CLASSIFICATION tasks as basis for assessment – Family resemblance as principle for choosing mathematical objects

Norbert Noster[1], Arnon Hershkovitz[2], Michal Tabach[2] and Hans-Stefan Siller[1]

[1]University of Würzburg, Germany; norbert.noster@uni-wuerzburg.de, hans-stefan.siller@ uni-wuerzburg.de

[2]Tel Aviv University, Israel; arnonhe@tauex.tau.ac.il, tabachm@tauex.tau.ac.il

Formative assessment is aimed at supporting learning processes. Classification tasks can be used as a basis for formative assessment, particularly for evaluating students' understanding of definitions and object properties. Here, we present six design principles for choosing mathematical objects in such tasks—across grade-levels and across the curriculum—using literature-based ideas about categories of objects, family resemblance, hierarchical structures of categories, and intuitiveness. We demonstrate how we used these principles for designing a digital classification task, and discuss further areas in which these design principles should be considered.

Keywords: Classification, Assessment, Design principles, Mathematics Education

Introduction

Learning processes in school are evaluated on a regular basis. Such evaluations that take place throughout a learning process are generally referred to formative assessment. Amongst other possible purposes it can be used to evaluate how well a learning process is going and to identify room for improvement. It can be served as aid for the teacher to plan further instructions to address issues that have been identified (Cizek, 2009). Before actions like these can take place, it is necessary to create events that serve as basis for assessment. Developing tasks that can be used for the purpose of formative assessment can be challenging, especially when trying to identify weaknesses in learners understanding of a given topic. Classification tasks have been found useful for assessing understanding (Vollrath, 1977). Utilizing these tasks, this contribution takes a conceptual and theoretical perspective aiming at providing design principles for choosing mathematical objects to be used in classification tasks for formative assessment. In this sense we use it for the purpose of testing for students' conceptions, that can serve as a tool and aid for teachers to plan following instruction. The principles focus on choosing objects for the classification task derived from ideas of classifications and relations between categories. These principles are then used to analyze an example of a digital classification task as well as findings related to that task.

Classification tasks

Classifying (also referred to as categorizing or sorting) means grouping objects that can be treated equivalent regarding a certain criterion (e.g. any object that can be used to sit on can be classified as a chair) (Rosch, 1978). It is a process that helps us structure our surroundings, for example by differentiating between edible and non-edible or living and non-living things (Richler & Palmeri, 2014; Rosch, 1978). It can be considered part or result of a learning process. Therefore, it is not surprising that it can be found in mathematics curricula all around the world (Mullis et al., 2016). Using *classification tasks*, that ask learners to classify given objects provide us with insight into their knowledge, which is why they are suitable for the purpose of assessment. Here we concur with Vollrath (1977), who asserts the importance of classification tasks by stating that they "can make the students conscious of the characterizing properties of the concept, guide them to a definition and control their understanding of the definition" (p. 212).

Depending on the choice of objects to be classified and the choice of a classification criterion, certain features can be highlighted, and student identification of these features may be assessed. Once objects are chosen and a classification criterion is set, there are still three main options for the design of a classification tasks (Vollrath, 1978): 1) Tasks in which the objects are not classified, and the classification criterion is known, e.g., "Given is a set of quadrilaterals; classify each of them based on the existence of a line of symmetry (Has / Does Not Have)"; 2) Tasks in which the objects are classified, and the classification criterion is unknown, e.g., "Given are two groups of quadrilaterals; find the property that all the objects in Group A has and all the objects in Group B do not have"; 3) Tasks in which the objects are not classified, and the classification criterion is unknown, e.g., "Given is a set of quadrilaterals; find a property and divide the objects into two groups so that all the objects in Group A have this property and all the objects in Group B do not have it".

Different aims can be associated with different types of classification tasks, such as assessing understanding of a property, initializing development of conceptual understanding, and providing an overview of the relationship between objects and properties (Vollrath, 1977). While all three variants can be used for assessing understanding up to an extent, the first one can be used to find out to what extent a property is being identified in a set of given objects. The second one focusses more on deriving a definition from the classified objects, whereas the third variant can lead to different classifications that are not associated with the learning goal. Our focus is on the first variant as it focusses on applying a definition (of the classification criterion) rather than identifying it. Furthermore, this kind of task could be designed digitally with correct and incorrect classifications being automatically identified, so individual assessment could be automated, and immediate feedback to students could be provided (Feldt-Caesar, 2017).

