultivating a digital growth mindset, (iii) independent social learning, (iv) enabling controlled autonomy, and (v) designing for adequate successes. Although we by no means claim to have fully resolved our guiding question, our design principles can inform practice and serve as an impetus for further discussion.

While this theoretically underpinned case study promises to be of great value to practitioners, particularly due to the urgency and relevance of the overarching goal of turning young people into data literates, the validity of the design principles, as well as the gains from the measures derived for the third DATAx project phase, will need to be further evaluated in forthcoming papers. This paper forms the outline for transitioning the DATAx projects tutorials into a more clearly structured experimental setting that can produce highly relevant scientific findings in the future.

REFERENCES

Bain, G., & Barnes, I. (2014). Why is programming so hard to learn? *In Proceedings of the 2014 conference on Innovation & technology in computer science education* (pp. 356-356).

Carretero, S., Vuorikari, R., & Punie, Y. (2017). DigComp 2.1: The Digital Competence Framework for Citizens with eight proficiency levels and examples of use. *Joint Research Centre, No. JRC106281* (Seville site).

European Commission. (n.d.). Council recommendation on key competences for lifelong learning. Retrieved May 30, 2021, from: https://ec.europa.eu/education/education-in-the-eu/council-recommendation-on-key-competences-for-lifelong-learning_en

Gopalan, V., Bakar, J. A. A., Zulkifli, A. N., Alwi, A., & Mat, R. C. (2017). A review of the motivation theories in learning. In *AIP Conference Proceedings* (Vol. 1891, No. 1, p. 020043). AIP Publishing LLC.

Hattie, John. *Visible learning for teachers: Maximizing impact on learning*. Routledge, 2012.

Greenwood, R., & Hinings, C. R. (1988). Organizational design types, tracks and the dynamics of strategic change. *Organization studies*, *9*(3), 293-316.

Keller, J. M. (2008). First principles of motivation to learn and e3-learning. *Distance education*, *29*(2), 175-185.

Li, T., & Lynch, R. (2016). The relationship between motivation for learning and academic achievement among basic and advanced level students studying Chinese as a foreign language in years 3 to 6 at Ascot International School in Bangkok, Thailand. *Scholar: Human Sciences*, *8*(1), 1-1.

Murphy, L., & Thomas, L. (2008). Dangers of a fixed mindset: implications of self-theories research for computer science education. In *Proceedings of the 13th annual conference on Innovation and technology in computer science education* (pp. 271-275).

Ozcelik, E., Cagiltay, N. E., & Ozcelik, N. S. (2013). The effect of uncertainty on learning in game-like environments. *Computers & Education*, *67*, 12-20.

Ridsdale, C., Rothwell, J., Smit, M., Ali-Hassan, H., Bliemel, M., Irvine, D., ... & Wuetherick, B. (2015). Strategies and best practices for data literacy education: Knowledge synthesis report.

Ryan, R. M., & Deci, E. L. (2000a). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary educational psychology*, *25*(1), 54-67.

Ryan, R. M., & Deci, E. L. (2000b). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist*, *55*(1), 68.

Schüller, K., Busch, P., & Hindinger, C. (2019). Future skills: ein framework für data literacy. *Hochschulforum Digitalisierung*, *46*, 1-128.

Tohidi, H., & Jabbari, M. M. (2012). The effects of motivation in education. *Procedia-Social and Behavioral Sciences*, *31*, 820-824.

Van den Broeck, A., Ferris, D. L., Chang, C. H., & Rosen, C. C. (2016). A review of self-determination theory's basic psychological needs at work. *Journal of Management*, *42*(5), 1195-1229.

Vansteenkiste, M., Ryan, R. M., Soenens, B. (2020). Basic psychological need theory: Advancements, critical themes, and future directions. *Motivation and Emotion*, *44* ,1-31.

Wenger, E. (2010). Conceptual tools for CoPs as social learning systems: Boundaries, identity, trajectories and participation. In *Social learning systems and communities of practice* (pp. 125-143). Springer, London.

Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational psychologist*, *47*(4), 302-314.