For designing classification tasks, the suiting set of objects and classification criteria need to be well defined, so that they will serve the assessment purposes. This yields our first two design principles.

Design Principle 1: Define the classification criterion along with its values. This criterion should be in line with the curriculum and is strongly related to the assessment requirements. Note that classification can be done to two groups or more; for example, we can ask students to classify angles by acuteness (Yes / No) or by their type (Acute, Right, Obtuse, Straight, Reflex). Another example: For a given set of quadrilaterals, we can define a classification criterion "Has (Property of) Reflective Symmetry" (Yes / No), "Has (Property of) Rotational Symmetry" (Yes / No), or "Has This Property of Symmetry" (Reflective / Rotational / None).

Desing Principle 2: Define the set of objects to be classified. This should be in line with the curriculum and with the assessment requirements. In many cases, it may be easy to rely on pre-defined mathematical sets of objects, e.g., polygons, two-dimensional geometrical shapes, simple fractions, integers, etc. A choice should be taken at this point whether each object would belong to a single classification group. For example, if we chose the classification criterion “Has This Property of Symmetry” (Reflective / Rotational / None), and we focus on quadrilaterals, we should decide whether we want to include rectangles, which can be classified to both groups, or not.

Once classification criterion and the types of objects are chosen, we move on to choosing the specific objects that will be used in the task. For this, we present the notion of family resemblance.

Relations between categories of objects as the basis for family resemblance

Classifying objects can be complex as it can be done based on various characteristics (Pothos et al., 2011). In this context, one key term is *category*, which is defined as a group of objects that are considered to be equivalent with respect to a criterion (Rosch, 1978). Importantly, as different classification criteria could be set for the very same set of objects, categories are not given a priori. For example, in the case of classifying quadrilaterals based on the existence of a line of symmetry (Has / Does Not Have), all the rectangles would be under a single category and all the parallelograms would be under single, different category; however, if the classification criterion would be color (Red / Blue), a red rectangle and a red parallelogram would be under a single category while a blue rectangle and a blue parallelogram would be under a single, different category. Even if we limit ourselves to considering only mathematical properties, different classification criteria may yield different categories. For example, think of the set of polynomials and the following objects: x; $x+1$; x^2, x^2+1; classifying them by the criterion “Is a Quadric Polynomial” (Yes / No) will yield the following categories: $\{x^2, x^2+1\}$ (Yes), $\{x, x+1\}$ (No), while setting the classification criterion “Is a Monomial” (Yes / No) will yield different categories: $\{x, x^2\}$ (Yes), $\{x+1, x^2+1\}$ (No). Objects of the same category share family resemblance (Rosch & Mervis, 1975) vis-à-vis the classification criterion. It is important to state that family resemblance refers to the extent to which an objects shares features (including irrelevant ones) with other objects of a category. From here, we derive a third design principle:

Design Principle 3: Identify categories of objects for each classification group. These categories should be in line with the curriculum and with the assessment requirements. For example, if we chose to classify based on parity of functions (Even Function / Not Even Function) and we are focused on the polynomials, possible categories for the “Even Function” group could be: monomials with even exponent, or quadric polynomials of the form $ax^2 + c$ $(with\ c \neq 0)$; categories for the “Not Even Function” could be: monomials with odd exponent, or linear polynomials.

In each of these groups, categories can be organized into a hierarchical taxonomy in which the classification criterion is inherited from a broader category to its sub-categories (Bernabeu et al., 2022; Rosch & Mervis, 1975). Constructing this hierarchy will help in identifying family resemblance, hence our fourth design principle:

Design Principle 4: Construct a hierarchical taxonomy of categories for each classification group. Think of a task for classifying functions based on parity that was presented in the previous paragraph. The two categories for the “Even Function” group could be seen as stemming from a higher-level category of polynomials with only even exponents to which another category could belong: constant polynomials.

Once the taxonomy for each classification group is set up, we can identify objects with different levels of family resemblance. The longer the path on the hierarchical structure from one category to another, the lower the family resemblance between objects of those categories. To assess understanding of a concept, it is important to include in the task objects of different characteristics, from which we derive another design principle:

Design Principle 5: For each classification group, choose objects of different levels of family resemblance.

So far, the choice of objects has been defined by their mathematical characteristics. The final step of choosing objects has to do with the students' point of view, specifically regarding their level of knowledge. For this, we regard the notion of intuition.

Intuition in mathematics education

Identifying an object as having a criterion could be done intuitively – that is, immediately, with confidence, without the need to justify this choice (Fischbein, 2002). For example, kindergarten children will identify a 3-sides polygon which has two equal sides and a third side parallel to a horizontal line as a triangle, while at the same time reject it from being a triangle if "it stands on its head" (Sinclair & Moss, 2012). Likewise, a circle would be intuitively rejected by kindergarten children from being a triangle, as it is an intuitive example for a different type of objects (Tsamir et al., 2008). Thus, an object, or a property of an object could be intuitively identified by students based on their prior experiences with these objects. In mathematics education textbooks, for example, geometrical figures are usually presented in an orientation that is parallel to horizontal and vertical lines. Hence such figures would be identified intuitively, while figures oriented differently would be less intuitive to identify (ibid). In other words, the intuitiveness of an object reflects on the level of difficulty it presents to students. For this, we yield the final design principle:

Design Principle 6: For each classification group, choose both intuitive and non-intuitive objects. Note that intuitiveness and non-intuitiveness of mathematical objects is highly sensitive to the types of objects and to the classification criterion in matter, so there cannot be general guidelines as how to design them.

Example of applying the design principles

In this section we have a look at a digital classification task that we have developed and studied (Hershkovitz et al., 2023; Noster, Hershkovitz, Siller, et al., 2022; Noster, Hershkovitz, Tabach, et al., 2022). We present it through the lens of the design principles stated above. Considering the curriculum for these grade levels in both Israel and Germany, where we studied using this task, we decided to design the task around the concept of symmetry, specifically reflective symmetry. Following Design Principle 1, we defined the classification criterion and values: Has at Least one Symmetry Line (Has / Has Not); hence this is a two-way classification task. Following Design Principle 2, we chose to focus on quadrilaterals, which are known mathematical objects for children at these ages.

In the context of quadrilaterals, there is a well-established categorization into, e.g., parallelograms, rectangles, squares, trapezoids, kites; these categories appeared in the textbooks of the populations we sampled, hence we based our design on them. Following Design Principle 3, we identified the following categories for the “Has” group: rectangles, squares, and kites; and the following groups for the “Has Not” group: Parallelograms, and Non-Isosceles Trapezoid.

Another well-established framework that was relevant to our choice of objects was the House of Quadrilaterals, which describes hierarchical relationships between different categories of quadrilaterals based on their definitions. This helped us in building taxonomies for the two classification groups. Following Design Principle 4, we constructed the following hierarchical taxonomy for the “Has” group: Squares are sub-category of Recatngles and Kites; Parallelograms and Non-Isosceles Trapezoids are sub-categories of general non-symmetrical quadrilaterals.

Now, following Design Principle 5, we chose objects to represent different levels of family resemblance. For the “Has” groups, we first chose a square and rectangle, which share high family resemblance to each other vis-à-vis reflective symmetry, and a kite, which share low family resemblance with both. For the “Has Not” group, we first chose three objects with low family resemblance between them: parallelogram, non-isosceles trapezoid, and a general non-symmetrical quadrilateral.

Finally, following Design Principle 6, we relied on a framework of intuitive and non-intuitive two-dimensional geometric objects (Tsamir et al., 2008), and added the notion that shapes with lines of symmetry that are either horizontal or vertical are more likely to be identified as symmetric than shapes with diagonal lines of symmetry (Götz & Gasteiger, 2022). Squares, rectangles, and kites are generally intuitively perceived as having a line of symmetry, while the general non-symmetry quadrilateral was built in a way that it would be intuitive to assume it had no lines of symmetry; parallelograms are non-intuitive as not having a line of symmetry, as children often mix this notion with rotational symmetry (which parallelogram do have). We added a tilted square, which is non-intuitive due to the diagonal lines of symmetry, and a rotated parallelogram, yet another non-intuitive objects, however, note that these two objects share high family resemblance with the square and parallelogram, respectively. See Figure 1.

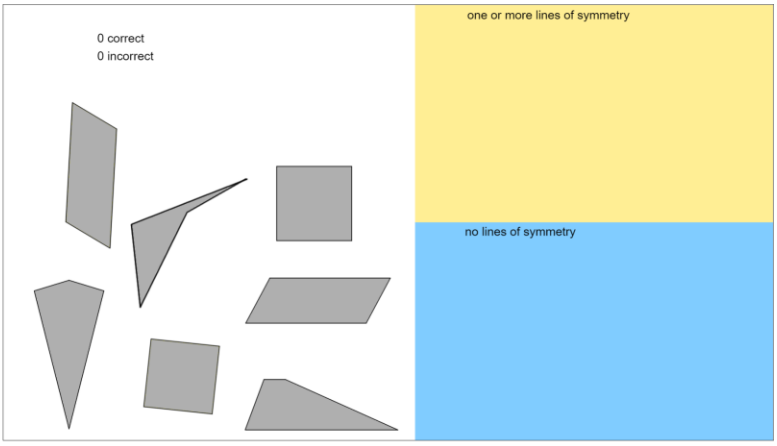

Figure 1: Examples for a digital classification task to differentiate between quadrilaterals with and without reflective symmetry

Discussion

This contribution aims at providing design principles for choosing mathematical objects to be used in classification tasks for formative assessment; these principles may be implemented across grade-levels and across topics. We focused particularly on classification tasks in which a set of objects is given along with a classification criterion, as this type of tasks provide a productive arena for assessing individual students' knowledge and can be easily assessed automatically. Based on the literature, we have identified the following design principles, that can guide the choice of the mathematical objects to be classified:

1. Define the classification criterion along with its values;
2. Define the set of objects to be classified;
3. Identify categories of objects for each classification group;
4. Construct a hierarchical taxonomy of categories for each classification group;
5. For each classification group, choose objects of different levels of family resemblance;
6. For each classification group, choose both intuitive and non-intuitive objects.

While explaining and demonstrating each principle as a "standalone", we also demonstrate the application of these principles together, to compose a classification task in a digital environment. Of course, these design principles should also be tested empirically. For example, the notion of intuitiveness of objects needs to be taken into field test, and cannot be based solely on the designers' assumptions (Noster, Hershkovitz, Tabach, et al., 2022). An iterative task design should follow, to verify the applicability of the design which resulted from applying the principles. Large data collection and its analysis could further inform the design choices.

As we are aiming at digital assessment tasks, there are also other aspects to be considered, to which we have not referred. These include for example the issue of feedback use. A digital environment could provide feedback of different types: simple vs. elaborated; immediate vs. delayed; or feedback on correctness vs. on strategies being implemented (Attali & van der Kleij, 2017; Shute, 2008; Tärning, 2018). To the best of our knowledge, there are no conclusive guidelines in the literature as for the most effective combination of these options for the purpose of serving assessment. This is an avenue for further research we are planning to pursue.

Another issue to be considered and tested relates to the layout of the objects on the screen. This may involve aspects like arrangement of the classification areas, e.g., unclassified objects are placed in the middle, in between the classification areas vs. to the left or right of the classification areas; arrangement of the objects, e.g., on a grid vs. randomly located; or issues related to size and color. We consider it an empirical question that needs to be tested based on large data collection, preferably in a set of randomized controlled studies, which again could lead to an iterative process of refining the design. To this we shall add more traditional design issues related to user interface, which are crucial in digital learning environments (Park & Song, 2015; Sagrario & Simbulan, 2007).

While the considerations presented here provide principles for choosing objects for classification tasks, it should be stated that this is only the first step in task design, as depending on the concrete task the effects may vary. Task design should be iterative (Liljedahl et al., 2007), but our principles serve as a foundation for a first predictive analysis, which has proven to be useful in the reflection of the data.

Acknowledgment

We thank the Bavarian Research Alliance and their Funding Programme for the Initiation of International Projects (BayIntAn) as well as their Scholarship Program for Bavarian and Israeli Early Career Researchers for supporting this project.

References

Attali, Y., & van der Kleij, F. (2017). Effects of feedback elaboration and feedback timing during computer-based practice in mathematics problem solving. *Computers & Education, 110*, 154–169. https://doi.org/10.1016/J.COMPEDU.2017.03.012

Bernabeu, M., Moreno, M., & Llinares, S. (2022). Preservice primary teachers' curricular reasoning when anticipating primary students' answers to geometrical figure classification tasks. *Proceedings of the Twelfth Congress of the European Society for Research in Mathematics Education*. https://hal.science/hal-03751484

Cizek, G. J. (2009). An introduction to formate assessment: History, characteristics, and challenges. In H. Andrade & G. J. Cizek (Eds.), *Handbook of formative assessment* (pp. 3–17). Routledge. https://doi.org/10.4324/9780203874851-6

Feldt-Caesar, N. (2017). Konzeptualisierung und Diagnose von mathematischem Grundwissen und Grundkönnen. *Konzeptualisierung Und Diagnose von Mathematischem Grundwissen Und Grundkönnen*. https://doi.org/10.1007/978-3-658-17373-9

Fischbein, E. (2002). *Intuition in Science and Mathematics: An educational approach*. Kluwer Academic Publishers. https://doi.org/10.1007/0-306-47237-6

Götz, D., & Gasteiger, H. (2022). Reflecting geometrical shapes: Approaches of primary students to reflection tasks and relations to typical error patterns. *Educational Studies in Mathematics*, *111*(1), 47–71. https://doi.org/10.1007/s10649-022-10145-5

Hershkovitz, A., Tabach, M., Noster, N., & Siller, H.-S. (2023). Student behavior while engaged with feedback-enhanced digital sorting tasks. In M. Ayalon, B. Koichu, R. Leikin, L. Rubel, & M. Tabach (Eds.), *Proceedings of the 46th Conference of the International Group for the Psychology of Mathematics Education (Volume 3)* (pp. 51–28). PME.

Liljedahl, P., Chernoff, E., & Zazkis, R. (2007). Interweaving mathematics and pedagogy in task design: A tale of one task. *Journal of Mathematics Teacher Education*, *10*(4–6), 239–249. https://doi.org/10.1007/s10857-007-9047-7

Mullis, I. V. S., Martin, M. O., Goh, S., & Cotter, K. (Eds.) (2016). TIMSS 2015 Encyclopedia: Education Policy and Curriculum in Mathematics and Science.

Noster, N., Hershkovitz, A., Siller, Hans.-S., & Tabach, M. (2022). Students’ strategies for identifying reflective symmetry of extra-mathematical shapes in a digital environment. *ERME Topic Conference on Mathematics Education in the Digital Age*.

Noster, N., Hershkovitz, A., Tabach, M., & Siller, H.-S. (2022). Learners’ strategies in interactive sorting tasks. In I. Hilliger, P. J. Muñoz-Merino, T. De Laet, & A. F. T. Ortega-Arranz (Eds.), *Educating for a New Future: Making Sense of Technology-Enhanced Learning Adoption. EC-TEL 2022. Lecture Notes in Computer Science, vol 13450* (pp. 285–298). Springer.

Park, H., & Song, H. D. (2015). Make e-learning effortless! Impact of a redesigned user interface on usability through the application of an affordance design approach. *Educational Technology & Society*, *18*(3), 185–196.

Pothos, E. M., Perlman, A., Bailey, T. M., Kurtz, K., Edwards, D. J., Hines, P., & McDonnell, J. V. (2011). Measuring category intuitiveness in unconstrained categorization tasks. *Cognition*, *121*(1), 83–100. https://doi.org/10.1016/j.cognition.2011.06.002

Richler, J. J., & Palmeri, T. J. (2014). Visual category learning. *WIREs Cognitive Science*, *5*(1), 75–94. https://doi.org/10.1002/wcs.1268

Rosch, E. (1978). Principles of Categorization. In E. Rosch & B. B. Lloyw (Eds.), *Cognition and Categorization* (pp. 27–48). Lawrence Erlbaum.

Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, *7*(4), 573–605. https://doi.org/10.1016/0010-0285(75)90024-9

Sagrario, M., & Simbulan, R. (2007). Learning objects’ user interface. In A. Koohang & K. Harman (Eds.), *Learning objects: Theory, praxis, issues, and trends* (pp. 259–336). Informing Science Press.

Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research*, *78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Sinclair, N., & Moss, J. (2012). The more it changes, the more it becomes the same: The development of the routine of shape identification in dynamic geometry environment. *International Journal of Educational Research*, *51–52*, 28–44. https://doi.org/10.1016/j.ijer.2011.12.009

Tärning, B. (2018). Review of feedback in digital applications - Does the feedback they provide support learning? *Journal of Information Technology Education: Research*, *17*, 247–283. https://doi.org/10.28945/4104

Tsamir, P., Tirosh, D., & Levenson, E. (2008). Intuitive nonexamples: The case of triangles. *Educational Studies in Mathematics*, *69*(2), 81–95. https://doi.org/10.1007/s10649-008-9133-5

Vollrath, H. J. (1977). The understanding of similarity and shape in classifying tasks. *Educational Studies in Mathematics 1977 8:2*, *8*(2), 211–224. https://doi.org/10.1007/BF00241026

Vollrath, H. J. (1978). Klassifikation nach Ähnlichkeit. *Der Mathematikunterricht* , *24*(2), 105–115.